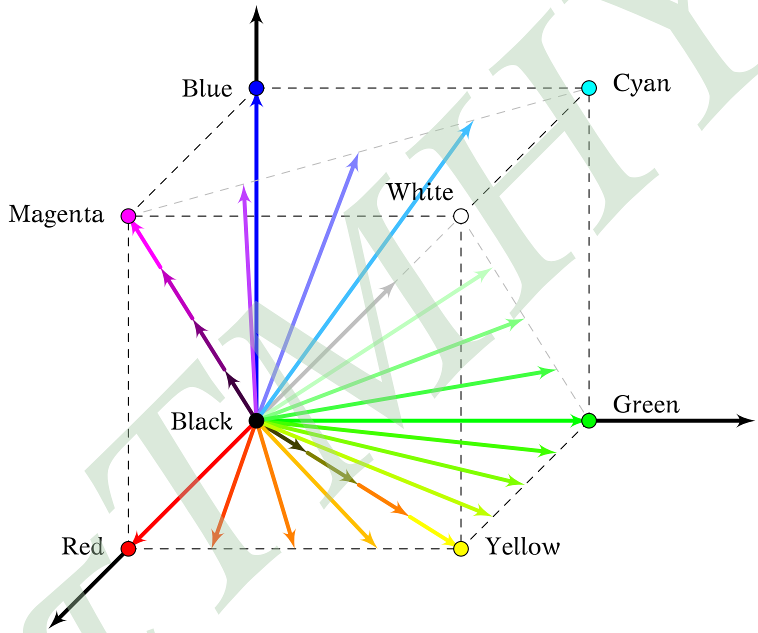
Γραμμική Άλγεβρα
μια Εισαγωγή

Σημειώσεις (σε εξέλιξη – V7)



Ε. Στεφανόπουλος

Τμήμα Μηχανικών Ηλεκτρονικών Υπολογιστών & Πληροφορικής
Πανεπιστήμιο Πατρών

Χριστούγεννα 2023

Περιεχόμενα

Κεφάλαιο 1

Διανύσματα

Εισαγωγή

Για τη Γραμμική Άλγεβρα τα διανύσματα είναι απλά τα σημεία ενός χώρου στον οποίο έχουν ορισθεί πράξεις με συγκεκριμένες ιδιότητες έτσι ώστε το αποτέλεσμα μιας τέτοιας πράξης είναι επίσης σημείο αυτού του χώρου. Το χώρο αυτόν θα τον ορίσουμε ως **Διανυσματικό χώρο**. Θα δούμε αργότερα ότι διάφορες κατηγορίες συναρτήσεων είναι σημεία τέτοιων χώρων, είναι δηλαδή διανύσματα. Για να φτάσουμε εκεί ας ξεκινήσουμε μελετώντας με συστηματικό τρόπο, ώστε να μπορέσουμε να γενικεύσουμε αργότερα, τα γνωστά μας διανύσματα που έχουμε συναντήσει κυρίως στη Φυσική, τα οποία αποτελούν τον πλέον φυσιολογικό Διανυσματικό χώρο.

1.1 Τα γνωστά διανύσματα

Η δύναμη του ανέμου που δρα σε ένα σώμα, για παράδειγμα νοτιοανανατολικός, ασθενής εντάσεως 3 BF, ή η ταχύτητα ενός κινητού είναι τυπικές φυσικές οντότητες οι οποίες για να περιγραφούν απαιτούνται περισσότερα δεδομένα απ΄ ότι ένας απλός αριθμός, που θα αρκούσε για παράδειγμα να περιγράψει τη θερμοκρασία ή τη μάζα ενός σώματος. Τα μεγέθη αυτά τα λέμε **διανυσματικά** και τα αντικείμενα που τα περιγράφουν τα λέμε **διανύσματα**. Τα διανυσματικά μεγέθη, όπως γνωρίζουμε από τη Φυσική, αποτυπώνονται γραφικά, στο επίπεδο ή στο χώρο, με ένα βέλος με αρχή τη θέση του σώματος που εφαρμόζονται, με συγκεκριμένη κατεύθυνση (κατεύθυνση του ανέμου) και με συγκεκριμένο μήκος (την ένταση του ανέμου). Μεγέθη τα οποία για να περιγραφούν αρκεί ένας μόνο αριθμός, όπως για παράδειγμα η θερμοκρασία, η μάζα, ο χρόνος, η ισχύς, λέγονται **βαθμωτά**.

Στο εξής ένα διάνυσμα, για παράδειγμα στο $\mathbb{R}^3$, θα παρίσταται με ένα βέλος με αρχή το σημείο $(0,0,0)$, την αρχή των αξόνων, και πέρας ένα σημείο (b_1, b_2, b_3), ταυτίζοντας έτσι όλα τα παράλληλα διανύσματα ίσου μήκους με αυτό η αρχή του οποίου είναι η αρχή των αξόνων. Τις συντεταγμένες του πέρατος τις λέμε **συνιστώσες** του διανύσματος.

Συμφωνούμε να συμβολίζουμε τα διανύσματα στο $\mathbb{R}^2$, $\mathbb{R}^3$ ή γενικότερα στο $\mathbb{R}^n$ με παχειά γράμματα $\boldsymbol{a}$, $\boldsymbol{u}$, $\boldsymbol{x}$, και να γράφουμε τις συνιστώσες κάθε διανύσματος ως στήλη, έτσι έχουμε

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \qquad \boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}.$$

Κάθε τέτοιο διάνυσμα μπορεί να γραφεί, εφαρμόζοντας την πράξη της αναστροφής, την οποία θα συζητήσουμε αργότερα, σε οριζόντια μορφή ως

$$\boldsymbol{a} = \begin{pmatrix} a_1 & a_2 \end{pmatrix}^{\mathrm{T}}, \qquad \boldsymbol{b} = \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix}^{\mathrm{T}}, \qquad \boldsymbol{x} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}^{\mathrm{T}}.$$

Σχόλιο 1.1. Παριστάνοντας ένα διάνυσμα του $\mathbb{R}^n$ με τις συνιστώσες του θα το γράφουμε ως στήλη. Όταν κάποιες φορές για λόγους οικονομίας χώρου θα γράφουμε ένα διάνυσμα σε οριζόντια μορφή, θα χρησιμοποιούμε πάντα τον συμβολισμό $\boldsymbol{u} = \begin{pmatrix} u_1 & u_2 & \cdots & u_n \end{pmatrix}^{\mathrm{T}}$, δίχως κόμματα. Για εμάς το $(u_1, u_2, \ldots, u_n)$ είναι απλά ένα σημείο στο $\mathbb{R}^n$.

1.2 Πράξεις μεταξύ διανυσμάτων

Εάν $\boldsymbol{a}$ και $\boldsymbol{b}$ είναι διανύσματα στο $\mathbb{R}^n$ και λ ένας πραγματικός αριθμός, στη συνέχεια θα λέμε μια πραγματική σταθερά ή απλά μια σταθερά, ορίζουμε τις πράξεις **πρόσθεση** και **πολλαπλασιασμό με σταθερά** σύμφωνα με τους νόμους

$$\boldsymbol{a} + \boldsymbol{b} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} + \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 + b_1 \\ a_2 + b_2 \\ \vdots \\ a_n + b_n \end{pmatrix}, \qquad \lambda \boldsymbol{a} = \lambda \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \begin{pmatrix} \lambda a_1 \\ \lambda a_2 \\ \vdots \\ \lambda a_n \end{pmatrix}. \tag{1.1}$$

Η πράξη της αφαίρεσης, προφανώς ανάλογη της πρόσθεσης, εκφράζεται ως συνδυασμός της πρόσθεσης και του πολλαπλασιασμού με σταθερά, όπως και στους πραγματικούς αριθμούς, μέσω της ισότητας

$$\boldsymbol{a} - \boldsymbol{b} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} - \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = \begin{pmatrix} a_1 - b_1 \\ a_2 - b_2 \\ \vdots \\ a_n - b_n \end{pmatrix} = \boldsymbol{a} + (-1)\boldsymbol{b} \tag{1.2}$$

Ορισμός 1.1. Εάν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα στο $\mathbb{R}^n$

(1) Το διάνυσμα $\boldsymbol{u} + \boldsymbol{v}$ το λέμε **άθροισμα** των $\boldsymbol{u}$ και $\boldsymbol{v}$.

(2) Το διάνυσμα του οποίου όλες οι συνιστώσες είναι ίσες με 0 το λέμε **μηδενικό διάνυσμα** ή απλά μηδέν και το συμβολίζουμε με $\mathbf{0}$.

(3) Ορίζουμε το **αντίθετο διάνυσμα** του $\boldsymbol{u}$ να είναι το διάνυσμα $-\boldsymbol{u}$ για το οποίο $-\boldsymbol{u} + \boldsymbol{u} = \mathbf{0}$.

(4) Το διάνυσμα $\boldsymbol{u} - \boldsymbol{v}$ λέμε **διαφορά** του $\boldsymbol{v}$ από το $\boldsymbol{u}$.

(5) Θα λέμε ότι τα διανύσματα $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι **ίσα** και θα γράφουμε $\boldsymbol{u} = \boldsymbol{v}$ αν και μόνον αν οι αντίστοιχες συνιστώσες των δύο διανυσμάτων είναι ίσες, δηλαδή $u_i = v_i$ για κάθε $i = 1, 2, \ldots, n$.

Παρατήρηση 1.1. Ως συνέπεια των πράξεων των πραγματικών αριθμών ισχύουν τα αποτελέσματα

(1) Αν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα στο $\mathbb{R}^n$, τότε $\boldsymbol{u} + \boldsymbol{v} = \boldsymbol{v} + \boldsymbol{u}$.

(2) Το μηδενικό διάνυσμα χαρακτηρίζεται μοναδικά από την ιδιότητα $\boldsymbol{u} + \boldsymbol{0} = \boldsymbol{u}$, όπου $\boldsymbol{u}$ είναι οποιοδήποτε διάνυσμα του $\mathbb{R}^n$, με την έννοια ότι αν για κάποιο διάνυσμα $\boldsymbol{z}$ του $\mathbb{R}^n$ ισχύει ότι $\boldsymbol{u} + \boldsymbol{z} = \boldsymbol{u}$, για κάθε $\boldsymbol{u}$ του $\mathbb{R}^n$, τότε και μόνο τότε όλες οι συνιστώσες του $\boldsymbol{z}$ είναι ίσες με μηδέν.

(3) Το αντίθετο διάνυσμα $-\boldsymbol{u}$ του $\boldsymbol{u}$ έχει συνιστώσες τις αντίθετες αντίστοιχες συνιστώσες αυτών του $\boldsymbol{u}$. Κατά συνέπεια το $-\boldsymbol{u}$ δεν είναι άλλο από το διάνυσμα $(-1)\boldsymbol{u}$.

(4) Αν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα στο $\mathbb{R}^n$, από τις ιδιότητες των πραγματικών αριθμών, ή από το (1) και την (1.2) έπεται ότι $\boldsymbol{u} - \boldsymbol{v} = -\boldsymbol{v} + \boldsymbol{u}$.

(5) Τα διανύσματα $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι ίσα αν και μόνον αν $\boldsymbol{u} - \boldsymbol{v} = \boldsymbol{0}$.

Θεώρημα 1.1 (Ιδιότητες των πράξεων). *Εάν $\boldsymbol{u}$, $\boldsymbol{v}$ και $\boldsymbol{w}$ είναι διανύσματα, και λ και μ είναι πραγματικές σταθερές, τότε ικανοποιούνται οι νόμοι*

(1) $\boldsymbol{u} + (\boldsymbol{v} + \boldsymbol{w}) = (\boldsymbol{u} + \boldsymbol{v}) + \boldsymbol{w}$

(2) $\lambda(\boldsymbol{u} + \boldsymbol{v}) = \lambda\boldsymbol{u} + \lambda\boldsymbol{v}$

(3) $(\lambda + \mu)\boldsymbol{u} = \lambda\boldsymbol{u} + \mu\boldsymbol{u}$

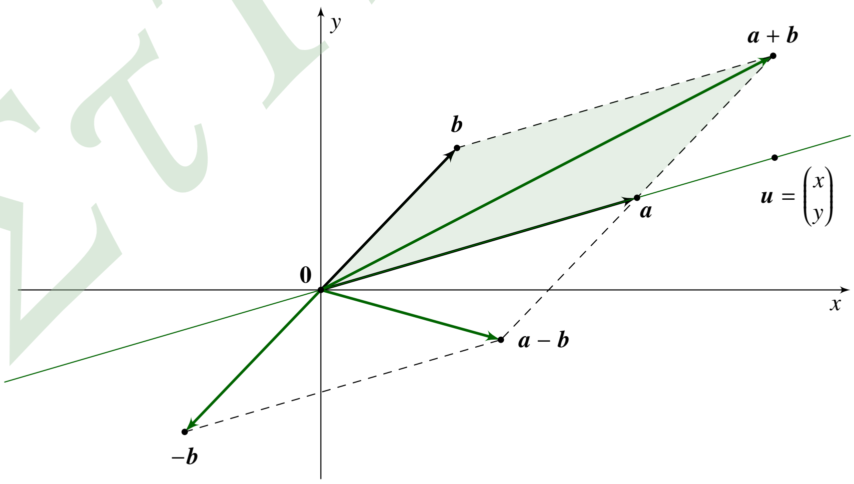
(4) $\lambda(\mu\boldsymbol{u}) = (\lambda\mu)\boldsymbol{u}$

(5) $1\boldsymbol{u} = \boldsymbol{u}$

(6) $0\boldsymbol{u} = \boldsymbol{0}$

Απόδειξη. Η απόδειξη αφήνεται σαν Άσκηση. □

Παρατήρηση 1.2 (Μηδείς αγεωμέτρητος εισίτω). Στο $\mathbb{R}^2$ αν δύο διανύσματα $\boldsymbol{a}$ και $\boldsymbol{b}$ δεν είναι το ένα πολλαπλάσιο του άλλου, τότε ορίζουν ένα (μοναδικό) παραλληλόγραμμο, βλέπε Σχήμα 1.1, και το άθροισμα $\boldsymbol{a} + \boldsymbol{b}$ αντιστοιχεί στη διαγώνιο του παραλληλογράμμου από την αρχή των αξόνων. Σε ανάλογο παραλληλόγραμμο εμφανίζεται και η διαφορά $\boldsymbol{a} - \boldsymbol{b}$. Θεωρώντας συνιστώσες στα διανύσματα $\boldsymbol{a}$ και $\boldsymbol{b}$ αποδείξτε ότι όντως αυτό συμβαίνει.



Σχήμα 1.1: Άθροισμα και διαφορά διανυσμάτων στο $\mathbb{R}^2$.

Αναφερόμενοι στο Σχήμα 1.1 ας υποθέσουμε ότι

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \qquad a_1 \neq 0.$$

Αν το σημείο (x, y) είναι τυχαίο σημείο της ευθείας που περιέχει το δάνυσμα $\boldsymbol{a}$, τότε $y = mx$, όπου m είναι η κλίση του τμήματος από το $(0, 0)$ στο (a_1, a_2), δηλαδή $m = a_2/a_1$. Έτσι αν για κάποιο $t \in \mathbb{R}$ είναι $x = ta_1$, τότε

$$\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} x \\ mx \end{pmatrix} = x\begin{pmatrix} 1 \\ m \end{pmatrix} = ta_1\begin{pmatrix} 1 \\ m \end{pmatrix} = t\begin{pmatrix} a_1 \\ ma_1 \end{pmatrix} = t\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} = t\boldsymbol{a}.$$

Κατά συνέπεια **η εξίσωση της ευθείας** η οποία περιέχει το διάνυσμα $\boldsymbol{a}$ εκφράζεται, σε διανυσματική μορφή, ως

$$\boldsymbol{r}(t) = t\boldsymbol{a}, \qquad t \in \mathbb{R}. \tag{1.3}$$

Η ίδια εξίσωση ισχύει και στην περίπτωση όπου $a_1 = 0$ και $a_2 \neq 0$, δηλαδή για $\boldsymbol{a} = \begin{pmatrix} 0 & a \end{pmatrix}^{\mathrm{T}}$ με $a \neq 0$ (γιατί;). Ανάλογα αποτελέσματα ισχύουν στο $\mathbb{R}^3$, αλλά και σε κάθε $\mathbb{R}^n$.

1.3 Γραμμικός συνδυασμός διανυσμάτων

Παρατήρηση 1.3. Παρατηρούμε ότι αν $\boldsymbol{u}$ είναι ένα διάνυσμα στο $\mathbb{R}^3$, από τις ιδιότητες των πράξεων έπεται ότι

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = \begin{pmatrix} u_1 \\ 0 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ u_2 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ 0 \\ u_3 \end{pmatrix} = u_1\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + u_2\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + u_3\begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

δηλαδή κάθε διάνυσμα στο $\mathbb{R}^3$ εκφράζεται ως συνδυασμός των διανυσμάτων

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Υπό αυτή την έννοια μπορούμε να πούμε ότι τα διανύσματα $\boldsymbol{e}_1$, $\boldsymbol{e}_2$ και $\boldsymbol{e}_3$ παράγουν τον τριδιάστατο χώρο. Γενικότερα αν στο $\mathbb{R}^n$ ορίσουμε το διάνυσμα $\boldsymbol{e}_k$ με 1 στην k γραμμή και 0 στις υπόλοιπες, $k = 1, 2, \ldots, n$, δηλαδή

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \qquad \cdots, \qquad \boldsymbol{e}_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}, \tag{1.4}$$

τότε κατ΄ αναλογία με το $\mathbb{R}^3$ κάθε διάνυσμα $\boldsymbol{u} = \begin{pmatrix} u_1 & u_2 & \ldots & u_n \end{pmatrix}^{\mathrm{T}} \in \mathbb{R}^3$ εκφράζεται ως ο ανάλογος συνδυασμός των $\boldsymbol{e}_k$, δηλαδή

$$\boldsymbol{u} = u_1\boldsymbol{e}_1 + u_2\boldsymbol{e}_2 + \cdots + u_n\boldsymbol{e}_n.$$

Ορισμός 1.2. Αν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα και α και β είναι σταθερές το διάνυσμα

$$\alpha\boldsymbol{u} + \beta\boldsymbol{v}$$

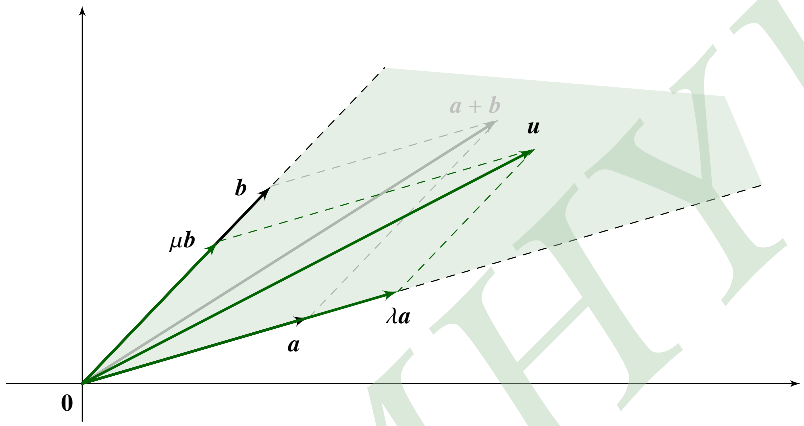
λέγεται **γραμμικός συνδυασμός** (linear combination) των $\boldsymbol{u}$ και $\boldsymbol{v}$. Γενικότερα αν

$\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$ είναι διανύσματα και $c_1, c_2, \ldots, c_k$ είναι σταθερές το διάνυσμα

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_k\boldsymbol{u}_k$$

λέγεται γραμμικός συνδυασμός των $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k$.

Παρατήρηση 1.4 (Η γεωμετρική ματιά). Ας υποθέσουμε ότι $\boldsymbol{a}$ και $\boldsymbol{b}$ είναι δύο διανύσματα στο $\mathbb{R}^2$ τα οποία δεν περιέχονται στην ίδια ευθεία. Ας υποθέσουμε επιπλέον ότι τα $\boldsymbol{a}$ και $\boldsymbol{b}$ βρίσκονται στο πρώτο τεταρτημόριο και ότι $\boldsymbol{u}$ είναι ένα τρίτο διάνυσμα "μεταξύ" των $\boldsymbol{a}$ και $\boldsymbol{b}$ όπως στο Σχήμα 1.2. Διαμορφώνοντας το παραλληλόγραμμο με διαγώνιο το $\boldsymbol{u}$ και ένα ζευγάρι



Σχήμα 1.2: Γραμμικός συνδυασμός διανυσμάτων στο $\mathbb{R}^2$.

πλευρών κατά μήκος των ευθειών που ορίζουν τα $\boldsymbol{a}$ και $\boldsymbol{b}$ συνάγεται το συμπέρασμα (γιατί;) ότι υπάρχουν θετικές σταθερές λ και μ ώστε

$$\boldsymbol{u} = \lambda\boldsymbol{a} + \mu\boldsymbol{b}.$$

Στη συγκεκριμένη περίπτωση $\lambda > 1$ και $0 < \mu < 1$. Γενικεύοντας συμπεραίνουμε ότι το σύνολο των διανυσμάτων

$$S = \{\lambda\boldsymbol{a} + \mu\boldsymbol{b} : \lambda \geq 0 \ \text{ και } \ \mu \geq 0\}$$

περιγράφει τον "κώνο" που ορίζουν τα διανύσματα $\boldsymbol{a}$ στο $\boldsymbol{b}$, δηλαδή το τμήμα του επιπέδου μεταξύ της γωνίας από το $\boldsymbol{a}$ στο $\boldsymbol{b}$ το οποίο περιέχει το $\boldsymbol{a} + \boldsymbol{b}$.

► **Άσκηση 1.1.** Εάν $\boldsymbol{a}$ και $\boldsymbol{b}$ είναι τα διανύσματα της Παρατήρησης 1.4 εξηγείστε τι περιγράφει καθένα από τα σύνολα

(α΄) $S_1 = \{\lambda\boldsymbol{a} + \mu\boldsymbol{b} : \lambda \in \mathbb{R} \ \text{ και } \ \mu \geq 0\}$.

(β΄) $S_2 = \{\lambda\boldsymbol{a} + \mu\boldsymbol{b} : \lambda \leq 0 \ \text{ και } \ \mu \leq 0\}$.

(γ΄) $S_3 = \{\boldsymbol{a} + \mu\boldsymbol{b} : \mu \in \mathbb{R}\}$.

(δ΄) $S_4 = \{\lambda\boldsymbol{a} + \mu\boldsymbol{b} : \lambda \in \mathbb{R} \ \text{ και } \ \mu \in \mathbb{R}\}$.

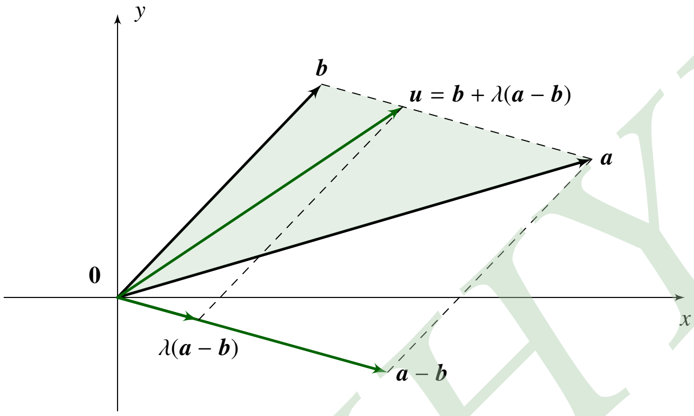
Παρατήρηση 1.5 (Κυρτός συνδυασμός). Ας υποθέσουμε ότι $\boldsymbol{a}$ και $\boldsymbol{b}$ είναι δύο διανύσματα στο $\mathbb{R}^2$ τα οποία δεν περιέχονται στην ίδια ευθεία, όπως ας πούμε στην Παρατήρηση 1.4, και ας θεωρήσουμε τον ιδιαίτερο γραμμικό συνδυασμό

$$\boldsymbol{u} = (1 - \lambda)\boldsymbol{b} + \lambda\boldsymbol{a}, \qquad 0 \leq \lambda \leq 1.$$

Παρατηρώντας ότι το $\boldsymbol{u}$ μπορεί να γραφεί ως

$$\boldsymbol{u} = \boldsymbol{b} + \lambda(\boldsymbol{a} - \boldsymbol{b}),$$

καταλαβαίνουμε ότι ότι το $\boldsymbol{u}$ είναι η διαγώνιος παραλληλογράμμου με πλευρές κατά μήκος των $\boldsymbol{b}$ και $\boldsymbol{a} - \boldsymbol{b}$, βλέπε Σχήμα 1.3. Κατά συνέπεια το $\boldsymbol{u}$ διαγράφει τη βάση του τριγώνου με "πλευρές" τα $\boldsymbol{b}$ και $\boldsymbol{a}$, καθώς το λ διαγράφει το διάστημα $[0, 1]$. Για $\lambda = 0$ είναι $\boldsymbol{u} = \boldsymbol{b}$, ενώ για $\lambda = 1$ είναι $\boldsymbol{u} = \boldsymbol{a}$. Το διάνυσμα $\boldsymbol{u}$ το λέμε **κυρτό συνδυασμό** (convex combination) των διανυσμάτων $\boldsymbol{a}$ και $\boldsymbol{b}$.



Σχήμα 1.3: Κυρτός συνδυασμός διανυσμάτων στο $\mathbb{R}^2$.

1.4 Μέτρο και εσωτερικό γινόμενο

Ορισμός 1.3. Εάν $\boldsymbol{u} = \begin{pmatrix} u_1 & u_2 & \dots & u_n \end{pmatrix}^{\mathrm{T}}$ είναι διάνυσμα στο $\mathbb{R}^n$ ορίζουμε το **μέτρο** του $\boldsymbol{u}$ να είναι ο μη αρνητικός αριθμός

$$\|\boldsymbol{u}\| := \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}.$$

Παρατήρηση 1.6. Παρατηρούμε ότι $\|\boldsymbol{u}\| = 0$ αν και μόνο αν $\boldsymbol{u} = \boldsymbol{0}$. Επίσης αν $\lambda \in \mathbb{R}$ και $\boldsymbol{u}$ είναι διάνυσμα στο $\mathbb{R}^n$, τότε

$$\begin{aligned} \|\lambda\boldsymbol{u}\| &= \sqrt{(\lambda u_1)^2 + (\lambda u_2)^2 + \cdots + (\lambda u_n)^2} \\ &= \sqrt{\lambda^2(u_1^2 + u_2^2 + \cdots + u_n^2)}, \end{aligned}$$

κατά συνέπεια, από τις ιδιότητες της ρίζας,

$$\|\lambda\boldsymbol{u}\| = |\lambda|\|\boldsymbol{u}\|, \tag{1.5}$$

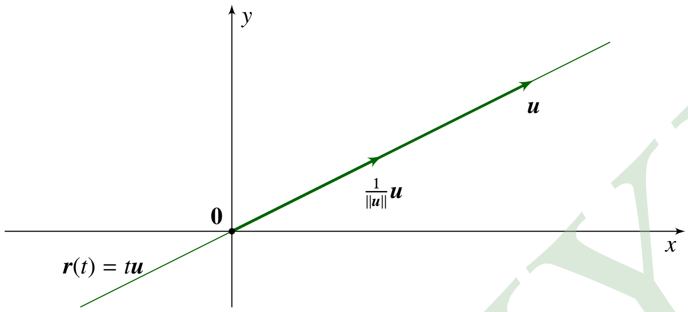
όπου $|\lambda|$ είναι η απόλυτη τιμή του λ.

Ορισμός 1.4. Ένα διάνυσμα $\boldsymbol{u}$ λέγεται **μοναδιαίο** αν $\|\boldsymbol{u}\| = 1$.

Παρατήρηση 1.7 (Μοναδιαία διανύσματα). Αν το $\boldsymbol{u}$ είναι μη μηδενικό διάνυσμα, τότε το διάνυσμα $(1/\|\boldsymbol{u}\|)\boldsymbol{u}$, περιέχεται στην ευθεία που παράγεται από το $\boldsymbol{u}$, βλέπε (1.3) και Σχήμα 1.4, και επιπλέον είναι μοναδιαίο, αφού μέσω της (1.5)

$$\left\| \frac{1}{\|\boldsymbol{u}\|}\boldsymbol{u} \right\| = \frac{1}{\|\boldsymbol{u}\|}\|\boldsymbol{u}\| = 1.$$

Το μοναδιαίο αυτό διάνυσμα το γράφουμε και ως $\boldsymbol{u}/\|\boldsymbol{u}\|$.



Σχήμα 1.4: Τα διανύσματα $\boldsymbol{u}$ και $\boldsymbol{u}/\|\boldsymbol{u}\|$ περιέχονται στην ευθεία $\boldsymbol{r}(t) = t\boldsymbol{u}$ στο $\mathbb{R}^2$.

Αν $\boldsymbol{a}$ και $\boldsymbol{b}$ είναι διανύσματα του $\mathbb{R}^3$ ως εσωτερικό γινόμενο $\boldsymbol{a} \cdot \boldsymbol{b}$ ορίζεται ο αριθμός

$$\boldsymbol{a} \cdot \boldsymbol{b} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} \cdot \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} = a_1 b_1 + a_2 b_2 + a_3 b_3.$$

Γενικά

Ορισμός 1.5. Αν

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \qquad \boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix},$$

είναι διανύσματα του $\mathbb{R}^n$ ορίζουμε το **εσωτερικό γινόμενο** των $\boldsymbol{u}$ και $\boldsymbol{v}$ με τη σχέση

$$\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n$$

Θεώρημα 1.2 (Ιδιότητες του εσωτερικού γινομένου). *Εάν $\boldsymbol{u}, \boldsymbol{v}$ και $\boldsymbol{w}$ είναι διανύσματα του $\mathbb{R}^n$ και λ είναι πραγματική σταθερά, τότε ισχύουν οι ιδιότητες*

(1) $\boldsymbol{u} \cdot \boldsymbol{u} \geq 0$ *και* $\boldsymbol{u} \cdot \boldsymbol{u} = 0$ *αν και μόνο αν* $\boldsymbol{u} = \boldsymbol{0}$.

(2) $(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}$.

(3) $(\lambda \boldsymbol{u}) \cdot \boldsymbol{v} = \lambda(\boldsymbol{u}, \boldsymbol{v})$.

(4) $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$.

Απόδειξη. Οι ιδιότητες είναι άμεσες συνέπειες των ιδιοτήτων των πράξεων των πραγματικών αριθμών και του ορισμού του εσωτερικού γινομένου. Πράγματι αν

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ \vdots \\ u_n \end{pmatrix}, \qquad \boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{pmatrix}, \qquad \boldsymbol{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{pmatrix},$$

και $\lambda \in \mathbb{R}$, τότε

(1) αφενός $\boldsymbol{u} \cdot \boldsymbol{u} = u_1^2 + u_2^2 + \cdots + u_n^2 \geq 0$, και αφετέρου

$$\boldsymbol{u} \cdot \boldsymbol{u} = 0 \Leftrightarrow u_1^2 + u_2^2 + \cdots + u_n^2 = 0 \Leftrightarrow u_1 = u_2 = \cdots = u_n = 0 \Leftrightarrow \boldsymbol{u} = \mathbf{0}.$$

(2) Από τις πράξεις των διανυσμάτων έπεται ότι

$$\begin{aligned}(\boldsymbol{u} + \boldsymbol{v}) \cdot \boldsymbol{w} &= (u_1 + v_1)w_1 + (u_2 + v_2)w_2 + \cdots + (u_n + v_n)w_n \\ &= (u_1 w_1 + u_2 w_2 + \cdots + u_n w_n) + (v_1 w_1 + v_2 w_2 + \cdots + v_n w_n) \\ &= \boldsymbol{u} \cdot \boldsymbol{w} + \boldsymbol{v} \cdot \boldsymbol{w}.\end{aligned}$$

(3) Όμοια από την πράξη του πολλαπλασιασμού στους πραγματικούς αριθμούς

$$\begin{aligned}(\lambda \boldsymbol{u}) \cdot \boldsymbol{v} &= (\lambda u_1)v_1 + (\lambda u_2)v_2 + \cdots + (\lambda u_n)v_n \\ &= \lambda(u_1 v_1 + u_2 v_2 + \cdots + u_n v_n) \\ &= \lambda(\boldsymbol{u} \cdot \boldsymbol{v}).\end{aligned}$$

(4) Το ότι $\boldsymbol{u} \cdot \boldsymbol{v} = \boldsymbol{v} \cdot \boldsymbol{u}$ είναι αποτέλεσμα της μεταθετικότητας του πολλαπλασιασμού στους πραγματικούς αριθμούς. □

▼ **Πόρισμα 1.1.** *Εάν $\boldsymbol{u}$, $\boldsymbol{v}$ και $\boldsymbol{w}$ είναι διανύσματα του $\mathbb{R}^n$ και λ είναι πραγματική σταθερά, τότε*

(1) $\boldsymbol{u} \cdot (\boldsymbol{v} + \boldsymbol{w}) = \boldsymbol{u} \cdot \boldsymbol{v} + \boldsymbol{u} \cdot \boldsymbol{w}$.

(2) $(\boldsymbol{u} - \boldsymbol{v}) \cdot \boldsymbol{w} = \boldsymbol{u} \cdot \boldsymbol{w} - \boldsymbol{v} \cdot \boldsymbol{w}$.

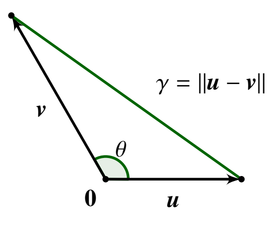
Απόδειξη. Η απόδειξη αφήνεται σαν Άσκηση. □

Παρατήρηση 1.8. Στο $\mathbb{R}^n$ το μέτρο ενός διανύσματος $\boldsymbol{u} = \begin{pmatrix} u_1 & u_2 & \ldots & u_n \end{pmatrix}^{\mathrm{T}}$ συνδέεται με το εσωτερικό γινόμενο μέσω της σχέσης

$$\|\boldsymbol{u}\| = \sqrt{\boldsymbol{u} \cdot \boldsymbol{u}}. \tag{1.6}$$

Αν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι δύο διανύσματα στο $\mathbb{R}^n$ που δεν περιέχονται στην ίδια ευθεία, τότε ορίζουν ένα τρίγωνο με μήκη πλευρών $\|\boldsymbol{u}\|$, $\|\boldsymbol{v}\|$ και $\|\boldsymbol{u} - \boldsymbol{v}\|$. Αν θ είναι η γωνία μεταξύ των διανυσμάτων $\boldsymbol{u}$ και $\boldsymbol{v}$, βλέπε Σχήμα 1.5, τότε $0 < \theta < \pi$ και από τον νόμο του συνημιτόνου έπεται ότι

$$\|\boldsymbol{u} - \boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - 2\|\boldsymbol{u}\|\|\boldsymbol{v}\| \cos\theta. \tag{1.7}$$



Σχήμα 1.5: Νόμος του συνημιτόνου: $\gamma^2 = \|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 - 2\|\boldsymbol{u}\|\|\boldsymbol{v}\| \cos\theta$.

Από την άλλη μεριά από την (1.6) και τις ιδιότητες του εσωτερικού γινομένου υπολογίζουμε

$$\begin{aligned} \|\boldsymbol{u}-\boldsymbol{v}\|^2 &= (\boldsymbol{u}-\boldsymbol{v})\cdot(\boldsymbol{u}-\boldsymbol{v}) \\ &= \boldsymbol{u}\cdot(\boldsymbol{u}-\boldsymbol{v})-\boldsymbol{v}\cdot(\boldsymbol{u}-\boldsymbol{v}) \\ &= \boldsymbol{u}\cdot\boldsymbol{u}-\boldsymbol{v}\cdot\boldsymbol{u}-\boldsymbol{u}\cdot\boldsymbol{v}+\boldsymbol{v}\cdot\boldsymbol{v} \\ &= \|\boldsymbol{u}\|^2-2\boldsymbol{u}\cdot\boldsymbol{v}+\|\boldsymbol{v}\|^2 \end{aligned} \tag{1.8}$$

Συγκρίνοντας τις (1.7) και (1.8) βρίσκουμε ότι

$$\boldsymbol{u}\cdot\boldsymbol{v} = \|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\theta.$$

Σημειώνουμε ότι αν $\boldsymbol{u}=\lambda\boldsymbol{v}$ η ισότητα είναι ισοδύναμη με την $\lambda=|\lambda|\cos\theta$ η οποία ισχύει, γιατί τότε για $\lambda>0$ είναι $\theta=0$, ενώ για $\lambda<0$ είναι $\theta=\pi$. Αποδείξαμε λοιπόν το

Θεώρημα 1.3. *Εάν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι δύο διανύσματα στο $\mathbb{R}^n$, και θ είναι η μεταξύ τους γωνία με $0\le\theta\le\pi$, τότε για το εσωτερικό γινόμενο ισχύει*

$$\boldsymbol{u}\cdot\boldsymbol{v} = \|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\theta. \tag{1.9}$$

Παρατήρηση 1.9 (Η τριγωνική ανισότητα). Για $\boldsymbol{u}$ και $\boldsymbol{v}$ στο $\mathbb{R}^n$, όπως στην (1.8), υπολογίζουμε

$$\begin{aligned} \|\boldsymbol{u}+\boldsymbol{v}\|^2 &= (\boldsymbol{u}+\boldsymbol{v})\cdot(\boldsymbol{u}+\boldsymbol{v}) \\ &= \|\boldsymbol{u}\|^2+2\boldsymbol{u}\cdot\boldsymbol{v}+\|\boldsymbol{v}\|^2 \\ &= \|\boldsymbol{u}\|^2+2\|\boldsymbol{u}\|\|\boldsymbol{v}\|\cos\theta+\|\boldsymbol{v}\|^2 \\ &\le \|\boldsymbol{u}\|^2+2\|\boldsymbol{u}\|\|\boldsymbol{v}\|+\|\boldsymbol{v}\|^2 \\ &= (\|\boldsymbol{u}\|+\|\boldsymbol{v}\|)^2 \end{aligned}$$

κατά συνέπεια, αφού οι εμπλεκόμενες ποσότητες είναι μη αρνητικές, τελικά παίρνουμε

$$\|\boldsymbol{u}+\boldsymbol{v}\| \le \|\boldsymbol{u}\|+\|\boldsymbol{v}\|. \tag{1.10}$$

Η ανισότητα (1.10) είναι η γνωστή τριγωνική ανισότητα, βλέπε Σχήμα 1.5.

Παρατήρηση 1.10. Σημειώνουμε ότι η σχέση (1.9) παρέχει πληροφορία για τη γεωμετρία των διανυσμάτων. Μας λέει, συγκεκριμένα, ότι δύο διανύσματα $\boldsymbol{u}$ και $\boldsymbol{v}$ στο $\mathbb{R}^n$ είναι **κάθετα** αν και μόνο αν $\boldsymbol{u}\cdot\boldsymbol{v}=0$. Στη περίπτωση όπου τα δύο διανύσματα $\boldsymbol{u}$ και $\boldsymbol{v}$ στο $\mathbb{R}^n$ είναι κάθετα ο νόμος του συνημιτόνου (1.7) εκφράζει το πυθαγόρειο θεώρημα. Θα επανέλθουμε και θα μελετήσουμε την έννοια της καθετότητας σε γενικότερο πλαίσιο σε επόμενο κεφάλαιο.

1.5 Μια εφαρμογή: Το χρωματικό μοντέλο rgb

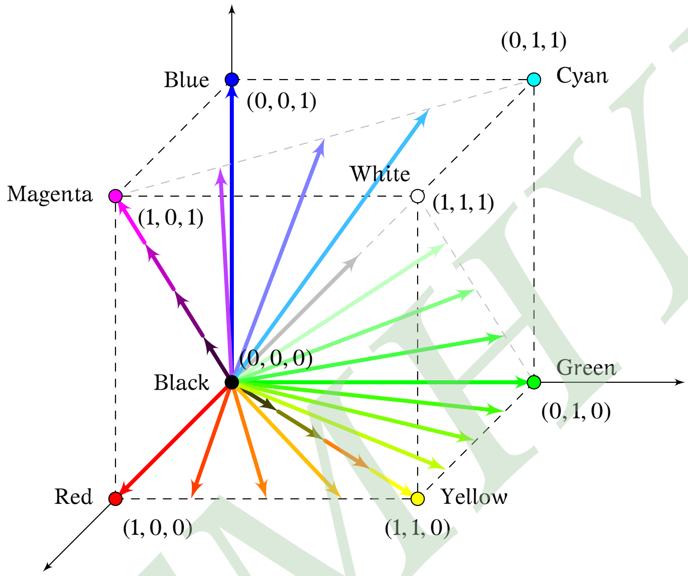
Ο σχηματισμός των χρωμάτων στην οθόνη ενός υπολογιστή στηρίζεται στο μοντέλο rgb. Το όνομα προέρχεται από τα βασικά, για το μοντέλο, χρώματα κόκκινο (**r**ed), πράσσινο (**g**reen) και μπλε (**b**lue). Το κόκκινο χρώμα αντιστοιχεί στο διάνυσμα $\boldsymbol{r}$, το πράσσινο στο $\boldsymbol{g}$ και το μπλε στο $\boldsymbol{b}$ όπου

$$\boldsymbol{r}=\begin{pmatrix}1\\0\\0\end{pmatrix},\qquad \boldsymbol{g}=\begin{pmatrix}0\\1\\0\end{pmatrix},\qquad \boldsymbol{b}=\begin{pmatrix}0\\0\\1\end{pmatrix},$$

βλέπε Σχήμα 1.6. Όλα τα υπόλοιπα χρώματα παράγονται αναμιγνύοντας κατάλληλες ποσότητες από τα τρία βασικά χρώματα, κατά συνέπεια κάθε χρώμα εκφράζεται ως γραμμικός συνδυασμός των $\boldsymbol{r}$, $\boldsymbol{g}$ και $\boldsymbol{b}$ με συντελεστές μεταξύ 0 και 1, ώστε καθένας από τους συντελεστές να εκφράζει το ποσοστό του αντίστοιχου βασικού χρώματος στο μίγμα. Έτσι το διάνυσμα

$$\boldsymbol{c} = c_1\boldsymbol{r} + c_2\boldsymbol{g} + c_3\boldsymbol{b} = \begin{pmatrix} c_1 & c_2 & c_3 \end{pmatrix}^{\mathrm{T}}$$

είναι το χρώμα που προκύπτει αναμιγνύοντας c_1 μέρη κόκκινου, c_2 μέρη πράσσινου, και c_3 μέρη μπλε. Τα διανύσματα $\boldsymbol{r}$, $\boldsymbol{g}$ και $\boldsymbol{b}$ ορίζουν με την έννοια της Παρατήρησης 1.3 τον **χρωματικό κύβο rgb**.



Σχήμα 1.6: Ο χρωματικός κύβος rgb.

Στις κορυφές του κύβου πέραν των τριών βασικών χρωμάτων, αντιστοιχούν τα χρώματα κίτρινο, φούξια ή βιολετί (magenta), κυανό, μαύρο στο $\begin{pmatrix} 0 & 0 & 0 \end{pmatrix}^{\mathrm{T}}$ και άσπρο στο $\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^{\mathrm{T}}$. Στο σχήμα 1.6 διακρίνονται τα διανύσματα κατά μήκος των διαγωνίων από το μαύρο στο φούξια και από το μαύρο στο κίτρινο αντίστοιχα

$$\lambda\boldsymbol{r} + \lambda\boldsymbol{b} = \begin{pmatrix} \lambda \\ 0 \\ \lambda \end{pmatrix}, \qquad \lambda\boldsymbol{r} + \lambda\boldsymbol{g} = \begin{pmatrix} \lambda \\ \lambda \\ 0 \end{pmatrix}, \qquad \text{για} \quad \lambda = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1.$$

Διακρίνονται επίσης διάφορα διανύσματα στη βάση του κύβου (στο κόκκινο-πρά;σινο επίπεδο), από το κόκκινο στο κίτρινο $\boldsymbol{y}$

$$(1-\lambda)\boldsymbol{r} + \lambda\boldsymbol{y} = (1-\lambda)\boldsymbol{r} + \lambda(\boldsymbol{r}+\boldsymbol{g}) = \boldsymbol{r} + \lambda\boldsymbol{g} = \begin{pmatrix} 1 \\ \lambda \\ 0 \end{pmatrix}, \qquad \text{για} \quad \lambda = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1$$

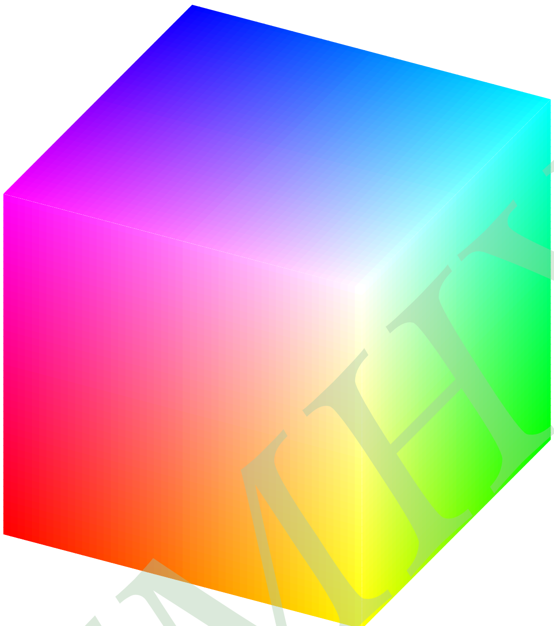
βλέπε Άσκηση 3, και από το πράσσινο στο κίτρινο

$$(1-\lambda)\boldsymbol{g} + \lambda\boldsymbol{y} = (1-\lambda)\boldsymbol{g} + \lambda(\boldsymbol{r}+\boldsymbol{g}) = \lambda\boldsymbol{r} + \boldsymbol{g} = \begin{pmatrix} \lambda \\ 1 \\ 0 \end{pmatrix}, \qquad \text{για} \quad \lambda = \frac{1}{4}, \frac{1}{2}, \frac{3}{4}, 1.$$

Έχουν επίσης σχεδιαστεί διανύσματα κατά μήκος των διαγωνίων των εδρών "φούξια-κυανό" και "πράσσινο-λευκό". Τα διανύσματα της μορφής

$$\boldsymbol{a} = a\boldsymbol{r} + a\boldsymbol{g} + a\boldsymbol{b} = \begin{pmatrix} a & a & a \end{pmatrix}^{\mathrm{T}} \qquad \text{με} \quad 0 < a < 1,$$

δηλαδή αυτά κατά μήκος της διαγωνίου από το μαύρο στο άσπρο, αντιστοιχούν στις, τόσες όσοι οι πραγματικοί αριθμοί στο διάστημα $(0, 1)$, αποχρώσεις του γκρι. Απεικονίζεται ένα με $a = 0.75$. Το Σχήμα 1.7 παρουσιάζει τμήμα του συνόρου του χρωματικού κύβου[1].



Σχήμα 1.7: Τρεις έδρες του χρωματικού κύβου rgb.

Σχόλιο 1.2. Το διάνυσμα μπορεί να αποτελέσει το ιδανικό, προς το παρόν τουλάχιστον, μέσο καταγραφής πληροφορίας ιδιαίτερα σε περιπτώσεις όπου απαιτείται και "αλγεβρικού τύπου" διαχείρηση της πληροφορίας αυτής. Για τον λόγο αυτό η έννοια του διανύσματος, όπως και του διανυσμματικού χώρου που θα συναντήσουμε στην επόμενη παράγραφο, εμφανίζεται σε διάφορες περιοχές, όπως στα οικονομικά, στη βιολογία, στη ψυχολογία και άλλες. Για παράδειγμα μπορούμε να υποθέσουμε ότι η ημερήσια παραγωγή γαλακτοκομικών προϊόντων μιας επιχείρησης περιγράφεται με ένα διάνυσμα

$$\boldsymbol{\eta} = \begin{pmatrix} p_1 \\ p_2 \\ p_3 \\ p_4 \end{pmatrix} \qquad \text{όπου} \qquad \begin{aligned} p_1 &= \text{ποσότητα γάλακτος σε λίτρα} \\ p_2 &= \text{ποσότητα βουτύρου σε λίτρα} \\ p_3 &= \text{ποσότητα γιαουρτιού σε λίτρα} \\ p_4 &= \text{ποσότητα τυριού φέτας σε λίτρα.} \end{aligned}$$

Τότε προφανώς η παραγωγή για τον μήνα Μάιο εκφράζεται σαν ένα άθροισμα $\boldsymbol{\eta}_1 + \boldsymbol{\eta}_2 + \cdots + \boldsymbol{\eta}_{31}$. Εάν δε η παραγωγή των σχετικών προϊόντων μιας ημέρας είναι διπλάσια αυτής κάποιας άλλης ημέρας το γεγονός αυτό εκφράζεται ως $\boldsymbol{\eta}_i = 2\boldsymbol{\eta}_j$ για τα ανάλογα i, j.

[1]Το Σχήμα 1.7 είναι μια παραλλαγή αρχείου TEX του Tom Bombadil για τον χρωματικό κύβο. Το αρχικό αρχείο βρίσκεται στον δεσμό: https://tex.stackexchange.com/questions/283729/tikz-draw-a-rgb-cube.

1.6 Ασκήσεις

1. Να βρεθούν σταθερές c_1, c_2, c_3 ώστε

$$c_1 \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} + c_2 \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} + c_3 \begin{pmatrix} 3 \\ 2 \\ -1 \end{pmatrix} = \begin{pmatrix} 6 \\ 14 \\ -2 \end{pmatrix}$$

2. Δείξτε ότι δεν υπάρχουν σταθερές c_1, c_2, c_3 ώστε

$$c_1 \begin{pmatrix} 1 \\ 11 \\ -3 \end{pmatrix} + c_2 \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix} + c_3 \begin{pmatrix} 6 \\ 9 \\ -2 \end{pmatrix} = \begin{pmatrix} 4 \\ 5 \\ 0 \end{pmatrix}$$

   Ερώτηση: Τί "δεν πάει καλά" εδώ; Μια πιθανή εξήγηση απορρέει από το ότι

$$\begin{pmatrix} 1 \\ 11 \\ -3 \end{pmatrix} + \begin{pmatrix} 5 \\ -2 \\ 1 \end{pmatrix} = \begin{pmatrix} 6 \\ 9 \\ -2 \end{pmatrix}.$$

3. Θυμίζουμε ότι εάν $\boldsymbol{a}$ και $\boldsymbol{b}$ είναι διανύσματα στο $\mathbb{R}^n$ ο ιδιαίτερος γραμμικός συνδυασμός

$$(1-t)\boldsymbol{a} + t\boldsymbol{b}, \qquad 0 \le t \le 1$$

   λέγεται **κυρτός συνδυασμός** των $\boldsymbol{a}$ και $\boldsymbol{b}$, βλέπε Παρατήρηση 1.5.

   (α΄) Θεωρείστε τον κυρτό συνδυσμό

$$\boldsymbol{v}(t) = (1-t)\begin{pmatrix} 0 \\ 2 \end{pmatrix} + t\begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad 0 \le t \le 1.$$

   Δείξτε ότι καθώς το t διαγράφει το διάστημα $[0,1]$ το διάνυσμα $\boldsymbol{v}(t)$ διαγράφει το ευθύγραμμο τμήμα με άκρα τα σημεία $(0,2)$ και $(1,1)$[2].

   (β΄) Ελευθερώνοντας το t ώστε να πάρει οποιαδήποτε πραγματική τιμή δικαιολογήστε γιατί το $\boldsymbol{v}(t)$ παριστάνει την εξίσωση της ευθείας του επιπέδου η οποία περιέχει τα σημεία $(0,2)$ και $(1,1)$.

   (γ΄) Γράφοντας

$$\boldsymbol{v}(t) = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + t\left[\begin{pmatrix} 1 \\ 1 \end{pmatrix} - \begin{pmatrix} 0 \\ 2 \end{pmatrix}\right] = \begin{pmatrix} 0 \\ 2 \end{pmatrix} + t\begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

   και διαπιστώνοντας ότι η εν λόγω ευθεία είναι παράλληλη στο διάνυσμα $\begin{pmatrix} 1 & -1 \end{pmatrix}^{\mathrm{T}}$, γενικεύστε και δείξτε ότι η ευθεία του επιπέδου η οποία περιέχει το σημείο (a,b) και είναι παράλληλη στο διάνυσμα $\boldsymbol{u} = \begin{pmatrix} p & q \end{pmatrix}^{\mathrm{T}}$ δίνεται από τη σχέση

$$\boldsymbol{r}(t) = \begin{pmatrix} a \\ b \end{pmatrix} + t\begin{pmatrix} p \\ q \end{pmatrix}, \qquad -\infty < t < +\infty.$$

---

[2] Υπενθυμίζουμε ότι για κάθε διάστημα $[a,b]$ της ευθείας ισχύει

$$[a,b] = \{x : x = (1-t)a + tb \text{ για κάποιο } t \in [0,1]\}.$$

Τι θα λέγατε να το αποδεικνύατε. Το ανάλογο ισχύει και για τα διαστήματα $[a,b)$, ή $(a,b]$, ή (a,b).

4. Αν $a, b, \theta \in \mathbb{R}$ χρησιμοποιώντας το εσωτερικό γινόμενο δείξτε ότι
$$(a\cos\theta + b\sin\theta)^2 \le a^2 + b^2.$$

5. **Η ανισότητα Cauchy-Bunyakovsky-Schwarz.** Εάν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα στο $\mathbb{R}^n$, τότε
$$|\boldsymbol{u}\cdot\boldsymbol{v}| \le \|\boldsymbol{u}\|\|\boldsymbol{v}\|.$$
Ισότητα ισχύει στην ανισότητα αν και μόνο αν τα $\boldsymbol{u} = \lambda\boldsymbol{v}$ για κάποια σταθερά λ.

6. Εάν $a_1, a_2, \ldots, a_n \in \mathbb{R}$ χρησιμοποιώντας το εσωτερικό γινόμενο δείξτε ότι
$$(a_1 + a_2 + \cdots + a_n)^2 \le n(a_1^2 + a_2^2 + \cdots + a_n^2).$$

7. Εάν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι μη μηδενικά διανύσματα στο $\mathbb{R}^n$ τότε η μεταξύ τους γωνία θ, με $0 \le \theta \le \pi$, δίνεται από τη σχέση
$$\cos\theta = \frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|}.$$
   (α΄) Εάν x, y είναι πραγματικοί αριθμοί και
$$\boldsymbol{u} = \begin{pmatrix} x \\ y \\ -x-y \end{pmatrix}, \qquad \boldsymbol{v} = \begin{pmatrix} -x-y \\ x \\ y \end{pmatrix}, \qquad \boldsymbol{w} = \begin{pmatrix} y \\ -x-y \\ x \end{pmatrix},$$
   δείξτε ότι
$$\frac{\boldsymbol{u}\cdot\boldsymbol{v}}{\|\boldsymbol{u}\|\|\boldsymbol{v}\|} = \frac{\boldsymbol{v}\cdot\boldsymbol{w}}{\|\boldsymbol{v}\|\|\boldsymbol{w}\|} = \frac{\boldsymbol{w}\cdot\boldsymbol{u}}{\|\boldsymbol{w}\|\|\boldsymbol{u}\|} = -\frac{1}{2}.$$
   Εξηγήστε γιατί συμβαίνει αυτό. Παρατηρήστε ότι το άθροισμα των συνυστωσών του $\boldsymbol{u}$ είναι ίσο με μηδέν.

   (β΄) Παρατηρήστε ότι καθένα από τα διανύσματα $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ είναι κάθετο στο διάνυσμα $\boldsymbol{a} = \begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^{\mathrm{T}}$.

8. **Το πυθαγόρειο θεώρημα.** Εάν τα διανύσματα $\boldsymbol{u}$ και $\boldsymbol{v}$ του $\mathbb{R}^3$ είναι κάθετα, δείξτε ότι $\|\boldsymbol{u}\|^2 + \|\boldsymbol{v}\|^2 = \|\boldsymbol{u}-\boldsymbol{v}\|^2$. Παρατηρήστε ότι $\|\boldsymbol{u}-\boldsymbol{v}\|$ είναι το μήκος της υποτείνουσας του ορθογωνίου τριγώνου με κάθετες "πλευρές" τα $\boldsymbol{u}$ και $\boldsymbol{v}$.

9. Θεωρούμε τα διανύσματα $\boldsymbol{u}$, $\boldsymbol{v}$, και $\boldsymbol{w}$ στο $\mathbb{R}^3$. Δείξτε ότι

   (α΄) Εάν το $\boldsymbol{u}$ είναι κάθετο στα $\boldsymbol{v}$ και $\boldsymbol{w}$, τότε είναι κάθετο σε κάθε γραμμικό συνδυασμό $\lambda\boldsymbol{v} + \mu\boldsymbol{w}$, όπου λ και μ είναι πραγματικές σταθερές.

   (β΄) Εάν τα $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι μοναδιαία, κάθετα διανύσματα, τότε $\|\boldsymbol{u}-\boldsymbol{v}\| = \sqrt{2}$.

1.7 Παράρτημα Ι: Η εξίσωση του επιπέδου

Αν $\boldsymbol{a}$ και $\boldsymbol{b}$ είναι διανύσματα στο $\mathbb{R}^3$ τα οποία δεν περιέχονται στην ίδια ευθεία τότε, όπως γνωρίζουμε από τη γεωμετρία, ορίζουν μοναδικό επίπεδο στο $\mathbb{R}^3$, αυτό που ορίζεται από τα σημεία $(0,0,0)$, (a_1, a_2, a_3) και (b_1, b_2, b_3), όπου a_j και b_j, $j = 1, 2, 3$, είναι οι συνιστώσες των $\boldsymbol{a}$ και $\boldsymbol{b}$ αντίστοιχα. Είναι το ίδιο επίπεδο που ορίζουν οι ευθείες οι οποίες περιέχουν τα $\boldsymbol{a}$ και $\boldsymbol{b}$ αντίστοιχα. Βλέπε Άσκηση 1.1 (β΄).

Ο κυρτός συνδυασμός $(1-t)\boldsymbol{a}+t\boldsymbol{b}$ με $0 \le t \le 1$, καθώς το t διαγράφει το διάστημα $[0,1]$ περιγράφει το ευθύγραμμο τμήμα από το (a_1,a_2,a_3) στο (b_1,b_2,b_3), βλέπε Παρατήρηση 1.5. Ένας κυρτός συνδυασμός μπορεί να παρασταθεί και ως

$$\alpha\boldsymbol{a}+\beta\boldsymbol{b}, \qquad \alpha+\beta=1 \quad \alpha\ge 0, \quad \beta\ge 0.$$

Αν τώρα α,β είναι τυχαίοι θετικοί αριθμοί, τότε μπορούμε να γράψουμε

$$\alpha\boldsymbol{a}+\beta\boldsymbol{b}=(\alpha+\beta)\Big(\frac{\alpha}{\alpha+\beta}\boldsymbol{a}+\frac{\beta}{\alpha+\beta}\boldsymbol{b}\Big)$$

και το διάνυσμα στην παρένθεση είναι κυρτός συνδυασμός των $\boldsymbol{a}$ και $\boldsymbol{b}$, κατά συνέπεια το $\alpha\boldsymbol{a}+\beta\boldsymbol{b}$, καθώς τα α και β παίρνουν όλες τις δυνατές θετικές τιμές, “σαρώνει” τη γωνία μεταξύ των θετικών ημιευθειών δια των $\boldsymbol{a}$ και $\boldsymbol{b}$. Αν $\alpha<0$ και $\beta>0$, τότε μπορούμε να γράψουμε

$$\alpha\boldsymbol{a}+\beta\boldsymbol{b}=-\alpha(-\boldsymbol{a})+\beta\boldsymbol{b}$$

οπότε σύμφωνα με το προηγούμενο βήμα ο γραμμικός συνδυασμός “σαρώνει” τη γωνία μεταξύ των θετικών ημιευθειών δια των $\boldsymbol{b}$ και $-\boldsymbol{a}$. Θεωρώντας και τις υπόλοιπες περιπτώσεις (ποιές;) καταλήγουμε στο συμπέρασμα ότι ο συνδυασμός $\alpha\boldsymbol{a}+\beta\boldsymbol{b}$ με $\alpha,\beta\in\mathbb{R}$ περιγράφει το επίπεδο που ορίζουν τα διανύσματα $\boldsymbol{a}$ και $\boldsymbol{b}$. Σημειώνουμε ότι για $\alpha=\beta=0$ ο γραμμικός συνδυασμός είναι το διάνυσμα $\mathbf{0}$.

Οι ευθείες στο $\mathbb{R}^2$ που περιέχουν το $(0,0)$ περιγράφονται από αλγεβρικές εξίσωσεις της μορφής

$$ax+by=0, \tag{1.11}$$

με ένα τουλάχιστον από τα a,b διάφορο του μηδενός. Ισχυριζόμαστε ότι η αλγεβρική εξίσωση ενός επιπέδου του $\mathbb{R}^3$ που περιέχει το $(0,0,0)$ είναι της μορφής

$$ax+by+cz=0 \tag{1.12}$$

με ένα τουλάχιστον από τα a,b,c διάφορο του μηδενός, που είναι το ακριβές ανάλογο της εξίσωσης της ευθείας στο $\mathbb{R}^2$. Αποδεικνύουμε τον ισχυρισμό υποθέτοντας ότι $a\neq 0$. Μπορούμε επομένως να υποθέσουμε ότι $a=1$, οπότε έχουμε την εξίσωση

$$x+by+cz=0. \tag{1.13}$$

Λύνοντας παίρνουμε $x=-by-cz$. Για $z=0$, $y=1$ βρίσκουμε $x=-b$, ενώ για $z=1$, $y=0$ βρίσκουμε $x=-c$, έτσι τα σημεία $(-b,1,0)$ και $(-c,0,1)$ ικανοποιούν την εξίσωση (1.13) και μαζί με το $(0,0,0)$ ορίζουν τη θέση μοναδικού επιπέδου έστω W. Τα διανύσματα

$$\boldsymbol{u}=\begin{pmatrix}-b\\1\\0\end{pmatrix}, \qquad \boldsymbol{v}=\begin{pmatrix}-c\\0\\1\end{pmatrix}$$

περιέχονται στο W. Δείχνουμε ότι το επίπεδο που παράγουν τα $\boldsymbol{u},\boldsymbol{v}$ περιγράφεται από την εξίσωση (1.12). Αρκεί να δείξουμε ότι αν λ και μ είναι πραγματικοί αριθμοί οι συνιστώσες του

$$\lambda\boldsymbol{u}+\mu\boldsymbol{v}=\begin{pmatrix}-\lambda b\\\lambda\\0\end{pmatrix}+\begin{pmatrix}-\mu c\\0\\\mu\end{pmatrix}=\begin{pmatrix}-\lambda b-\mu c\\\lambda\\\mu\end{pmatrix}$$

ικανοποιούν την (1.13). Πράγματι

$$-\lambda b - \mu c + b\lambda + c\mu = 0,$$

κατά συνέπεια η (1.13) είναι η αλγεβρική εξίσωση επιπέδου.

Σημειώνουμε ότι τόσο η εξίσωση (1.11) όσο και η (1.12) μπορούν να ιδωθούν ως εσωτερικά γινόμενα κάθετων διανυσμάτων, στη μεν πρώτη περίπτωση του σταθερού διανύσματος

$$\boldsymbol{w} = \begin{pmatrix} a & b \end{pmatrix}^{\mathrm{T}}$$

με το τυχαίο διάνυσμα $\begin{pmatrix} x & y \end{pmatrix}^{\mathrm{T}}$ της ευθείας, στη δε δεύτερη περίπτωση του σταθερού διανύσματος

$$\boldsymbol{w} = \begin{pmatrix} a & b & c \end{pmatrix}^{\mathrm{T}}$$

με το τυχαίο διάνυσμα $\begin{pmatrix} x & y & z \end{pmatrix}^{\mathrm{T}}$ του επιπέδου. Το ερώτημα είναι πώς δοθέντων δύο διανυσμάτων $\boldsymbol{a}$ και $\boldsymbol{b}$, που παράγουν το επίπεδο, υπολογίζεται το διάνυσμα $\boldsymbol{w}$. Αν

$$\boldsymbol{a} = \begin{pmatrix} a_1 & a_2 & a_3 \end{pmatrix}^{\mathrm{T}}, \qquad \boldsymbol{b} = \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix}^{\mathrm{T}}$$

το ζητούμενο διάνυσμα $\boldsymbol{w} = \begin{pmatrix} w_1 & w_2 & w_3 \end{pmatrix}^{\mathrm{T}}$ είναι λύση του συστήματος

$$\begin{aligned} a_1w_1 + a_2w_2 + a_3w_3 &= 0 \\ b_1w_1 + b_2w_2 + a_3w_3 &= 0. \end{aligned}$$

Θα δούμε αργότερα ότι ένα τέτοιο κάθετο διάνυσμα, μια λύση του συστήματος, είναι το **εξωτερικό γινόμενο** των $\boldsymbol{a}$ και $\boldsymbol{b}$, το οποίο θα γράφουμε ως $\boldsymbol{w} = \boldsymbol{a} \times \boldsymbol{b}$.

Σε σχέση με την Άσκηση 7 παρατηρούμε ότι τα διανύσματα $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ περιέχονται στο επίπεδο με κάθετο διάνυσμα το $\begin{pmatrix} 1 & 1 & 1 \end{pmatrix}^{\mathrm{T}}$. Επειδή το πολύ δύο διανύσματα αρκούν για να περιγράψουν ένα επίπεδο θα πρέπει το ένα από τα τρία να είναι γραμμικός συνδυασμός των άλλων δύο. Το αποτέλεσμα αυτό προκύπτει από το γεγονός ότι $\boldsymbol{u} + \boldsymbol{v} + \boldsymbol{w} = \boldsymbol{0}$, οπότε πράγματι $\boldsymbol{w} = -\boldsymbol{u} - \boldsymbol{v}$.

1.8 Παράρτημα II: Οι μιγαδικοί αριθμοί

Οι μιγαδικοί αριθμοί εμφανίζονται φυσιολογικά κατά την διαδικασία επίλυσης εξισώσεων. Η επιθυμία κάθε εξίσωση να έχει λύση οδήγησε στην αναζήτηση ενός συστήματος αριθμών που κατά κάποια έννοια επεκτείνει τους πραγματικούς αριθμούς και είναι τέτοιο ώστε η εξίσωση $x^2 + 1 = 0$ να έχει λύση. Ένα τέτοιο σύστημα υπάρχει και είναι οι μιγαδικοί αριθμοί. Σε αυτό το σύστημα οι ζητούμενες λύσεις δεν θα μπορούσαν να είναι άλλες από τις

$$x = \sqrt{-1}, \qquad \text{και} \qquad x = -\sqrt{-1}.$$

ΚΑΤΑΣΚΕΥΗ. Στο σύνολο $\mathbb{R}^2 = \mathbb{R} \times \mathbb{R}$, όπου

$$(x_1, y_1) = (x_2, y_2) \Leftrightarrow x_1 = x_2 \quad \text{και} \quad y_1 = y_2,$$

με τη γνωστή πρόσθεση

$$(x_1, y_1) + (x_2, y_2) = (x_1 + x_2, y_1 + y_2) \tag{1.14}$$

ορίζουμε την πράξη του πολαπλασιασμού με τη σχέση

$$(x_1, y_1)(x_2, y_2) = (x_1x_2 - y_1y_2, x_1y_2 + x_2y_1). \tag{1.15}$$

Παρατηρούμε ότι για $(x, y) \in \mathbb{R}^2$

$$(x, y) + (0, 0) = (x, y) \tag{1.16}$$

$$(x, y) + (-x, -y) = (0, 0) \tag{1.17}$$

$$(x, y)(1, 0) = (x, y), \tag{1.18}$$

δηλαδή το $(0, 0)$ είναι το ουδέτερο στοιχείο της πρόσθεσης, το $(-x, -y)$ είναι το αντίθετο του (x, y), ενώ το $(1, 0)$ είναι ουδέτερο στοιχείο του πολλαπλασιασμού, η μονάδα. Εξετάζοντας εάν υπάρχει το αντίστροφο του (x, y), δηλαδή εκείνο το (x', y') για το οποίο

$$(x, y)(x', y') = (1, 0)$$

και παρατηρώντας ότι

$$(0, 0)(x', y') = (0, 0) \tag{1.19}$$

υποθέτουμε ότι $(x, y) \neq (0, 0)$. Εάν (a, b) είναι το αντίστροφο στοιχείο του (x, y), εάν αυτό υπάρχει, τότε θα πρέπει

$$(x, y)(a, b) = (xa - yb, xb + ya) = (1, 0).$$

Από την παραπάνω ισότητα στο $\mathbb{R}^2$ προκύπτουν οι σχέσεις $xa - yb = 1$ και $xb + ya = 0$. Λύνοντας αυτό το σύστημα βρίσκουμε

$$a = \frac{x}{x^2 + y^2}, \qquad b = \frac{-y}{x^2 + y^2}.$$

Οι αριθμοί a και b υπάρχουν, καθόσον $x^2 + y^2 > 0$ οποτεδήποτε $(x, y) \neq (0, 0)$, επομένως το αντίστροφο του (x, y) το οποίο συμβολίζουμε με $(x, y)^{-1}$ είναι το

$$(x, y)^{-1} = \left(\frac{x}{x^2 + y^2}, \frac{-y}{x^2 + y^2}\right). \tag{1.20}$$

Το σύνολο των σημείων $z = (x, y) \in \mathbb{R} \times \mathbb{R}$ εφοδιασμένο με τις πράξεις (1.14) και (1.15) συμβολίζουμε με $\mathbb{C}$ και τα στοιχεία του καλούμε **μιγαδικούς αριθμούς** (complex numbers). Είναι εύκολο να αποδειχθεί ότι στο $\mathbb{C}$ ισχύουν οι νόμοι

(C1) $z_1 + z_2 = z_2 + z_1$, για κάθε z_1, z_2 στο $\mathbb{C}$.

(C2) $(z_1 + z_2) + z_3 = z_1 + (z_2 + z_3)$, για κάθε z_1, z_2, z_3 στο $\mathbb{C}$.

(C3) Υπάρχει ο μοναδικός μιγαδικός αριθμός $\mathbf{0} = (0, 0)$, έτσι ώστε $z + \mathbf{0} = z$, για κάθε $z \in \mathbb{C}$.

(C4) Για κάθε $z \in \mathbb{C}$ υπάρχει μοναδικός μιγαδικός αριθμός $-z$, έτσι ώστε $z + (-z) = \mathbf{0}$.

(C5) $z_1 z_2 = z_2 z_1$, για κάθε z_1, z_2 στο $\mathbb{C}$.

(C6) $(z_1 z_2) z_3 = z_1 (z_2 z_3)$, για κάθε z_1, z_2, z_3 στο $\mathbb{C}$.

(C7) Υπάρχει ο μοναδικός μιγαδικός αριθμός $\mathbf{1} = (1, 0)$, έτσι ώστε $z \cdot \mathbf{1} = z$, για κάθε $z \in \mathbb{C}$.

(C8) Για κάθε $z \in \mathbb{C}$ με $z \neq 0$ υπάρχει μοναδικός μιγαδικός αριθμός z^{-1} έτσι ώστε $z \cdot z^{-1} = \mathbf{1}$.

(C9) $z_1(z_2 + z_3) = z_1 z_2 + z_1 z_3$, για κάθε z_1, z_2, z_3 στο $\mathbb{C}$.

Απόρροια των πράξεων (1.14) και (1.15) είναι ότι

$$(x, y) = (x, 0) + (0, y)$$
$$(0, 1)(y, 0) = (0, y)$$

έτσι κάθε μιγαδικός αριθμός μπορεί να γραφεί στη μορφή

$$(x, y) = (x, 0) + (0, 1)(y, 0). \tag{1.21}$$

Εάν x είναι ένας πραγματικός αριθμός, σημείο της ευθείας, μπορεί να ταυτοποιηθεί με το $(x, 0)$, σημείο του επιπέδου. Επιπλέον παρατηρούμε ότι

$$(x_1, 0) + (x_2, 0) = (x_1 + x_2, 0), \qquad (x_1, 0)(x_2, 0) = (x_1 x_2, 0),$$

δηλαδή το σύστημα των μιγαδικών αριθμών επεκτείνει κατά φυσιολογικό τρόπο τους πραγματικούς αριθμούς και υπό το πρίσμα της ταυτοποίησης $x \equiv (x, 0)$ μπορούμε να θεωρούμε ότι $\mathbb{R} \subset \mathbb{C}$. Στη συνέχεια θα γράφουμε 0 αντί για **0** και 1 αντί για **1**. Θέτοντας $i = (0, 1)$ σύμφωνα με την παραπάνω ταυτοποίηση η (1.21) γράφεται

$$(x, y) = x + iy. \tag{1.22}$$

Ο μιγαδικός αριθμός i λέγεται **φανταστική μονάδα** (imaginary unit) για λόγους που θα γίνουν κατανοητοί παρακάτω. Εάν $z = (x, y)$ είναι ένας μιγαδικός αριθμός από εδώ και στο εξής θα γράφουμε $z = x + iy$. Εάν $z_1 = x_1 + iy_1$ και $z_2 = x_2 + iy_2$ είναι μιγαδικοί αριθμοί τότε το άθροισμα $z_1 + z_2$ και το γινόμενο $z_1 z_2$ δίνονται, μέσω των (1.14) και (1.15), από τις σχέσεις

$$z_1 + z_2 = (x_1 + iy_1) + (x_2 + iy_2) = (x_1 + x_2) + i(y_1 + y_2) \tag{1.23}$$
$$z_1 z_2 = (x_1 + iy_1)(x_2 + iy_2) = (x_1 x_2 - y_1 y_2) + i(x_1 y_2 + x_2 y_1). \tag{1.24}$$

Όπως και στους πραγματικούς αριθμούς, επαγωγικά ορίζουμε $z^{n+1} = z^n z$, για κάθε φυσικό αριθμό n. Παρατηρούμε ότι

$$i^2 = (0, 1)(0, 1) = (-1, 0) = -1,$$

σύμφωνα με την ταυτοποίηση, γεγονός που δικαιολογεί την ονομασία φανταστική μονάδα. Επειδή $-i = (0, -1)$ θα είναι

$$(-i)^2 = (0, -1)(0, -1) = (-1, 0) = -1.$$

Βλέπουμε λοιπόν ότι $i^2 + 1 = 0$ και $(-i)^2 + 1 = 0$.

► **Άσκηση 1.2.** Εάν $z \in \mathbb{C}$ και $w \in \mathbb{C}$ να δειχθεί ότι: $z^2+1 = (z+i)(z-i)$ και $z^2+w^2 = (z+iw)(z-iw)$.

Παρατήρηση 1.11. Ας θεωρήσουμε τον μιγαδικό αριθμό $z = x + iy$. Από τον αντιμεταθετικό νόμο (νόμος C5) έχουμε $iy = yi$ οπότε ο μπορούμε να γράφουμε

$$z = x + iy, \qquad \text{ή} \qquad z = x + yi.$$

Επίσης από την μοναδικότητα του αντίθετου μιγαδικού αριθμού έπεται ότι

$$i(-y) = (-1)iy = -iy.$$

Έτσι από τις (1.22), (1.17) και (1.20) έπεται ότι οι $-z$ και z^{-1}, εφόσον $z \neq 0$, δίνονται αντίστοιχα από τις σχέσεις

$$-z = -x + i(-y) = -x - iy \tag{1.25}$$
$$z^{-1} = \frac{x}{x^2 + y^2} + i\frac{-y}{x^2 + y^2} = \frac{x}{x^2 + y^2} - i\frac{y}{x^2 + y^2} \tag{1.26}$$

Παρατήρηση 1.12. Έστω $z_1 = x_1 + iy_1$ και $z_2 = x_2 + iy_2$, τότε κάνοντας χρήση του νόμου C9 (επιμεριστική ιδιότητα του πολλαπλασιασμού ως προς την πρόσθεση) υπολογίζουμε

$$
\begin{aligned}
z_1z_2 &= (x_1 + iy_1)(x_2 + iy_2) \\
&= x_1(x_2 + iy_2) + iy_1(x_2 + iy_2) && \text{(νόμος C9)} \\
&= x_1x_2 + x_1iy_2 + iy_1x_2 + iy_1iy_2 && \text{(νόμος C9)} \\
&= x_1x_2 + ix_1y_2 + iy_1x_2 + i^2y_1y_2 && \text{(νόμος C5)} \\
&= x_1x_2 + ix_1y_2 + iy_1x_2 - y_1y_2 && (i^2 = -1) \\
&= (x_1x_2 - y_1y_2) + i(x_1y_2 + x_2y_1) && \text{(νόμος C9)}
\end{aligned}
$$

που είναι η (1.24). Ο πολλαπλασιασμός δηλαδή, μιγαδικών αριθμών μπορεί να εκτελεσθεί με χρήση της οικείας, από τους πραγματικούς αριθμούς, επιμεριστικής ιδιότητας.

Παρατήρηση 1.13. Εάν $z_1 = x_1 + iy_1$ και $z_2 = x_2 + iy_2$, είναι μιγαδικοί αριθμοί, όπως στους πραγματικούς αριθμούς, η αφαίρεση και το πηλίκο ορίζονται, αντίστοιχα, με τις σχέσεις

$$z_1 - z_2 = z_1 + (-z_2) = (x_1 + iy_1) + (-x_2 + i(-y_2)) = (x_1 - x_2) + i(y_1 - y_2) \tag{1.27}$$

$$\frac{z_1}{z_2} = z_1z_2^{-1} = (x_1 + iy_1)\left(\frac{x_2}{x_2^2 + y_2^2} + i\frac{-y_2}{x_2^2 + y_2^2}\right) = \frac{x_1x_2 + y_1y_2}{x_2^2 + y_2^2} + i\frac{-x_1y_2 + x_2y_1}{x_2^2 + y_2^2}. \tag{1.28}$$

Παρατηρούμε ότι για $z_1 = 1 = 1 + i0$ και $z_2 = z = x + iy$ από την τελευταία σχέση έπεται

$$\frac{1}{z} = \frac{x}{x^2 + y^2} + i\frac{-y}{x^2 + y^2} = z^{-1}. \tag{1.29}$$

Επακόλουθο της τελευταίας αυτής σχέσης είναι η

$$\frac{z_1}{z_2} = z_1\frac{1}{z_2}. \tag{1.30}$$

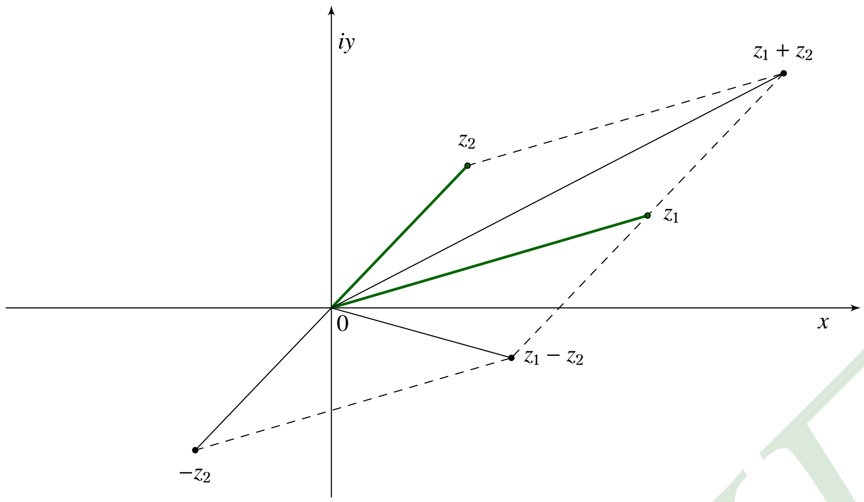
• Έστω ο μιγαδικός αριθμός $z = x + iy$, με $x \in \mathbb{R}$ και $y \in \mathbb{R}$. Ο x λέγεται **πραγματικό μέρος** (real part) του z και γράφουμε $x = \operatorname{Re} z$, και ο y λέγεται **φανταστικό μέρος** (imaginary part) του z και γράφουμε $y = \operatorname{Im} z$. Έτσι εάν $z \in \mathbb{R}$ τότε $\operatorname{Re} z = z$ και $\operatorname{Im} z = 0$, ενώ εάν $z = iy$, με $y \in \mathbb{R}$, τότε $\operatorname{Re} z = 0$ και $\operatorname{Im} z = -iz$.

• Οι μιγαδικοί αριθμοί $z_1 = x_1 + iy_1$ και $z_2 = x_2 + iy_2$ είναι **ίσοι** και γράφουμε $z_1 = z_2$, εάν και μόνον εάν $x_1 = x_2$ και $y_1 = y_2$, ισοδύναμα $\operatorname{Re} z_1 = \operatorname{Re} z_2$ και $\operatorname{Im} z_1 = \operatorname{Im} z_2$.

Δείξαμε λοιπόν ότι το σώμα των μιγαδικών αριθμών $\mathbb{C}$ αποτελεί μία φυσιολογική επέκταση των πραγματικών αριθμών, όπου στο σύστημα αυτό η εξίσωση $z^2 + 1 = 0$ έχει λύση.

1.8.1 Το μιγαδικό επίπεδο

Από τον ορισμό των μιγαδικών αριθμών έπεται ότι υπάρχει μια ένα προς ένα αντιστοιχία μεταξύ του μιγαδικού αριθμού $z = x + iy$ και του σημείου (x, y) του επιπέδου. Έτσι το επίπεδο του οποίου κάθε σημείο (x, y) ταυτίζουμε με τον μιγαδικό αριθμό $z = x+iy$ ονομάζουμε **μιγαδικό επίπεδο** (complex plane). Ο άξονας των x λέγεται **πραγματικός άξονας** (real axis), ενώ αυτός των y λέγεται **φανταστικός άξονας** (imaginary axis). Ισοδύναμα οι μιγαδικοί αριθμοί μπορούν να ιδοθούν ως διανύσματα του $\mathbb{R}^2$ στα οποία έχει ορισθεί, επιπλέον της πρόσθεσης η οποία αντιστοιχεί στη συνήθη πρόσθεση διανυσμάτων και του πολλαπλασιασμού με πραγματική σταθερά, μια πράξη πολλαπλασιασμού μεταξύ διανυσμάτων το αποτέλεσμα της οποίας είναι επίσης διάνυσμα του $\mathbb{R}^2$. Γεωμετρικά οι πράξεις της πρόσθεσης και της διαφοράς μιγαδικών αριθμών αποτυπώνεται στο Σχήμα 1.8.



Σχήμα 1.8: Το άθροισμα και η διαφορά μιγαδικών αριθμών.

Ορισμός 1.6. Έστω $z = x + iy$ ένας μιγαδικός αριθμός.

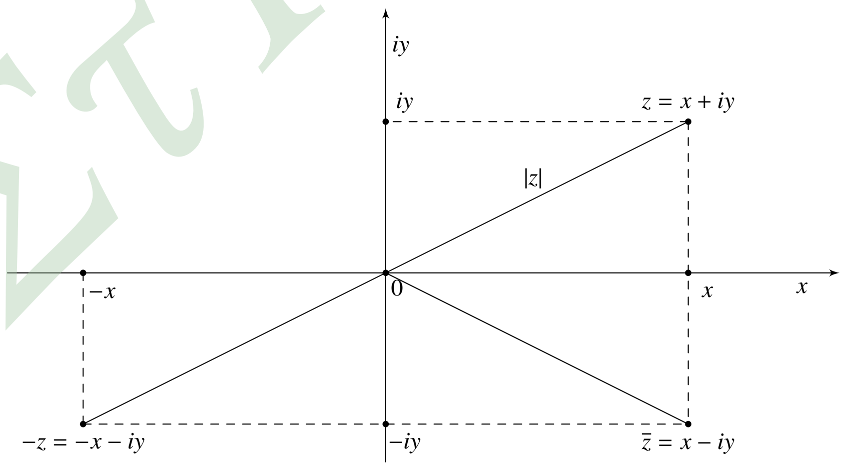
(1) **Το μέτρο** (modulus) του z, συμβολίζεται με $|z|$, ορίζεται να είναι ο πραγματικός αριθμός

$$|z| = \sqrt{x^2 + y^2}. \tag{1.31}$$

(2) **Ο συζυγής** (conjugate) του z, συμβολίζεται με $\bar{z}$, ορίζεται να είναι ο μιγαδικός αριθμός

$$\bar{z} = x - iy. \tag{1.32}$$

Το μέτρο $|z|$ του z αντιστοιχεί στο μέτρο διανύσματος, ενώ ο συζυγής $\bar{z}$ του z αντιστοιχεί στο συμμετρικό διάνυσμα του z ως προς τον πραγματικό άξονα.



Σχήμα 1.9: Γραφική απεικόνιση των μιγαδικών αριθμών z, $\bar{z}$ και $-z$.

Παρατηρούμε ότι αν $z \in \mathbb{R}$, ισοδύναμα $y = 0$, τότε $|z| = \sqrt{x^2} = |x|$, δηλαδή το μέτρο μιγαδικού αριθμού γενικεύει την απόλυτη τιμή πραγματικού αριθμού. Για το λόγο αυτό το μέτρο το λέμε και απόλυτη τιμή. Επιπλέον αν $z \in \mathbb{R}$, τότε $\bar{z} = z$.

Παράδειγμα 1.1. Να βρεθεί το μέτρο και ο συζυγής του μιγαδικού αριθμού $-i(2-i3)$.
Εάν $z = -i(2-i3)$, τότε $z = -i2 + 3i^2 = -3 - i2$, οπότε

$$
\begin{aligned}
|z| &= |-3-i2| = \sqrt{(-3)^2 + (-2)^2} = \sqrt{13} \\
\bar{z} &= \overline{-3-i2} = -3 + i2.
\end{aligned}
$$

Οι ιδιότητες του μέτρου και του συζυγούς μιγαδικού αριθμού συνοψίζονται στη

▲ **Πρόταση 1.1.** *Ισχύουν οι ιδιότητες:*

(1) $|z| \geq 0$, *για κάθε* $z \in \mathbb{C}$, *και* $|z| = 0$ *εάν και μόνον εάν* $z = 0$.

(2) $|z_1 z_2| = |z_1||z_2|$, *για κάθε ζευγάρι μιγαδικών αριθμών* z_1 *και* z_2.

(3) $|z_1/z_2| = |z_1|/|z_2|$, *για κάθε ζευγάρι μιγαδικών αριθμών* z_1 *και* z_2 *με* $z_2 \neq 0$.

(4) $z = \bar{z}$ *εάν και μόνον εάν* $z \in \mathbb{R}$.

(5) $z = \bar{\bar{z}}$, *για κάθε* $z \in \mathbb{C}$.

(6) $|z| = |\bar{z}|$, *για κάθε* $z \in \mathbb{C}$.

(7) $|z|^2 = z\bar{z}$, *για κάθε* $z \in \mathbb{C}$.

(8) $\overline{z_1 + z_2} = \bar{z}_1 + \bar{z}_2$, *για κάθε ζευγάρι μιγαδικών αριθμών* z_1 *και* z_2.

(9) $\overline{z_1 z_2} = \bar{z}_1 \bar{z}_2$, *για κάθε ζευγάρι μιγαδικών αριθμών* z_1 *και* z_2.

(10) $\overline{(z_1/z_2)} = \bar{z}_1/\bar{z}_2$, *για κάθε ζευγάρι μιγαδικών αριθμών* z_1 *και* z_2 *με* $z_2 \neq 0$.

Απόδειξη. Έστω $z = x + iy$, $z_1 = x_1 + iy_1$, και $z_2 = x_2 + iy_2$ να είναι μιγαδικοί αριθμοί.

(1) Επειδή $|z| = \sqrt{x^2 + y^2}$, είναι προφανές ότι $|z| \geq 0$, ενώ $|z| = 0 \Leftrightarrow x^2 + y^2 = 0 \Leftrightarrow x = 0$ και $y = 0 \Leftrightarrow z = 0$.

(2) Από την σχέση (1.24) έπεται ότι $|z_1 z_2| = |(x_1 x_2 - y_1 y_2) + i(x_1 y_1 + x_2 y_2)|$ έτσι έχουμε

$$
\begin{aligned}
|z_1 z_2| &= \sqrt{(x_1 x_2 - y_1 y_2)^2 + (x_1 y_1 + x_2 y_2)^2} = \sqrt{x_1^2 x_2^2 + y_1^2 y_2^2 + x_1^2 y_1^2 + x_2^2 y_2^2} \\
&= \sqrt{(x_1^2 + y_1^2)(x_2^2 + y_2^2)} = \sqrt{x_1^2 + y_1^2}\,\sqrt{x_2^2 + y_2^2} \\
&= |z_1||z_2|.
\end{aligned}
$$

(3) Από την (1.29) για $z \neq 0$ έχουμε

$$
1 = z\frac{1}{z} \Longrightarrow 1 = \left|z\frac{1}{z}\right| = |z|\left|\frac{1}{z}\right| \Longrightarrow \frac{1}{|z|} = \left|\frac{1}{z}\right|
$$

με χρήση της ιδιότητας 2, επομένως για $z_2 \neq 0$

$$
\left|\frac{z_1}{z_2}\right| = \left|z_1 \frac{1}{z_2}\right| = |z_1|\left|\frac{1}{z_2}\right| = |z_1|\frac{1}{|z_2|} = \frac{|z_1|}{|z_2|}.
$$

(4) Έστω $z = x + iy$, τότε $z = \bar{z} \Leftrightarrow x + iy = x - iy \Leftrightarrow 0 = i2y \Leftrightarrow y = 0 \Leftrightarrow z \in \mathbb{R}$.

(5) Έστω $z = x + iy$, τότε $\overline{\overline{z}} = \overline{x - iy} = x + iy = z$.

(6) Έστω $z = x + iy$, τότε $|\overline{z}| = |x - iy| = \sqrt{x^2 + (-y)^2} = \sqrt{x^2 + y^2} = |z|$.

(7) Εάν $z = x + iy$, έχουμε $z\overline{z} = (x + iy)(x - iy) = x^2 - ixy + ixy - i^2y^2 = x^2 + y^2 = |z|^2$.

(8) Από την (1.23) έπεται ότι $\overline{z_1 + z_2} = x_1 + x_2 - i(y_1 + y_2) = (x_1 - iy_1) + (x_2 - iy_2) = \overline{z_1} + \overline{z_2}$.

(9) Έχουμε $\overline{z}_1\overline{z}_2 = (x_1 - iy_1)(x_2 - iy_2) = x_1x_2 - y_1y_2 - i(x_1y_1 + x_2y_2) = \overline{z_1z_2}$, μιας και από την (1.24) $z_1z_2 = (x_1x_2 - y_1y_2) + i(x_1y_1 + x_2y_2)$.

(10) Από την (1.29) για $z \neq 0$ έχουμε

$$1 = z\frac{1}{z} \Longrightarrow 1 = \overline{1} = \overline{\left(z\frac{1}{z}\right)} = \overline{z}\overline{\left(\frac{1}{z}\right)} \Longrightarrow \frac{1}{\overline{z}} = \overline{\left(\frac{1}{z}\right)}$$

με χρήση της ιδιότητας 9, επομένως για $z_2 \neq 0$

$$\overline{\left(\frac{z_1}{z_2}\right)} = \overline{z_1\frac{1}{z_2}} = \overline{z}_1\overline{\left(\frac{1}{z_2}\right)} = \overline{z}_1\frac{1}{\overline{z}_2} = \frac{\overline{z}_1}{\overline{z}_2}.$$

Η απόδειξη είναι πλήρης. □

Παρατήρηση 1.14. Εάν $z = x + iy$ είναι ένας μιγαδικός αριθμός και $z + \overline{z} = x + iy + x - iy = 2x$, και $z - \overline{z} = x + iy - (x - iy) = i2y$, συμπεραίνουμε

$$\operatorname{Re} z = \frac{z + \overline{z}}{2}, \qquad \operatorname{Im} z = \frac{z - \overline{z}}{2i}. \tag{1.33}$$

Επίσης $x \leq |x| \leq \sqrt{x^2 + y^2}$, όμοια $y \leq |y| \leq \sqrt{x^2 + y^2}$, οπότε

$$\operatorname{Re} z \leq |\operatorname{Re} z| \leq |z|, \qquad \operatorname{Im} z \leq |\operatorname{Im} z| \leq |z|. \tag{1.34}$$

► **Άσκηση 1.3.** Να δειχθεί ότι ο αριθμός a είναι πραγματικός εάν και μόνον εάν $\operatorname{Re} a = a$.

Παρατήρηση 1.15. Εάν a και b είναι πραγματικοί αριθμοί θυμίζουμε τη γνωστή ιδιότητα της απόλυτης τιμής $|a + b| \leq |a| + |b|$. Το ίδιο ισχύει και για μιγαδικούς αριθμούς. Ας είναι z_1, και z_2 δύο μιγαδικοί αριθμοί. Τότε ισχύει η **τριγωνική ανισότητα**

$$|z_1 + z_2| \leq |z_1| + |z_2|. \tag{1.35}$$

Κάνοντας χρήση της Πρότασης 1.1 έχουμε

$$\begin{aligned}
|z_1 + z_2|^2 &= (z_1 + z_2)\overline{(z_1 + z_2)} = (z_1 + z_2)(\overline{z}_1 + \overline{z}_2) && (\text{ ιδιότητες 7 και 8 })\\
&= z_1\overline{z}_1 + z_1\overline{z}_2 + z_2\overline{z}_1 + z_2\overline{z}_2 &&\\
&= |z_1|^2 + z_1\overline{z}_2 + \overline{z_1\overline{z}_2} + |z_2|^2 && (\text{ ιδιότητες 7 και 5 })\\
&= |z_1|^2 + 2\operatorname{Re}(z_1\overline{z}_2) + |z_2|^2 && (\text{ σχέση (1.33) })\\
&\leq |z_1|^2 + 2|z_1\overline{z}_2| + |z_2|^2 && (\text{ σχέση (1.34) })\\
&= |z_1|^2 + 2|z_1||\overline{z}_2| + |z_2|^2 && (\text{ ιδιότητα 2 })\\
&= (|z_1| + |z_2|)^2 && (\text{ ιδιότητα 6 })
\end{aligned}$$

απ΄ όπου έπεται η ζητούμενη τριγωνική ανισότητα. Επιπλέον οι αριθμοί z_1 και z_2 ορίζουν ένα παραλληλόγραμο, βλέπε Σχήμα 1.8, στο οποίο τα μεγέθη $|z_1 + z_2|$ και $|z_1 - z_2|$ είναι τα μέτρα των διαγωνίων αυτού του παραλληλογράμου. Από την απόδειξη της τριγωνικής ανισότητας (1.35) προκύπτει ο γνωστός **ο νόμος του παραλληλογράμου**

$$|z_1 + z_2|^2 + |z_1 - z_2|^2 = 2(|z_1|^2 + |z_2|^2) \tag{1.36}$$

ο οποίος μας λέει ότι το άθροισμα των τετραγώνων των διαγωνίων παραλληλογράμου ισούται με το άθροίσμα των τετραγώνων των πλευρών του.

▶ **Άσκηση 1.4.** Εάν $z \in \mathbb{C}$ δείξτε ότι

$$|z| \leq |\operatorname{Re} z| + |\operatorname{Im} z| \leq \sqrt{2}|z|.$$

Παρατήρηση 1.16. Από τις ιδιότητες που περιγράφονται στη Πρόταση 1.1 έπεται ότι για $z \neq 0$

$$\frac{1}{z} = \frac{\bar{z}}{z\bar{z}} = \frac{\bar{z}}{|z|^2} \tag{1.37}$$

που είναι ακριβώς η σχέση (1.29). Επειδή $|i| = \sqrt{0^2 + 1^2} = 1$, έπεται αμέσως ότι

$$\frac{1}{i} = \frac{\bar{i}}{i\bar{i}} = \frac{-i}{|i|^2} = -i. \tag{1.38}$$

γενικότερα εάν $z \in \mathbb{C}$ και $|z| = 1$, από την (1.37) έπεται ότι $1/z = \bar{z}$.

Αν $z = x + iy \neq 0$, τότε $|z| > 0$, οπότε τα κλάσματα

$$\frac{x}{|z|} = \frac{x}{\sqrt{x^2 + y^2}} \qquad \text{και} \qquad \frac{y}{|z|} = \frac{y}{\sqrt{x^2 + y^2}}$$

ορίζονται και ικανοποιούν τη σχέση

$$\left(\frac{x}{|z|}\right)^2 + \left(\frac{y}{|z|}\right)^2 = \frac{x^2}{x^2 + y^2} + \frac{y^2}{x^2 + y^2} = 1,$$

κατά συνέπεια υπάρχει $\theta \in \mathbb{R}$ τέτοιο ώστε

$$\cos\theta = \frac{x}{|z|} \qquad \text{και} \qquad \sin\theta = \frac{y}{|z|}.$$

1.8.2 Τριγωνομετρική μορφή μιγαδικού αριθμού

Εάν r και θ είναι οι πολικές συντεταγμένες του σημείου $(x, y) \neq (0, 0)$, δηλαδή

$$x = r\cos\theta, \qquad y = r\sin\theta,$$

τότε ο μη μηδενικός μιγαδικός αριθμός $z = x + iy$ μπορεί να γραφεί ως $z = r\cos\theta + ir\sin\theta$.

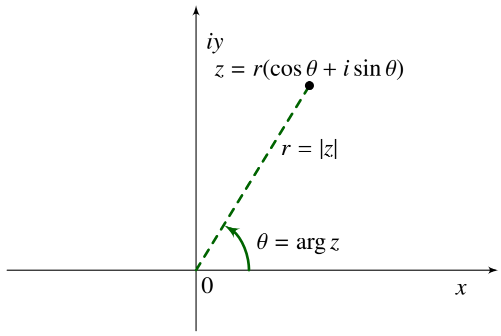
Παρατηρούμε ότι

$$|z| = \sqrt{(r\cos\theta)^2 + (r\sin\theta)^2} = r$$

ενώ το $\theta = \arg z$, είναι καταληπτό ως τη γωνία (σε ακτίνια) μεταξύ της πραγματικής θετικής ημιευθείας και του ευθυγράμμου τμήματος από το 0 στο z. Η έκφραση

$$z = r(\cos\theta + i\sin\theta) \tag{1.39}$$

λέγεται **τριγωνομετρική μορφή** (trigonometric form) ή **πολική μορφή** (polar form) του μιγαδικού αριθμού z.



Σχήμα 1.10: Τριγωνομετρική μορφή μιγαδικού αριθμού.

Παράδειγμα 1.2. Να γραφούν σε πολική μορφή οι αριθμοί $z_1 = 1 + i$, $z_2 = 1$, $z_3 = -2$.

Επειδή $|1 + i| = \sqrt{2}$, έχουμε

$$z_1 = 1 + i = \sqrt{2}\Big(\frac{1}{\sqrt{2}} + i\frac{1}{\sqrt{2}}\Big) = \sqrt{2}\Big(\cos\frac{\pi}{4} + i\sin\frac{\pi}{4}\Big)$$

$z_2 = 1 = 1 + i0 = \cos 0 + i \sin 0.$

$z_3 = -2 = 2(-1 + i0) = 2(\cos\pi + i\sin\pi).$

Για τους μιγαδικούς αριθμούς $z_1 = r_1(\cos\theta_1 + i\sin\theta_1)$ και $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$ παρατηρούμε ότι

$$\begin{aligned} z_1z_2 &= r_1r_2(\cos\theta_1 + i\sin\theta_1)(\cos\theta_2 + i\sin\theta_2) \\ &= r_1r_2[(\cos\theta_1\cos\theta_2 - \sin\theta_1\sin\theta_2) + i(\sin\theta_1\cos\theta_2 + \cos\theta_1\sin\theta_2)], \end{aligned}$$

οπότε η πολική μορφή του γινομένου z_1z_2 δίνεται από τη σχέση

$$z_1z_2 = r_1r_2[\cos(\theta_1 + \theta_2) + i\sin(\theta_1 + \theta_2)]. \tag{1.40}$$

Εάν $z = r(\cos\theta + i\sin\theta)$ είναι μη μηδενικός αριθμός, ισοδύναμα $r \neq 0$, τότε από τη σχέση (1.40) έπεται ότι

$$\frac{1}{z} = \frac{1}{r}[\cos(-\theta) + i\sin(-\theta)]. \tag{1.41}$$

Σημειώνουμε ότι η σχέση αυτή προκύπτει επίσης από την (1.37). Εάν τώρα $z_2 = r_2(\cos\theta_2 + i\sin\theta_2)$ είναι διάφορος του μηδενός, τότε συνδυάζοντας τις (1.40) και (1.41) έχουμε

$$\frac{z_1}{z_2} = \frac{r_1}{r_2}[\cos(\theta_1 - \theta_2) + i\sin(\theta_1 - \theta_2)]. \tag{1.42}$$

Παράδειγμα 1.3. Εάν $z_1 = 2\sqrt{3} - i2$ και $z_2 = -1 + i\sqrt{3}$ να γραφούν οι z_1 και z_2 σε πολική μορφή και να υπολογισθούν οι z_1z_2 και z_1/z_2.

Επειδή $|z_1| = |2\sqrt{3} - i2| = \sqrt{16} = 4$ και $|z_2| = |-1 + i\sqrt{3}| = \sqrt{4} = 2$ θα έχουμε αντίστοιχα

$$z_1 = 4\left[\frac{\sqrt{3}}{2} + i\left(-\frac{1}{2}\right)\right] = 4\left[\cos\left(-\frac{\pi}{6}\right) + i\left(-\frac{\pi}{6}\right)\right],$$
$$z_2 = 2\left[\left(-\frac{1}{2}\right) + i\frac{\sqrt{3}}{2}\right] = 2\left[\cos\frac{2\pi}{3} + i\sin\frac{2\pi}{3}\right].$$

Από την σχέση (1.40) υπολογίζουμε το γινόμενο

$$\begin{aligned} z_1 z_2 &= 4 \cdot 2\left[\cos\left(-\frac{\pi}{6} + \frac{2\pi}{3}\right) + i\left(-\frac{\pi}{6} + \frac{2\pi}{3}\right)\right] \\ &= 8\left[\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right] \\ &= i8, \end{aligned}$$

ενώ από την (1.42) το πηλίκο

$$\begin{aligned} \frac{z_1}{z_2} &= \frac{4}{2}\left[\cos\left(-\frac{\pi}{6} - \frac{2\pi}{3}\right) + i\sin\left(-\frac{\pi}{6} - \frac{2\pi}{3}\right)\right] \\ &= 2\left[\cos\left(-\frac{5\pi}{6}\right) + i\sin\left(-\frac{5\pi}{6}\right)\right] \\ &= 2\left[-\frac{\sqrt{3}}{2} + i\left(-\frac{1}{2}\right)\right] \\ &= -\sqrt{3} - i. \end{aligned}$$

Εάν $z_k = r_k(\cos\theta_k + i\sin\theta_k)$, $k = 1, 2, \ldots, n$ με μαθηματική επαγωγή μέσω της (1.40) έχουμε

$$z_1 z_2 \cdots z_n = r_1 r_2 \cdots r_n[\cos(\theta_1 + \theta_2 + \cdots + \theta_n) + i\sin(\theta_1 + \theta_2 + \cdots + \theta_n)], \tag{1.43}$$

και ειδικά για $z_1 = z_2 = \cdots = z_n = z$ προκύπτει

$$z^n = r^n[\cos(n\theta) + i\sin(n\theta)], \tag{1.44}$$

για κάθε φυσικό αριθμό n. Εάν $z = r(\cos\theta + i\sin\theta)$, $r \neq 0$, από τις σχέσεις (1.41) και (1.44) έπεται ότι για $n = 1, 2, 3, \ldots$

$$z^{-n} = (z^{-1})^n = \left(\frac{1}{z}\right)^n = \frac{1}{r^n}[\cos(-n\theta) + i\sin(-n\theta)],$$

και επειδή $z^0 = 1$, τελικά η σχέση (1.44) ισχύει για κάθε ακέραιο $0, \pm 1, \pm 2, \ldots$.

Εάν $z = \cos\theta + i\sin\theta$ η (1.44) μετασχηματίζεται στην

$$(\cos\theta + i\sin\theta)^n = \cos(n\theta) + i\sin(n\theta). \tag{1.45}$$

Η τελευταία σχέση είναι γνωστή ως **τύπος του de Moivre**.

Μία ενδιαφέρουσα εφαρμογή του τύπου του de Moivre είναι η εύρεση ριζών μιγαδικών αριθμών. Θέτουμε λοιπόν το εξής

ΠΡΟΒΛΗΜΑ: Εάν w είναι ένας μιγαδικός αριθμός και $n \geq 2$ είναι ένας φυσικός αριθμός να βρεθούν μιγαδικοί z τέτοιοι ώστε $z^n = w$. Έστω ότι $w = r(\cos\theta + i\sin\theta)$, τότε παρατηρούμε ότι ο αριθμός $z_0 = r^{1/n}(\cos(\theta/n) + i\sin(\theta/n))$ ικανοποιεί την

$$z_0^n = \left[\sqrt[n]{r}\left(\cos\frac{\theta}{n} + i\sin\frac{\theta}{n}\right)\right]^n = r(\cos\theta + i\sin\theta) = w, \tag{1.46}$$

δηλαδή ο z_0 είναι μία λύση του προβλήματος. Όμως και οι $z_k = r^{1/n}(\cos[(\theta+2k\pi)/n] + i\sin[(\theta+2k\pi)/n])$, $k = 1, 2, \ldots$ είναι λύσεις μιας και

$$z_k^n = \left[\sqrt[n]{r}\left(\cos\frac{\theta+2k\pi}{n} + i\sin\frac{\theta+2k\pi}{n}\right)\right]^n = r(\cos(\theta+2k\pi) + i\sin(\theta+2k\pi)) = w. \tag{1.47}$$

Από τις (1.46) και (1.47) βλέπουμε ότι οι αριθμοί $z_k = r^{1/n}(\cos[(\theta+2k\pi)/n] + i\sin[(\theta+2k\pi)/n])$ ικανοποιούν $z_k^n = w$ για $k = 0, 1, 2, \ldots$. Στη συνέχεια θυμίζουμε ότι για n σταθερό κάθε $k \in \mathbb{N}$ γράφεται μοναδικά στη μορφή $k = m + ln$ όπου $m = 0, 1, \ldots, n-1$ και $l \in \mathbb{N}$ (διαίρεση του k δια n). Έτσι εάν $k \geq n$, τότε

$$\begin{aligned} z_k &= \sqrt[n]{r}\left(\cos\frac{\theta+2k\pi}{n} + i\sin\frac{\theta+2k\pi}{n}\right) \\ &= \sqrt[n]{r}\left(\cos\frac{\theta+2m\pi+2ln\pi}{n} + i\sin\frac{\theta+2m\pi+2ln\pi}{n}\right) \\ &= \sqrt[n]{r}\left(\cos\left(\frac{\theta+2m\pi}{n} + 2l\pi\right) + i\sin\left(\frac{\theta+2m\pi}{n} + 2l\pi\right)\right) \\ &= \sqrt[n]{r}\left(\cos\frac{\theta+2m\pi}{n} + i\sin\frac{\theta+2m\pi}{n}\right) \\ &= z_m \end{aligned}$$

όπου $m = 0, 1, 2, \ldots, n-1$. **Συμπέρασμα**: Οι n το πλήθος μιγαδικοί αριθμοί

$$z_k = \sqrt[n]{r}\left(\cos\frac{\theta+2k\pi}{n} + i\sin\frac{\theta+2k\pi}{n}\right), \qquad k = 0, 1, 2, \ldots, n-1 \tag{1.48}$$

είναι οι λύσεις της εξίσωσης $z^n = r(\cos\theta + i\sin\theta)$ και λέγονται *ν*οστες ρίζες του $w = r(\cos\theta + i\sin\theta)$.

Παράδειγμα 1.4. Επειδή $1 = \cos 0 + i\sin 0$, οι **n-οστες ρίζες της μονάδας** είναι οι αριθμοί

$$\zeta_k = \cos\frac{2k\pi}{n} + i\sin\frac{2k\pi}{n}, \qquad k = 0, 1, 2, \ldots, n-1. \tag{1.49}$$

Εάν ορίσουμε

$$\omega_n = \cos\frac{2\pi}{n} + i\sin\frac{2\pi}{n}, \tag{1.50}$$

τότε από τον τύπο του de Moivre έπεται ότι οι *ν*οστες ρίζες της μονάδας είναι οι 1, ω_n, ω_n^2, …, ω_n^{n-1}. Παρατηρούμε ότι $\omega_n^n = 1$. Οι *ν*οστες ρίζες της μονάδας είναι οι κορυφές ενός κανονικού πολυγώνου με n πλευρές εγγεγραμμένου στο μοναδιαίο κύκλο.

Παράδειγμα 1.5. Να βρεθούν αριθμοί z τέτοιοι ώστε $z^2 = -2$ (τετραγωνικές ρίζες του -2).
Είναι $-2 = 2(\cos\pi + i\sin\pi)$, οπότε οι αριθμοί που ζητούμε δίνονται από τη σχέση

$$z_k = \sqrt{2}\left(\cos\frac{\pi+2k\pi}{2} + i\sin\frac{\pi+2k\pi}{2}\right), \qquad k = 0, 1.$$

Έτσι έχουμε

$$z_0 = \sqrt{2}\left(\cos\frac{\pi}{2} + i\sin\frac{\pi}{2}\right) = i\sqrt{2}, \qquad z_1 = \sqrt{2}\left(\cos\frac{3\pi}{2} + i\sin\frac{3\pi}{2}\right) = -i\sqrt{2}.$$

Πράγματι $z_0^2 = (i\sqrt{2})^2 = i^2 2 = -2$ και $z_1^2 = (-i\sqrt{2})^2 = i^2 2 = -2$.

Κεφάλαιο 2

Μητρώα

2.1 Γραμμικά συστήματα και μητρώα

Ας θεωρήσουμε το σύστημα τριών εξισώσεων

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned} \tag{2.1}$$

με αγνώστους τις x, y, z. Την (2.1) μπορούμε να την δούμε σαν ισότητα μεταξύ δύο διανυσμάτων του $\mathbb{R}^3$, συγκεκριμένα

$$\begin{pmatrix} a_1x + b_1y + c_1z \\ a_2x + b_2y + c_2z \\ a_3x + b_3y + c_3z \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}$$

απ΄ όπου, ως απόρροια των πράξεων των διανυσμάτων προκύπτει ότι το αρχικό σύστημα μπορεί να γραφεί στη μορφή

$$\begin{pmatrix} a_1x \\ a_2x \\ a_3x \end{pmatrix} + \begin{pmatrix} b_1y \\ b_2y \\ b_3y \end{pmatrix} + \begin{pmatrix} c_1z \\ c_2z \\ c_3z \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix},$$

ή

$$\begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} x + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} y + \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} z = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}. \tag{2.2}$$

Κωδικοποιώντας μπορούμε να γράψουμε το σύστημα (2.1) στη μορφή

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix}, \tag{2.3}$$

όπου το αντικείμενο

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix},$$

περιέχει καταχωρημένο σε κατάλληλη θέση τον συντελεστή κάθε αγνώστου κάθε εξίσωσης. Συγκεκριμένα η πρώτη γραμμή περιέχει τους συντελεστές της πρώτης εξίσωσης, ή πρώτη στήλη τους συντελεστές του πρώτου αγνώστου και ούτω καθεξής. Η διάταξη αυτή μπορεί να ιδωθεί ως ένα πολύστηλο διάνυσμα, και ονομάζεται **μητρώο**, ή **μήτρα** (matrix)[1].[2] Επισημαίνουμε ότι ο τρόπος γραφής (2.3) του αρχικού συστήματος εμπεριέχει την πράξη του πολλαπλασιασμού του μητρώου των συντελεστών του συστήματος με το διάνυσμα των αγνώστων του συστήματος

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix}.$$

Συγκεκριμένα από την ισοδυναμία των (2.1), (2.2) και (2.3) προκύπτει

$$\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \\ a_3 & b_3 & c_3 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} a_1 \\ a_2 \\ a_3 \end{pmatrix} x + \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} y + \begin{pmatrix} c_1 \\ c_2 \\ c_3 \end{pmatrix} z = \begin{pmatrix} a_1x + b_1y + c_1z \\ a_2x + b_2y + c_2z \\ a_3x + b_3y + c_3z \end{pmatrix}. \tag{2.4}$$

Σημειώνουμε ότι το αρχικό σύστημα μπορεί να παρασταθεί με το μητρώο όλων των δεδομένων (data) του συστήματος, το **επαυξημένο** (augmented) μητρώο όπως ονομάζεται,

$$\begin{pmatrix} a_1 & b_1 & c_1 & d_1 \\ a_2 & b_2 & c_2 & d_2 \\ a_3 & b_3 & c_3 & d_3 \end{pmatrix} \tag{2.5}$$

στο οποίο, όπως και στην περίπτωση του μητρώου των συντελεστών, αναγνωρίζουμε τί περιέχει η κάθε γραμμή και η κάθε στήλη. Για παράδειγμα η πρώτη στήλη περιέχει τους συντελεστές του πρώτου αγνώστου με τη σειρά που εμφανίζονται σε κάθε εξίσωση του συστήματος, και ανάλογα η δεύτερη γραμμή περιέχει τα δεδομένα της δεύτερης εξίσωσης με τη σωστή σειρά παράθεσης.

Ορισμός 2.1. Έστω ότι τα $\boldsymbol{a}_1, \boldsymbol{a}_2, \ldots, \boldsymbol{a}_m$ είναι διανύσματα του $\mathbb{R}^n$, όπου γράφουμε

$$\boldsymbol{a}_k = \begin{pmatrix} a_{1k} \\ a_{2k} \\ \vdots \\ a_{nk} \end{pmatrix}, \qquad k = 1, 2, \ldots, m.$$

[1] Λατινικά για το uterus, womb δηλαδή μήτρα.

[2] Ο James Joseph Sylvester είναι, μάλλον, ο πρώτος που έδωσε σημασία σε "μια ορθογώνια δάταξη όρων σε γραμμές και στήλες" και σε εργασία του το 1850 [11] έγραψε: "For this purpose we must commence, not with a square, but with an oblong arrangement of terms consisting, suppose, of m lines and n columns. This will not in itself represent a determinant, but is, as it were, a Matrix out of which we may form various systems of determinants by fixing upon a number p, and selecting at will p lines and p columns, the squares corresponding of pth order." Αλλού δε [12] αναφέρει: "I have (...) defined a Matrix as a rectangular array of terms, out of which different systems of determinants may be engendered, as from the womb of a common parent ..." Χρησιμοποίησε λοιπόν ο Sylvester τον όρο matrix με τη συμβατική του σημασία, να σημαίνει δηλαδή "το μέσο από το οποίο κάτι άλλο προέρχεται ή αναπτύσσεται". Ο Sylvester, ο οποίος ενδιαφερόταν κυρίως για τις ορίζουσες, χρησιμοποίησε το αντικείμενο matrix μάλλον τυπικά. Αργότερα ο Arthur Cayley, ο οποίος υπήρξε στενός συνεργάτης του Sylvester, μελετώντας τους μετασχηματισμούς της γραμμικής κίνησης (Παράδειγμα 2.15) συνειδητοποίησε ότι το αντικείμενο matrix είναι μια αλγεβρική οντότητα και σε εργασίες του το 1855 και 1858 όρισε το αντικείμενο matrix καθώς και τον πολλαπλασιασμό μεταξύ τέτοιων αντικειμένων, ώστε να αντιστοιχεί στη σύνθεση των μετασχηματισμών, και την πρόσθεση θεμελιώνοντας έτσι την Άλγεβρα Μητρώων.

Τη διάταξη-παράθεση

$$\begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_m\end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix}$$

λέμε **μητρώο** με n γραμμές και m στήλες, ή απλά $n \times m$ μητρώο. Λέμε επίσης ότι το μητρώο A έχει **διάσταση** $n \times m$. Ένα $n \times n$ μητρώο λέγεται **τετραγωνικό**. Εάν A είναι ένα $n \times m$ μητρώο γράφουμε $A = \begin{pmatrix} a_{ij} \end{pmatrix}$, $i = 1, 2, \ldots, n$, $j = 1, 2, \ldots, m$. Το σύνολο των $n \times m$ μητρώων $A = \begin{pmatrix} a_{ij} \end{pmatrix}$ συμβολίζουμε με $\mathbb{M}^{n,m}(\mathbb{K})$, όπου με $\mathbb{K}$ συμβολίζουμε το σώμα των πραγματικών, ή μιγαδικών αριθμών[3], δηλαδή $\mathbb{K} = \mathbb{R}$ αν οι a_{ij} είναι πραγματικοί αριθμοί, ή $\mathbb{K} = \mathbb{C}$ αν οι a_{ij} είναι μιγαδικοί αριθμοί. Πολλές φορές γράφουμε απλά $\mathbb{M}^{n,m}$ αν δεν χρειάζεται να γίνει ιδιαίτερη αναφορά στο σώμα. Ενίοτε γράφουμε $\mathbb{R}^{n\times m}$ αντί του $\mathbb{M}^{n,m}(\mathbb{R})$ και $\mathbb{C}^{n\times m}$ αντί του $\mathbb{M}^{n,m}(\mathbb{C})$.

Παράδειγμα 2.1. Σύμφωνα με τον ορισμό τα

$$\begin{pmatrix} a & b & c \end{pmatrix}, \quad \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \begin{pmatrix} a & b \\ c & d \\ e & f \end{pmatrix}, \quad \begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}, \quad \begin{pmatrix} a \\ b \\ c \end{pmatrix}, \quad \begin{pmatrix} a \end{pmatrix}$$

είναι μητρώα διαστάσεων 1×3, 2×2, 3×2, 2×3, 3×1, και 1×1 αντίστοιχα.

Ένα $n \times 1$ μητρώο είναι ένα διάνυσμα στο $\mathbb{R}^n$, ή στο $\mathbb{C}^n$. Πολλές φορές όταν δεν δημιουργείται σύγχυση ένα 1×1 μητρώο $\begin{pmatrix} a \end{pmatrix}$ μπορούμε να το ταυτίζουμε με τον αριθμό a.

Σημειώνουμε ότι με τη χρήση μητρώων μπορούμε να γράψουμε, τελικά μέσω της (2.3), το

[3]Το σύνολο των πραγματικών αριθμών όπως και αυτό των μιγαδικών αριθμών με τις πράξεις της πρόσθεσης και του πολλαπλασιασμού αποτελούν αλγεβρικές δομές στις οποίες οι δύο πράξεις ικανοποιούν τις ίδιες χαρακτηριστικές ιδιότητες. Η δομή αυτή λέγεται σώμα. Συγκεκριμένα έχουμε

Ορισμός. Έστω $\mathcal{F}$ ένα μη κενό σύνολο στο οποίο έχουν ορισθεί δύο πράξεις, πρόσθεση "+" και πολλαπλασιασμός "·". Η δομή $(\mathcal{F}, +, \cdot)$ λέγεται **σώμα** (field) εάν ικανοποιούνται οι νόμοι:

(F1) $x + y \in \mathcal{F}$, για κάθε $x, y \in \mathcal{F}$.

(F2) $x + y = y + x$, για κάθε $x, y \in \mathcal{F}$.

(F3) $(x + y) + z = x + (y + z)$, για κάθε $x, y, z \in \mathcal{F}$.

(F4) Υπάρχει $0 \in \mathcal{F}$, ώστε $x + 0 = x$, για κάθε $x \in \mathcal{F}$.

(F5) Για κάθε $x \in \mathcal{F}$ υπάρχει $-x \in \mathcal{F}$, ώστε $x + (-x) = 0$.

(F6) $x \cdot y \in \mathcal{F}$, για κάθε $x, y \in \mathcal{F}$.

(F7) $x \cdot y = y \cdot x$, για κάθε $x, y \in \mathcal{F}$.

(F8) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$, για κάθε $x, y, z \in \mathcal{F}$.

(F9) Υπάρχει $1 \in \mathcal{F}$, $1 \neq 0$ ώστε $x \cdot 1 = x$, για κάθε $x \in \mathcal{F}$.

(F10) Για κάθε $x \neq 0$ υπάρχει $x^{-1} \in \mathcal{F}$, ώστε $x \cdot x^{-1} = 1$.

(F11) $x \cdot (y + z) = x \cdot y + x \cdot z$, για κάθε $x, y, z \in \mathcal{F}$.

σύστημα (2.1) σαν εξίσωση στη μορφή

$$A\boldsymbol{x} = \boldsymbol{b} \tag{2.6}$$

όπου A είναι το μητρώο των συντελεστών, $\boldsymbol{x}$ το διάνυσμα των αγνώστων και $\boldsymbol{b}$ το διάνυσμα των σταθερών. Γενικότερα, επιβεβαιώστε μέσω της (2.2) ότι, ένα ανάλογο σύστημα n εξισώσεων με m αγνώστους μπορεί να γραφεί στη μορφή (2.6) όπου το A θα είναι ένα $n \times m$ μητρώο, $\boldsymbol{x} \in \mathbb{R}^m$ και $\boldsymbol{b} \in \mathbb{R}^n$.

2.2 Πράξεις μεταξύ μητρώων

Αν σκεφτούμε τα μητρώα ως πολύστηλα διανύσματα, τότε μπορούμε να μιλάμε για ισότητα μητρώων. Επιπλέον είναι λογικό να θελήσουμε να επεκτείνουμε τόσο την πράξη της πρόσθεσης διανυσμάτων όσο και αυτή του πολλαπλασιασμού διανύσματος με σταθερά και στα μητρώα.

Ορισμός 2.2. Θα λέμε ότι τα μητρώα A και B είναι ίσα, αν έχουν την ίδια διάσταση, έστω $n \times m$, και αν $A = \left(a_{ij}\right)$ και $B = \left(b_{ij}\right)$, τότε $a_{ij} = b_{ij}$ για κάθε $i \in \{1, 2, \dots, n\}$ και για κάθε $j \in \{1, 2, \dots, m\}$.

Αν τώρα $A = \left(a_{ij}\right)$ και $B = \left(b_{ij}\right)$ είναι $n \times m$ μητρώα και λ μια σταθερά ορίζουμε αντίστοιχα

$$A + B = \begin{pmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots & a_{1m} + b_{1m} \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots & a_{2m} + b_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} + b_{n1} & a_{n2} + b_{n2} & \cdots & a_{nm} + b_{nm} \end{pmatrix}, \qquad \lambda A = \begin{pmatrix} \lambda a_{11} & \lambda a_{12} & \cdots & \lambda a_{1m} \\ \lambda a_{21} & \lambda a_{22} & \cdots & \lambda a_{2m} \\ \vdots & \vdots & & \vdots \\ \lambda a_{n1} & \lambda a_{n2} & \cdots & \lambda a_{nm} \end{pmatrix} \tag{2.7}$$

και παρατηρούμε ότι και η αφαίρεση μπορεί, όπως στα διανύσματα, να ορισθεί έμμεσα ως

$$A - B = A + (-1)B = \begin{pmatrix} a_{11} - b_{11} & a_{12} - b_{12} & \cdots & a_{1m} - b_{1m} \\ a_{21} - b_{21} & a_{22} - b_{22} & \cdots & a_{2m} - b_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} - b_{n1} & a_{n2} - b_{n2} & \cdots & a_{nm} - b_{nm} \end{pmatrix}.$$

Παρατήρηση 2.1 (Το μηδενικό μητρώο). Με O συμβολίζουμε, για κάθε διάσταση, το μητρώο του οποίου όλα τα στοιχεία είναι ίσα με 0. Έτσι αν A είναι ένα $n \times m$ μητρώο, τότε $A + O = O + A = A$. Εδώ το O έχει διάσταση $n \times m$. Το O ονομάζουμε **μηδενικό μητρώο** ή απλά μηδέν.

▲ **Πρόταση 2.1.** *Εάν A, B, C είναι $n \times m$ μητρώα, O είναι το μηδενικό $n \times m$ μητρώο, και λ, μ είναι σταθερές τότε ισχύουν οι νόμοι*

(1) $A + B = B + A$

(2) $A + (B + C) = (A + B) + C$

(3) $\lambda(A + B) = \lambda A + \lambda B$

(4) $(\lambda + \mu)A = \lambda A + \mu A$

(5) $\lambda(\mu A) = (\lambda\mu)A$

(6) $0A = O$

Απόδειξη. Αφήνεται σαν Άσκηση. □

Πολλαπλασιασμός μητρώων

Εάν $A = \left(a_{ij}\right) \in \mathbb{M}^{n,m}(\mathbb{R})$, και $\boldsymbol{b}$ ένα διάνυσμα στο $\mathbb{R}^m$, όπως στην (2.4), ορίζουμε το γινόμενο $A\boldsymbol{b}$ με τη σχέση

$$A\boldsymbol{b} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \begin{pmatrix} a_{11}b_1 + a_{12}b_2 + \cdots + a_{1m}b_m \\ a_{21}b_1 + a_{22}b_2 + \cdots + a_{2m}b_m \\ \vdots \\ a_{n1}b_1 + a_{n2}b_2 + \cdots + a_{nm}b_m \end{pmatrix}, \tag{2.8}$$

και ως συνέπεια των πράξεων των διανυσμάτων έχουμε επιπλέον ότι

$$A\boldsymbol{b} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix} = \begin{pmatrix} a_{11} \\ a_{21} \\ \vdots \\ a_{n1} \end{pmatrix} b_1 + \begin{pmatrix} a_{12} \\ a_{22} \\ \vdots \\ a_{n2} \end{pmatrix} b_2 + \cdots + \begin{pmatrix} a_{1m} \\ a_{2m} \\ \vdots \\ a_{nm} \end{pmatrix} b_m. \tag{2.9}$$

Γενικεύοντας το πολλαπλασιασμό, αν $A = \left(a_{ij}\right)$ είναι ένα $n \times m$ μητρώο και $B = \left(b_{ij}\right)$ είναι ένα $m \times k$ μητρώο γράφοντας

$$B = \left(\boldsymbol{b}_1 \ \ \boldsymbol{b}_2 \ \ \cdots \ \ \boldsymbol{b}_k\right),$$

οπου $\boldsymbol{b}_j \in \mathbb{R}^m$ για $j = 1, 2, \ldots, k$, ορίζουμε το γινόμενο AB με τη σχέση

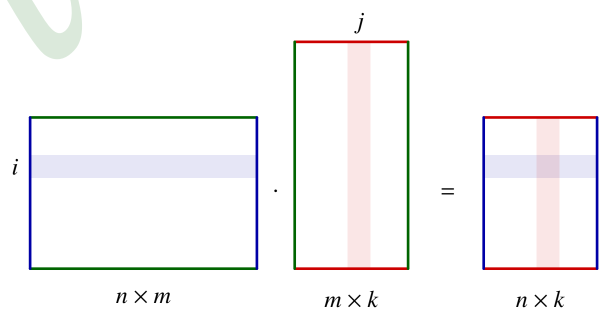
$$AB = \left(A\boldsymbol{b}_1 \ \ A\boldsymbol{b}_2 \ \ \cdots \ \ A\boldsymbol{b}_k\right). \tag{2.10}$$

Έτσι το γινόμενο AB είναι το $n \times k$ μητρώο $C = AB$ με

$$C = \left(\boldsymbol{c}_1 \ \ \boldsymbol{c}_2 \ \ \cdots \ \ \boldsymbol{c}_k\right), \qquad \text{όπου} \qquad \boldsymbol{c}_j = A\boldsymbol{b}_j, \qquad j = 1, 2, \ldots, k \tag{2.11}$$

ειδικότερα αν $C = \left(c_{ij}\right)$, τότε μέσω της (2.9) προκύπτει ότι

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{im}b_{mj}. \tag{2.12}$$



Σχήμα 2.1: Οπτικοποίηση του γινομένου μητρώων, η συμβατότητα στη διάσταση αποδίδεται με το ίδιο χρώμα. Η i-στήλη του μητρώου στα αριστερά πολλαπλασιάζει την j-στήλη του μητρώου στα δεξιά για να δώσει το στοιχείο στην i-γραμμή και j-στήλη του γινομένου.

Σημείωση 2.1. Από τις (2.8) και (2.10) συνάγονται τα παρακάτω συμπεράσματα:

(1) Το γινόμενο ενός $n \times m$ μητρώου με διάνυσμα του $\mathbb{R}^k$ ορίζεται μόνο αν $k = m$.

(2) Το γινόμενο ενός $n \times m$ μητρώου με ένα $m \times k$ μητρώο είναι ένα $n \times k$ μητρώο, σχηματικά

$$(n \times m) \cdot (m \times k) \to n \times k$$

(3) Εάν τα A και B είναι μητρώα συμβατών διαστάσεων ώστε να ορίζεται το γινόμενο και $C = AB$, τότε το στοιχείο c_{ij} του C είναι το γινόμενο του μητρώου "i-γραμμή του A" με το μητρώο "j-στήλη του B", βλέπε (2.12).

Παράδειγμα 2.2. Να υπολογισθούν τα a και b στο γινόμενο που ακολουθεί

$$\begin{pmatrix} 1 & 2 & -2 \\ 2 & 3 & 0 \end{pmatrix} \begin{pmatrix} 3 & -1 & 0 & 2 \\ 2 & -3 & 1 & 2 \\ 4 & 5 & -2 & 1 \end{pmatrix} = \begin{pmatrix} \square & \square & a & \square \\ \square & b & \square & \square \end{pmatrix}.$$

Το μητρώο στα αριστερά είναι 2×3 και αυτό στα δεξιά 3×4 κατά συνέπεια το γινόμενο ορίζεται και είναι ένα 2×4 μητρώο.
Το a βρίσκεται στην πρώτη γραμμή και στην τρίτη στήλη του γινομένου, άρα

$$a = \begin{pmatrix} 1 & 2 & -2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ -2 \end{pmatrix} = (1)(0) + (2)(1) + (-2)(-2) = 6$$

Το b βρίσκεται στην δεύτερη γραμμή και στην δεύτερη στήλη του γινομένου, άρα

$$b = \begin{pmatrix} 2 & 3 & 0 \end{pmatrix} \begin{pmatrix} -1 \\ -3 \\ 5 \end{pmatrix} = (2)(-1) + (3)(-3) + (0)(5) = -11.$$

Παράδειγμα 2.3. Παρατηρείστε ότι αν

$$A = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} \quad \text{και} \quad P = \begin{pmatrix} p_1 & q_1 \\ p_2 & q_2 \\ p_3 & q_3 \end{pmatrix},$$

τότε

$$\begin{aligned} AP = \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} \begin{pmatrix} p_1 & q_1 \\ p_2 & q_2 \\ p_3 & q_3 \end{pmatrix} &= \left(\begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} \begin{pmatrix} p_1 \\ p_2 \\ p_3 \end{pmatrix} \quad \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} \begin{pmatrix} q_1 \\ q_2 \\ q_3 \end{pmatrix} \right) \\ &= \begin{pmatrix} a_1p_1 + b_1p_2 + c_1p_3 & a_1q_1 + b_1q_2 + c_1q_3 \\ a_2p_1 + b_2p_2 + c_2p_3 & a_2q_1 + b_2q_2 + c_2q_3 \end{pmatrix} \end{aligned}$$

ενώ

$$PA = \begin{pmatrix} p_1 & q_1 \\ p_2 & q_2 \\ p_3 & q_3 \end{pmatrix} \begin{pmatrix} a_1 & b_1 & c_1 \\ a_2 & b_2 & c_2 \end{pmatrix} = \begin{pmatrix} p_1a_1 + q_1a_2 & p_1b_1 + q_1b_2 & p_1c_1 + q_1c_2 \\ p_2a_1 + q_2a_2 & p_2b_1 + q_2b_2 & p_2c_1 + q_2c_2 \\ p_3a_1 + q_3a_2 & p_3b_1 + q_3b_2 & p_3c_1 + q_3c_2 \end{pmatrix}$$

κατά συνέπεια $AP \neq PA$ αφού δεν έχουν την ίδια διάσταση.

Παράδειγμα 2.4. Αν

$$A = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix} \qquad \text{και} \qquad B = \begin{pmatrix} 5 & 2 \\ 1 & -1 \end{pmatrix},$$

τότε

$$AB = \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 5 & 2 \\ 1 & -1 \end{pmatrix} = \begin{pmatrix} 9 & 5 \\ 3 & -3 \end{pmatrix}$$

$$BA = \begin{pmatrix} 5 & 2 \\ 1 & -1 \end{pmatrix} \begin{pmatrix} 2 & -1 \\ 0 & 3 \end{pmatrix} = \begin{pmatrix} 10 & 1 \\ 2 & -4 \end{pmatrix},$$

οπότε $AB \neq BA$.

Παράδειγμα 2.5. Αν

$$P = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \qquad \text{και} \qquad Q = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & 4 \end{pmatrix},$$

τότε

$$PQ = \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & 4 \end{pmatrix} = \begin{pmatrix} 4 & 6 & 6 \\ 7 & 3 & 13 \end{pmatrix}$$

αλλά το

$$QP = \begin{pmatrix} 1 & 3 & 1 \\ 2 & 0 & 4 \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 3 \end{pmatrix}$$

δεν ορίζεται.

▲ **Πρόταση 2.2.** *Εάν τα μητρώα A, B, C έχουν διατάσεις ώστε η κάθε μια από τις πράξεις να μπορεί να εκτελεστεί, και λ, μ είναι σταθερές, τότε ισχύουν οι νόμοι*

(1) $A(BC) = (AB)C$

(2) $A(B + C) = AB + AC$

(3) $(A + B)C = AC + BC$

(4) $A(\lambda B) = (\lambda A)B = \lambda(AB)$

Απόδειξη. Αφήνεται σαν Άσκηση. □

Παρατήρηση 2.2 (Η ανατομία του γινομένου). Πέρα από τον μηχανιστικό τρόπο εκτέλεσης του πολλαπλασιασμού μητρώων μέσω της (2.12), η κατανόηση της δομής του γινομένου επιτρέπει την εξαγωγή ιδιαίτερα χρήσιμων συμπερασμάτων για το γινόμενο. Ας δούμε ένα απλό παράδειγμα. Είδαμε ότι

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix} \begin{pmatrix} u & v \\ w & x \\ y & z \end{pmatrix} = \begin{pmatrix} au + bw + cy & av + bx + cz \\ du + ew + fy & dv + ex + fz \end{pmatrix},$$

οπότε παριστώντας τις στήλες μητρώου ως μητρώα με μία στήλη, και τις γραμμές ως μητρώα με μία γραμμή βλέπουμε αφενός ότι

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}\begin{pmatrix} u & v \\ w & x \\ y & z \end{pmatrix} = \left(\begin{pmatrix} a \\ d \end{pmatrix}u + \begin{pmatrix} b \\ e \end{pmatrix}w + \begin{pmatrix} c \\ f \end{pmatrix}y \quad \begin{pmatrix} a \\ d \end{pmatrix}v + \begin{pmatrix} b \\ e \end{pmatrix}x + \begin{pmatrix} c \\ f \end{pmatrix}z\right) \tag{2.13}$$

απ΄ όπου έπεται ότι *κάθε στήλη του γινομένου είναι γραμμικός συνδυασμός των στηλών του πρώτου μητρώου με συντελεστές τα στοιχεία της αντίστοιχης στήλης του δεύτερου μητρώου*, και αφετέρου

$$\begin{pmatrix} a & b & c \\ d & e & f \end{pmatrix}\begin{pmatrix} u & v \\ w & x \\ y & z \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} u & v \end{pmatrix}a + \begin{pmatrix} w & x \end{pmatrix}b + \begin{pmatrix} y & z \end{pmatrix}c \\ \begin{pmatrix} u & v \end{pmatrix}d + \begin{pmatrix} w & x \end{pmatrix}e + \begin{pmatrix} y & z \end{pmatrix}f \end{pmatrix} \tag{2.14}$$

απ΄ όπου έπεται ότι *κάθε γραμμή του γινομένου είναι γραμμικός συνδυασμός των γραμμών του δεύτερου μητρώου με συντελεστές τα στοιχεία της αντίστοιχης γραμμής του πρώτου μητρώου*. Για γενικά μητρώα ας επεκτείνουμε τον γνωστό συμβολισμό ως εξής: Για το $n \times m$ μητρώο Q γράφουμε $\boldsymbol{q}_j$ για την j-στήλη του Q ως $n \times 1$ μητρώο και $\boldsymbol{q}^i$ για την i-γραμμή του Q ως $1 \times m$ μητρώο, ώστε

$$Q = \begin{pmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 & \ldots & \boldsymbol{q}_m \end{pmatrix} = \begin{pmatrix} \boldsymbol{q}^1 \\ \boldsymbol{q}^2 \\ \vdots \\ \boldsymbol{q}^n \end{pmatrix}.$$

Αν $A = \begin{pmatrix} a_{ij} \end{pmatrix}$ είναι ένα $n \times m$ μητρώο και $B = \begin{pmatrix} b_{ij} \end{pmatrix}$ είναι ένα $m \times k$ μητρώο, μια προσεκτική ανάγνωση των (2.8), (2.9) και (2.10), βλέπε και (2.13), (2.14) αποκαλύπτει για το γινόμενο $C = AB$ τα εξής:

(1) Κάθε στήλη του AB είναι γραμμικός συνδυασμός των στηλών του A με συντελεστές τα στοιχεία της αντίστοιχης στήλης του B, με άλλα λόγια αν

$$AB = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_m \end{pmatrix}\begin{pmatrix} \boldsymbol{b}_1 & \boldsymbol{b}_2 & \cdots & \boldsymbol{b}_k \end{pmatrix} = \begin{pmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \cdots & \boldsymbol{c}_k \end{pmatrix} = C$$

τότε

$$\boldsymbol{c}_j = \boldsymbol{a}_1 b_{1j} + \boldsymbol{a}_2 b_{2j} + \cdots + \boldsymbol{a}_m b_{mj} = A\boldsymbol{b}_j, \qquad j = 1, 2, \ldots, k. \tag{2.15}$$

(2) Κάθε γραμμή του AB είναι γραμμικός συνδυασμός των γραμμών του B με συντελεστές τα στοιχεία της αντίστοιχης γραμμής του A, με άλλα λόγια αν

$$AB = \begin{pmatrix} \boldsymbol{a}^1 \\ \boldsymbol{a}^2 \\ \vdots \\ \boldsymbol{a}^n \end{pmatrix}\begin{pmatrix} \boldsymbol{b}^1 \\ \boldsymbol{b}^2 \\ \vdots \\ \boldsymbol{b}^m \end{pmatrix} = \begin{pmatrix} \boldsymbol{c}^1 \\ \boldsymbol{c}^2 \\ \vdots \\ \boldsymbol{c}^n \end{pmatrix} = C$$

τότε

$$\boldsymbol{c}^i = \boldsymbol{b}^1 a_{i1} + \boldsymbol{b}^2 a_{i2} + \cdots + \boldsymbol{b}^m a_{im} = \boldsymbol{a}^i B, \qquad i = 1, 2, \ldots, n. \tag{2.16}$$

(3) Παρατηρήστε ότι

$$C = \begin{pmatrix} \boldsymbol{a}^1\boldsymbol{b}_1 & \boldsymbol{a}^1\boldsymbol{b}_2 & \cdots & \boldsymbol{a}^1\boldsymbol{b}_k \\ \boldsymbol{a}^2\boldsymbol{b}_1 & \boldsymbol{a}^2\boldsymbol{b}_2 & \cdots & \boldsymbol{a}^2\boldsymbol{b}_k \\ \vdots & \vdots & & \vdots \\ \boldsymbol{a}^n\boldsymbol{b}_1 & \boldsymbol{a}^n\boldsymbol{b}_2 & \cdots & \boldsymbol{a}^n\boldsymbol{b}_k \end{pmatrix}.$$

► **Άσκηση 2.1.** Δείξτε ότι

$$AB = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_m \end{pmatrix} \begin{pmatrix} \boldsymbol{b}^1 \\ \boldsymbol{b}^2 \\ \vdots \\ \boldsymbol{b}^m \end{pmatrix} = \sum_{k=1}^{n} \boldsymbol{a}_k \boldsymbol{b}^k.$$

► **Άσκηση 2.2 (Συνέπειες του πολλαπλασιασμού).** Αν A και B είναι $n \times n$ μητρώα δείξτε ότι:

(α΄) Εάν το A έχει μια μηδενική γραμμή (με όλα τα στοιχεία ίσα με μηδέν), τότε η αντίστοιχη γραμμή του γινομένου AB είναι μηδενική.

(β΄) Εάν το B έχει μια μηδενική στήλη (με όλα τα στοιχεία ίσα με μηδέν), τότε η αντίστοιχη στήλη του γινομένου AB είναι μηδενική.

Παρατήρηση 2.3. Εάν το A είναι ένα $n \times m$ μητρώο, τότε γράφοντας, κατ΄ εξαίρεση, $O_{p\times q}$ το μηδενικό μητρώο διάστασης $p \times q$, έχουμε

$$O_{k\times n}A = O_{k\times m}, \qquad \text{και} \qquad AO_{m\times l} = O_{n\times l},$$

για κάθε $k, l \in \mathbb{N}$.

Το ταυτοτικό μητρώο

Παρατήρηση 2.4. Παρατηρούμε ότι

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 2 & -2 \\ 2 & 3 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -2 \\ 2 & 3 & 0 \end{pmatrix}$$

$$\begin{pmatrix} 1 & 2 & -2 \\ 2 & 3 & 0 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -2 \\ 2 & 3 & 0 \end{pmatrix}$$

δηλαδή το τετραγωνικό $n \times n$ μητρώο $I = \left(i_{ij}\right)$ με $i_{ii} = 1$, και $i_{ij} = 0$ για $i \neq j$, για κατάλληλο n, συμπεριφέρεται όπως το 1 στον πολλαπλασιασμό. Επιπλέον για τετραγωνικά μητρώα

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Ορισμός 2.3. Το τετραγωνικό $n \times n$ μητρώο $I = \left(\delta_{ij}\right)$, όπου δ_{ij} είναι το δέλτα του Kronecker,

$$\delta_{ij} = \begin{cases} 1 & \text{εάν } i = j \\ 0 & \text{εάν } i \neq j \end{cases}$$

θα το λέμε **ταυτοτικό μητρώο** (identity matrix) και θα το συμβολίζουμε με I ή με I_n αν χρειάζεται να δηλωθεί η διάσταση.

Παρατηρούμε ότι αν I είναι το $n \times n$ ταυτοτικό μητρώο, τότε

$$I = \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \ldots & \boldsymbol{e}_n \end{pmatrix},$$

όπου $\boldsymbol{e}_j$ είναι το διάνυσμα του $\mathbb{R}^n$ με 1 στην j-γραμμή και 0 στις υπόλοιπες, βλέπε (1.4).

► **Άσκηση 2.3.** Έστω $A = \begin{pmatrix} \boldsymbol{a}_1 & \boldsymbol{a}_2 & \dots & \boldsymbol{a}_n \end{pmatrix}$ ένα $n \times n$ μητρώο. Δείξτε ότι:

(α΄) Εάν $\boldsymbol{e}_j$, $j = 1, 2, \dots, n$ είναι όπως παραπάνω, δείξτε ότι $A\boldsymbol{e}_j = \boldsymbol{a}_j$, δηλαδή ο πολλαπλασιασμός με το $\boldsymbol{e}_j$ δίνει την j-στήλη του A.

(β΄) Εάν το A είναι τέτοιο ώστε $A\boldsymbol{x} = \mathbf{0}$ για κάθε διάνυσμα $\boldsymbol{x} \in \mathbb{R}^n$, δείξτε ότι το A είναι το μηδενικό $n \times n$ μητρώο O.

(γ΄) Εάν το B είναι ένα $n \times n$ μητρώο τέτοιο ώστε $A\boldsymbol{x} = B\boldsymbol{x}$ για κάθε διάνυσμα $\boldsymbol{x} \in \mathbb{R}^n$, δείξτε ότι $A = B$.

Δυνάμεις και πολυώνυμα μητρώων

Εάν το A είναι τετραγωνικό μητρώο ορίζουμε για $n = 0, 1, 2, \dots$ τις **δυνάμεις** του A

$$A^0 = I, \qquad A^1 = A, \qquad A^2 = AA, \qquad A^{n+1} = AA^n.$$

Έτσι αν το A είναι τετραγωνικό μητρώο και m και n είναι μη αρνητικοί ακέραιοι, τότε

$$\begin{aligned} A^m A^n &= \underbrace{AA\cdots A}_{m \text{ φορές}}\underbrace{AA\cdots A}_{n \text{ φορές}} = A^{m+n} \\ (A^m)^n &= \underbrace{A^m A^m \cdots A^m}_{n \text{ φορές}} = A^{mn}. \end{aligned}$$

Αν A είναι ένα τετραγωνικό μητρώο και m και n είναι μη αρνητικοί ακέραιοι, τότε για $\lambda, \mu \in \mathbb{R}$ ορίζεται ο γραμμικός συνδυασμός

$$\lambda A^m + \mu A^n$$

το οποίο είναι τετραγωνικό μητρώο. Έτσι αν p είναι ένα πολυώνυμο, για παράδειγμα αν $p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_n t^n$, ορίζεται το μητρώο $p(A)$ με τη σχέση

$$p(A) = a_0 I + a_1 A + a_2 A^2 + \cdots + a_n A^n.$$

Επίσης από τις ιδιότητες του πολλαπλασιασμού μητρώων, Πρόταση 2.2, για μητρώα A και B, για τα οποία το γινόμενο AB ορίζεται, και σταθερές λ και μ έπεται ότι

$$(\lambda A)(\mu B) = (\mu\lambda A)B = \mu\lambda(AB) = \lambda\mu(AB).$$

Κατά συνέπεια για τετραγωνικό μητρώο A και σταθερά λ, έπεται ότι

$$(\lambda A)^n = \lambda^n A^n$$

για κάθε θετικό ακέραιο αριθμό n.

► **Άσκηση 2.4.** Να βρεθούν όλα τα 2×2 μητρώα που ικανοποιούν την εξίσωση $I + A^2 = O$, όπου O είναι τη μηδενικό μητρώο.

► **Άσκηση 2.5.** Εάν p και q είναι πολυώνυμα και A είναι ένα τετραγωνικό μητρώο δείξτε ότι

$$(pq)(A) = p(A)q(A),$$

όπου pq είναι το πολυώνυμο γινόμενο των p και q και $p(A)q(A)$ το γινόμενο των μητρώων $p(A)$ και $q(A)$.

Παράδειγμα 2.6. Έστω A ένα τετραγωνικό μητρώο, και $p(t) = 1 + t^2$. Εάν το A είναι "ρίζα" του p, δηλαδή $p(A) = O$, να υπολογισθούν οι θετικές δυνάμεις του μητρώου A.
Από τη σχέση $p(A) = O$ βρίσκουμε

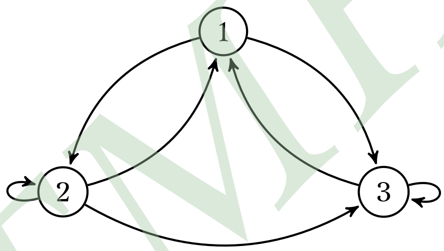
$$I + A^2 = O \Rightarrow A^2 = -I,$$

έτσι από τις ιδιότητες των δυνάμεων μητρώου για κάθε $n = 0, 1, 2, \dots$ υπολογίζουμε

$$\begin{aligned} A^{2n} &= (A^2)^n = (-I)^n = (-1)^n I \\ A^{2n+1} &= A^{2n}A = (-1)^n IA = (-1)^n A. \end{aligned}$$

Σε αναλογία με το Σχόλιο 1.1, που αφορά σε διανύσματα, το μητρώο με την "πολυδιανυσματική" δομή του είναι τελικά ένα από τα πλέον ιδανικά μέσα καταγραφής και διαχείρησης κωδικοποιημένης πληροφορίας.

■ **Εφαρμογή 2.1.** Στη θεωρία γραφημάτων ένα κατευθυνόμενο γράφημα με n κόμβους μπορεί να παρασταθεί με ένα κατάλληλο $n \times n$ μητρώο $M = \left(m_{ij}\right)$. Αν $i, j \in \{1, 2, \dots, n\}$ είναι κόμβοι και υπάρχει κατευθυνόμενη ακμή από το i στο j, όπου επιτρέπεται $i = j$, γράφουμε $m_{ij} = 1$ διαφορετικά $m_{ij} = 0$. Έτσι για το γράφημα στο σχήμα



Σχήμα 2.2: Κατευθυνόμενο γράφημα

το μητρώο που το περιγράφει είναι το

$$M = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Το μητρώο

$$M^2 = \begin{pmatrix} 2 & 1 & 2 \\ 2 & 2 & 3 \\ 1 & 1 & 2 \end{pmatrix}$$

καταμετρά όλα τα μονοπάτια μήκους δύο μεταξύ των κόμβων. Πράγματι αν γράψουμε $M^2 = \left(m_{ij}^{(2)}\right)$, τότε

$$m_{ij}^{(2)} = m_{i1}m_{1j} + m_{i2}m_{2j} + m_{i3}m_{3j}.$$

Κάθε όρος του αθροίσματος αντιστοιχεί στο μονοπάτι $i \to 1 \to j$, $i \to 2 \to j$, $i \to 3 \to j$, αντίστοιχα, και είναι μηδέν ή ένα. Όταν, ο όρος αυτός είναι ένα σημαίνει, λόγω του γινομένου, ότι υπάρχει μονοπάτι από το i στο j μέσω του αντίστοιχου κόμβου 1, 2, ή 3, δηλαδή υπάρχει μονοπάτι μήκους δύο από το i στο j. Έτσι το $m_{ij}^{(2)}$ είναι το πλήθος των μονοπατιών από το i στο j, μήκους δύο, και $m_{ij}^{(2)} \in \{0,1,2,3\}$. Για παράδειγμα

$$\begin{aligned} m_{23}^{(2)} &= m_{21}m_{13} + m_{22}m_{23} + m_{23}m_{33} \\ &= (1)(1) + (1)(1) + (1)(1) = 3 \end{aligned}$$

δηλαδή ο κόμβος 2 συνδέεται με τον 3 με τρία μονοπάτια, τα 213, 223, και 233. Όμοια επειδή $m_{11}^{(2)} = 2$ ο κόμβος 1 θα πρέπει να συνδέεται με τον εαυτό του με δύο μονοπάτια μήκους 2, τα οποία όπως βλέπουμε στο σχήμα είναι τα 121 και 131. Το ανάλογο αποτέλεσμα γενικεύεται και ισχύει για κάθε δύναμη του M, δηλαδή το M^k καταμετρά όλα τα μονοπάτια μήκους k στο γράφημα. Η πλήρης απόδειξη γίνεται με επαγωγή αφού $M^{k+1} = M^k M$.

2.2.1 Το ανάστροφο μητρώο

Ορισμός 2.4. Εάν A είναι ένα $n\times m$ μητρώο ορίζουμε το **ανάστροφο** (transposed) μητρώο A^{T} του A, να είναι το $m\times n$ μητρώο του οποίου η j-στήλη αποτελείται από τα στοιχεία της j-γραμμής του A με την ίδια διάταξη, δηλαδή η πρώτη στήλη του A^{T} είναι η πρώτη γραμμή του A, η δεύτερη στήλη του A^{T} είναι η δεύτερη γραμμή του A και ούτω καθεξής.

Παρατήρηση 2.5. Έτσι αν $A \in \mathbb{M}^{n,m}$, τότε $A^{\mathrm{T}} \in \mathbb{M}^{m,n}$, επιπλέον αν $A = \left(a_{ij}\right)$ τότε $A^{\mathrm{T}} = \left(a'_{ij}\right)$ με $a'_{ij} = a_{ji}$. Άμεση συνέπεια του ορισμού είναι ότι $(A^{\mathrm{T}})^{\mathrm{T}} = A$. Πράγματι αν $(A^{\mathrm{T}})^{\mathrm{T}} = \left(a''_{ij}\right)$, και A^{T} είναι όπως προηγούμενα τότε

$$(A^{\mathrm{T}})^{\mathrm{T}} = \left(a''_{ij}\right) = \left(a'_{ji}\right) = \left(a_{ij}\right) = A. \tag{2.17}$$

Παράδειγμα 2.7. Εάν

$$A = \begin{pmatrix} 1 & 2 & -2 \\ 2 & 3 & 0 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 0 & -2 \end{pmatrix}, \qquad C = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix},$$

από τον ορισμό του ανάστροφου μητρώου βρίσκουμε

$$A^{\mathrm{T}} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ -2 & 0 \end{pmatrix}, \qquad B^{\mathrm{T}} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix}, \qquad C^{\mathrm{T}} = \begin{pmatrix} 1 & 0 & -2 \end{pmatrix}.$$

Παρατηρούμε ότι αφενός $C = B^{\mathrm{T}}$ και $C^{\mathrm{T}} = B$ και αφετέρου ότι $(C^{\mathrm{T}})^{\mathrm{T}} = B^{\mathrm{T}} = C$. Επιπλέον

$$C^{\mathrm{T}}C = \begin{pmatrix} 1 & 0 & -2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} = \begin{pmatrix} 1+0+4 \end{pmatrix} = \begin{pmatrix} 5 \end{pmatrix} = 5$$

ως 1×1 μητρώο, ενώ

$$CC^{\mathrm{T}} = \begin{pmatrix} 1 \\ 0 \\ -2 \end{pmatrix} \begin{pmatrix} 1 & 0 & -2 \end{pmatrix} = \begin{pmatrix} 1\cdot 1 & 1\cdot 0 & 1\cdot(-2) \\ 0\cdot 1 & 0\cdot 0 & 0\cdot(-2) \\ (-2)\cdot 1 & (-2)\cdot 0 & (-2)\cdot(-2) \end{pmatrix} = \begin{pmatrix} 1 & 0 & -2 \\ 0 & 0 & 0 \\ -2 & 0 & 4 \end{pmatrix}.$$

Παρατήρηση 2.6. Ένα $n \times 1$ πραγματικό μητρώο A είναι ένα διάνυσμα στο $\mathbb{R}^n$ και ως τέτοιο έχουμε

$$A^{\mathrm{T}} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \Rightarrow (A^{\mathrm{T}})^{\mathrm{T}} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}^{\mathrm{T}} = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = A$$

γεγονός το οποίο συμφωνεί με τον τρόπο γραφής ενός διανύσματος του $\mathbb{R}^n$ ως διάνυσμα-γραμμή. Παρατηρούμε επίσης ότι

$$A^{\mathrm{T}}A = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = a_1^2 + a_2^2 + \cdots + a_n^2 \tag{2.18}$$

το οποίο είναι αριθμός ως 1×1 μητρώο. Αν $\boldsymbol{a}$ και $\boldsymbol{b}$ είναι διανύσματα στον $\mathbb{R}^n$, κατανοώντας την πράξη $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}$ ως γινόμενο μητρώων θα γράφουμε στη συνέχεια

$$\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} = a_1b_1 + a_2b_2 + \cdots + a_nb_n$$

ισοδύναμα $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = \boldsymbol{a} \cdot \boldsymbol{b}$, όπου "$\cdot$" είναι το σύνηθες εσωτερικό γινόμενο. Παρατηρήστε ότι αν $\boldsymbol{a} \in \mathbb{R}^n$, τότε

$$\|\boldsymbol{a}\| = \sqrt{\boldsymbol{a} \cdot \boldsymbol{a}} = \sqrt{\boldsymbol{a}^{\mathrm{T}}\boldsymbol{a}}.$$

Θεώρημα 2.1. *Εάν A και B είναι μητρώα καταλλήλων διαστάσεων ώστε οι σχετικές πράξεις να μπορούν να εκτελεστούν, τότε*

(1) $(A^{\mathrm{T}})^{\mathrm{T}} = A$.

(2) $(A + B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}$.

(3) $(\lambda A)^{\mathrm{T}} = \lambda A^{\mathrm{T}}$, *για κάθε σταθερά* λ.

(4) $(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}}$.

Απόδειξη. Έστω $A = \left(a_{ij}\right)$ ένα $n \times m$ μητρώο.

(1) Βλέπε (2.17).

(2) Αν $B = \left(b_{ij}\right)$ είναι ένα επίσης $n \times m$ μητρώο, τότε

$$(A + B)^{\mathrm{T}} = \left(a_{ij} + b_{ij}\right)^{\mathrm{T}} = \left(a_{ji} + b_{ji}\right) = \left(a_{ji}\right) + \left(b_{ji}\right) = \left(a_{ij}\right)^{\mathrm{T}} + \left(b_{ij}\right)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}}.$$

(3) Για κάθε σταθερά λ έχουμε

$$(\lambda A)^{\mathrm{T}} = \left(\lambda a_{ij}\right)^{\mathrm{T}} = \left(\lambda a_{ji}\right) = \lambda\left(a_{ji}\right) = \lambda A^{\mathrm{T}}.$$

(4) Έστω $B = \left(b_{ij}\right)$ ένα $m \times k$ μητρώο, τότε το AB ορίζεται και είναι ένα $n \times k$ μητρώο, έτσι το $(AB)^{\mathrm{T}}$ είναι ένα $k \times n$ μητρώο καθώς και το $B^{\mathrm{T}}A^{\mathrm{T}}$ (γιατί;). Αν $(AB)^{\mathrm{T}} = \left(c_{ij}\right)$, τότε

$$\begin{aligned} c_{ij} &= (j\text{-γραμμή του } A)(i\text{-στήλη του } B) \\ &= a_{j1}b_{1i} + a_{j2}b_{2i} + \cdots + a_{jm}b_{mi} \end{aligned}$$

ενώ αν $B^{\mathrm{T}}A^{\mathrm{T}} = \left(d_{ij}\right)$, τότε

$$\begin{aligned} d_{ij} &= (i\text{-γραμμή του } B^{\mathrm{T}})(j\text{-στήλη του } A^{\mathrm{T}}) \\ &= (i\text{-στήλη του } B)^{\mathrm{T}}(j\text{-γραμμή του } A)^{\mathrm{T}} \\ &= b_{1i}a_{j1} + b_{2i}a_{j2} + \cdots + b_{mi}a_{jm}. \end{aligned}$$

Συμπεραίνουμε λοιπόν ότι $c_{ij} = d_{ij}$ που είναι το ζητούμενο. □

▼ **Πόρισμα 2.1.** *Εάν το A είναι τετραγωνικό μητρώο τότε*

$$(A^2)^{\mathrm{T}} = (A^{\mathrm{T}})^2, \qquad (A^n)^{\mathrm{T}} = (A^{\mathrm{T}})^n$$

για $n = 0, 1, 2, \ldots$.

Απόδειξη. Αφήνεται σαν Άσκηση. □

Ορισμός 2.5. Ένα τετραγωνικό μητρώο A για το οποίο ισχύει $A^{\mathrm{T}} = A$ λέγεται **συμμετρικό**.

Παρατήρηση 2.7. Σε τετραγωνικό $n \times n$ μητρώο $A = \left(a_{ij}\right)$ λέγοντας **κύρια διαγώνιο** εννούμε τα "διαγώνια" στοιχεία $a_{11}, a_{22}, \ldots, a_{nn}$. Αν το $A = \left(a_{ij}\right)$ είναι συμμετρικό, τότε από τον ορισμό έπεται ότι $a_{ij} = a_{ji}$, συνεπώς ένα τέτοιο μητρώο είναι συμμετρικό ως προς την κύρια διαγώνιό του, γεγονός που δικαιολογεί τον χαρακτηρισμό συμμετρικό.

► **Άσκηση 2.6.** Εάν το A είναι ένα $n \times m$ μητρώο δείξτε ότι το καθένα από τα μητρώα $A^{\mathrm{T}}A$ και AA^{T} ορίζεται και είναι συμμετρικό.

Θεώρημα 2.2. *Αν A και B είναι $n \times n$ συμμετρικά μητρώα και k είναι μια σταθερά, τότε*

(1) *Το A^{T} είναι συμμετρικό.*

(2) *Το $A + B$ είναι συμμετρικό.*

(3) *Το kA είναι συμμετρικό.*

Απόδειξη. Οι προς απόδειξη ισχυρισμοί είναι άμεσες συνέπειες του Θεωρήματος 2.1 αφού

(1) $(A^{\mathrm{T}})^{\mathrm{T}} = A = A^{\mathrm{T}}$

(2) $(A + B)^{\mathrm{T}} = A^{\mathrm{T}} + B^{\mathrm{T}} = A + B$

(3) $(kA)^{\mathrm{T}} = kA^{\mathrm{T}} = kA$

από συμμετρία των A και B. □

Παρατήρηση 2.8. Στη συνέχεια εξετάζουμε αν το γινόμενο συμμετρικών μητρώων είναι συμμετρικό. Έτσι αν A και B είναι $n \times n$ συμμετρικά μητρώα, τότε και πάλι μέσω του Θεωρήματος 2.1 έχουμε ότι

$$(AB)^{\mathrm{T}} = B^{\mathrm{T}}A^{\mathrm{T}} = BA$$

και το τελευταίο μητρώο είναι εν γένει διαφορετικό από το AB. Συμπεραίνουμε λοιπόν ότι: *αν τα $n \times n$ μητρώα A και B είναι συμμετρικά, τότε το AB είναι συμμετρικό αν και μόνο αν τα A και B αντιμετατίθενται, ισχύει δηλαδή ότι $AB = BA$.*

Πρακτικός συμβολισμός

Η πρακτική να γράφουμε τα διανύσματα ως στήλες, δηλαδή ως μητρώα διάστασης $n \times 1$ οδηγεί στον ορισμό-συμβολισμό που ακολουθεί και αποδεικνύεται ιδιαίτερα χρήσιμος στην πράξη.

Ορισμός 2.6 (Συμβολισμός). Αν M είναι ένα τυπικό μητρώο συμβολίζουμε με M_{i*} την i γραμμή του M, ως διάνυσμα-στήλη και με M_{*j} την j στήλη του M ως διάνυσμα.

Έτσι αν

$$M = \begin{pmatrix} 1 & -2 & 3 \\ 2 & 5 & -1 \end{pmatrix}$$

τότε

$$M_{1*} = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}, \qquad M_{2*} = \begin{pmatrix} 2 \\ 5 \\ -1 \end{pmatrix}, \qquad M_{*1} = \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \qquad M_{*2} = \begin{pmatrix} -2 \\ 5 \end{pmatrix}, \qquad M_{*3} = \begin{pmatrix} 3 \\ -1 \end{pmatrix}.$$

Γενικά αν M είναι ένα $n \times m$ μητρώο, τότε έχουμε τους σχετικούς τρόπους αναγραφής των M και M^{T} αντίστοιχα

$$M = \begin{pmatrix} M_{1*}^{\mathrm{T}} \\ M_{2*}^{\mathrm{T}} \\ \vdots \\ M_{n*}^{\mathrm{T}} \end{pmatrix} = \begin{pmatrix} M_{*1} & M_{*2} & \cdots & M_{*m} \end{pmatrix}, \tag{2.19}$$

και

$$M^{\mathrm{T}} = \begin{pmatrix} M_{1*} & M_{2*} & \cdots & M_{n*} \end{pmatrix} = \begin{pmatrix} M_{*1}^{\mathrm{T}} \\ M_{*2}^{\mathrm{T}} \\ \vdots \\ M_{*m}^{\mathrm{T}} \end{pmatrix}. \tag{2.20}$$

Παρατήρηση 2.9 (Αναστροφή και πολλαπλασιασμός). Αν A είναι ένα $n \times m$ μητρώο και B ένα $m \times k$ μητρώο, τότε το γινόμενο AB μπορεί να γραφεί ως

$$AB = \begin{pmatrix} A_{1*}^{\mathrm{T}} B_{*1} & A_{1*}^{\mathrm{T}} B_{*2} & \dots & A_{1*}^{\mathrm{T}} B_{*k} \\ A_{2*}^{\mathrm{T}} B_{*1} & A_{2*}^{\mathrm{T}} B_{*2} & \dots & A_{2*}^{\mathrm{T}} B_{*k} \\ \vdots & \vdots & & \vdots \\ A_{n*}^{\mathrm{T}} B_{*1} & A_{n*}^{\mathrm{T}} B_{*2} & \dots & A_{n*}^{\mathrm{T}} B_{*k} \end{pmatrix} = \left(A_{i*}^{\mathrm{T}} B_{*j}\right). \tag{2.21}$$

Σε σχέση δε με την Παρατήρηση 2.2 έχουμε αφενός

$$AB = A\begin{pmatrix} B_{*1} & B_{*2} & \cdots & B_{*k} \end{pmatrix} = \begin{pmatrix} AB_{*1} & AB_{*2} & \cdots & AB_{*k} \end{pmatrix}. \tag{2.22}$$

και αφετέρου

$$AB = \begin{pmatrix} A_{1*}^{\mathrm{T}} \\ A_{2*}^{\mathrm{T}} \\ \vdots \\ A_{n*}^{\mathrm{T}} \end{pmatrix} B = \begin{pmatrix} A_{1*}^{\mathrm{T}} B \\ A_{2*}^{\mathrm{T}} B \\ \vdots \\ A_{n*}^{\mathrm{T}} B \end{pmatrix}. \tag{2.23}$$

Τέλος η (2.21) γράφεται ως

$$AB = \left(A_{i*} \cdot B_{*j}\right), \tag{2.24}$$

όπου "$\cdot$" είναι το σύνηθες εσωτερικό γινόμενο στο $\mathbb{R}^m$.

2.2.2 Το αντίστροφο μητρώο

Για κάθε πραγματικό ή μιγαδικό αριθμό $a \neq 0$ υπάρχει αντίστοιχα πραγματικός ή μιγαδικός αριθμός b ώστε το γινόμενο των δύο να είναι ίσο με το ουδέτερο στοιχείο του πολλαπλασιασμού, ισοδύναμα $ab = 1$. Το b είναι το αντίστροφο του a. Ισχύει κάτι ανάλογο για μητρώα;

Παράδειγμα 2.8. Για τα μητρώα

$$A = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \qquad \text{και} \qquad B = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix}$$

παρατηρούμε

$$AB = \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I$$

$$BA = \begin{pmatrix} 3 & -5 \\ -1 & 2 \end{pmatrix} \begin{pmatrix} 2 & 5 \\ 1 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = I,$$

κατά συνέπεια το γινόμενο των A, B με όποια σειρά είναι ίσο με το ταυτοτικό μητρώο, το ουδέτερο δηλαδή στοιχείο του πολλαπλασιασμού μητρώων.

Ορισμός 2.7. Έστω ότι το A είναι ένα τετραγωνικό μητρώο. Αν υπάρχει μητρώο B ώστε

$$AB = BA = I$$

λέμε ότι το A είναι **αντιστρέψιμο** και το B θα λέγεται το **αντίστροφο** του A.

Συνεπώς, σε σχέση με το Παράδειγμα 2.8, το μητρώο B είναι το αντίστροφο μητρώο του A, αλλά και και A είναι το αντίστροφο B.

Παράδειγμα 2.9. Εξετάστε αν το μητρώο

$$A = \begin{pmatrix} 2 & 5 & 1 \\ 1 & 3 & 2 \\ 0 & 0 & 0 \end{pmatrix}$$

είναι αντιστρέψιμο.
Για να είναι το A αντιστρέψιμο θα πρέπει να υπάρχει 3×3 μητρώο B, ώστε $AB = BA = I$. Έστω

$$B = \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix}.$$

Εξετάζουμε το γινόμενο AB. Υπολογίζουμε

$$AB = \begin{pmatrix} 2 & 5 & 1 \\ 1 & 3 & 2 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{pmatrix} = \begin{pmatrix} \square & \square & \square \\ \square & \square & \square \\ 0 & 0 & 0 \end{pmatrix} \neq I,$$

κατά συνέπεια το A δεν είναι αντιστρέψιμο.

► **Άσκηση 2.7.** Σε σχέση με το Παράδειγμα 2.9, δείξτε ότι αν ένα 3×3 μητρώο έχει μια μηδενική στήλη, ας πούμε τη δεύτερη, δεν είναι αντιστρέψιμο.

Παρατήρηση 2.10. Γενικεύοντας, αν το $n \times n$ μητρώο A έχει μια μηδενική γραμμή, ή μια μηδενική στήλη, δεν είναι αντιστρέψιμο.

Παράδειγμα 2.10. Αν a και b είναι μη μηδενικοί αριθμοί εξετάστε αν το μητρώο A αντιστρέφεται, όπου

$$A = \begin{pmatrix} a & a \\ b & b \end{pmatrix}$$

Ας υποθέσουμε ότι υπάρχει 2×2 μητρώο B ώστε $AB = I$, και έστω

$$B = \begin{pmatrix} p & q \\ r & s \end{pmatrix},$$

τότε θα πρέπει

$$AB = \begin{pmatrix} a & a \\ b & b \end{pmatrix}\begin{pmatrix} p & q \\ r & s \end{pmatrix} = \begin{pmatrix} a(p+r) & a(q+s) \\ b(p+r) & b(q+s) \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Το αποτέλεσμα είναι αδύνατο αφού την ίδια στιγμή θα πρέπει να ισχύει αφενός $a(p+r) = 1$ και αφετέρου $p + r = 0$, καθόσον $b \neq 0$. Καταλήξαμε σε άτοπο επειδή υποθέσαμε ότι υπάρχει το αντίστροφο του A, κατά συνέπεια το μητρώο δεν έχει αντίστροφο.

Θεώρημα 2.3. *Το αντίστροφο μητρώο, αν υπάρχει, είναι μοναδικό.*

Απόδειξη. Έστω A ένα αντιστρέψιμο μητρώο και έστω ότι B και C είναι τετραγωνικά μητρώα τέτοια ώστε $AB = BA = I$ και $AC = CA = I$. Τα B, C έχουν την ίδια διάσταση, επιπλέον από τις ιδιότητες του γινομένου και του αντιστρόφου έχουμε

$$B = BI = B(AC) = (BA)C = IC = C$$

γεγονός που αποδεικνύει το ζητούμενο. □

Ορισμός 2.8. Το αντίστροφο μητρώο του τετραγωνικού μητρώου A, εφόσον αυτό υπάρχει, συμβολίζουμε με A^{-1}.

► **Άσκηση 2.8.** Έστω

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}.$$

Αν $ad - bc \neq 0$, δείξτε ότι

$$A^{-1} = \frac{1}{J}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}, \qquad \text{όπου} \qquad J = ad - bc.$$

► **Άσκηση 2.9.** Εάν A είναι ένα τετραγωνικό μητρώο για το οποίο υπάρχει το A^{-1}

(α΄) Το A^{-1} είναι αντιστρέψιμο και $(A^{-1})^{-1} = A$.

(β΄) Εάν $\lambda \neq 0$, το λA είναι αντιστρέψιμο και

$$(\lambda A)^{-1} = \frac{1}{\lambda} A^{-1}.$$

Θεώρημα 2.4. *Εάν A και B είναι αντιστρέψιμα μητρώα ίδιας διάστασης, τότε το μητρώο AB είναι αντιστρέψιμο και $(AB)^{-1} = B^{-1}A^{-1}$.*

Απόδειξη. Παρατηρούμε ότι

$$\begin{aligned}(AB)(B^{-1}A^{-1}) &= A(BB^{-1})A^{-1} = AIA^{-1} = AA^{-1} = I\\ (B^{-1}A^{-1})(AB) &= B^{-1}(A^{-1}A)B = B^{-1}IB = B^{-1}B = I,\end{aligned}$$

κατά συνέπεια το μητρώο AB είναι αντιστρέψιμο και από τη μοναδικότητα του αντίστροφου έπεται ότι $(AB)^{-1} = B^{-1}A^{-1}$. □

Παρατήρηση 2.11. Εάν το τετραγωνικό μητρώο A είναι αντιστρέψιμο από το Θεώρημα 2.4 έπεται ότι το A^2 είναι αντιστρέψιμο και $(A^2)^{-1} = (A^{-1})^2$. Έτσι επαγωγικά έχουμε ότι το A^n είναι αντιστρέψιμο και $(A^n)^{-1} = (A^{-1})^n$ για $n = 1, 2, \ldots$, κατά συνέπεια μπορούμε να ορίσουμε τις αρνητικές δυνάμεις

$$A^{-n} = (A^n)^{-1} = (A^{-1})^n, \qquad n = 1, 2, \ldots.$$

Έτσι έχουμε ορίσει τις ακέραιες δυνάμεις του A με τη σχέση

$$A^0 = I, \qquad A^n = \begin{cases} \underbrace{AA\cdots A}_{n \text{ φορές}} & n = 1, 2\ldots \\ (A^{-1})^{-n} & n = -1, -2, \ldots \end{cases}$$

Εύκολα επαληθεύεται ότι εάν k και l είναι ακέραιοι αριθμοί, τότε

$$A^k A^l = A^{k+l}, \qquad \text{και} \qquad (A^k)^l = A^{kl}.$$

Επίσης από τη σχέση

$$(A^{-1})^{\mathrm{T}} A^{\mathrm{T}} = (AA^{-1})^{\mathrm{T}} = I = (A^{-1}A)^{\mathrm{T}} = A^{\mathrm{T}}(A^{-1})^{\mathrm{T}}$$

συμπεραίνουμε ότι αν το A είναι αντιστρέψιμο, τότε και το A^{T} αντιστρέφεται και $(A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}$, από την μοναδικότητα του αντίστροφου. Έτσι μπορούμε να ορίσουμε

$$A^{-\mathrm{T}} = (A^{\mathrm{T}})^{-1} = (A^{-1})^{\mathrm{T}}.$$

Σημείωση 2.2. Σύμφωνα με τον ορισμό το $n \times n$ μητρώο A έχει αντίστροφο αν υπάρχει $n \times n$ μητρώο B ώστε $AB = I$ και $BA = I$. Στη πραγματικότητα αποδεικνύεται, και θα το δείξουμε σε επόμενο κεφάλαιο όταν θα έχουμε αναπτύξει τα κατάλληλα εργαλεία, ότι αν ισχύει η μία από τις δύο συνθήκες, δηλαδή $AB = I$ ή $BA = I$, τότε αναγκαστικά θα ισχύει και η άλλη, συγκεκριμένα θα αποδείξουμε το: **Θεώρημα.** *Αν για το $n \times n$ μητρώο A υπάρχει $n \times n$ μητρώο B ώστε $AB = I$, τότε θα είναι $BA = I$, συνεπώς το A αντιστρέφεται και $A^{-1} = B$.*

2.3 Μητρώα ειδικής μορφής

Λέγοντας μητρώα ειδικής μορφής εννοούμε τετραγωνικά, κυρίως, μητρώα των οποίων η μορφή έχει συγκεκριμένη χαρακτηριστική δομή. Ένα τέτοιο μητρώο είναι για παράδειγμα το ταυτοτικό μητρώο, το οποίο αν το δούμε σαν μια τετραγωνική διάταξη με πλευρές παράλληλες στους άξονες και κέντρο στην αρχή των αξόνων, τότε όλα τα στοιχεία του είναι μηδενικά εκτός από αυτά που βρίσκονται στη διαγώνιο από την κορυφή στο δεύτερο τεταρτημόριο προς αυτή στο τέταρτο καθένα από τα οποία είναι ίσο με ένα.

Ορισμός 2.9. Έστω $A = \left(a_{ij}\right)$ ένα τετραγωνικό μητρώο.

(1) Λέγοντας **κύρια διαγώνιο** του A εννοούμε τη "διαγώνιο" του A η οποία αποτελείται από τα στοιχεία a_{ii}. Τα στοιχεία a_{ii} τα λέμε **διαγώνια στοιχεία**.

(2) Το A λέγεται **διαγώνιο** μητρώο (diagonal) αν $a_{ij} = 0$ όταν $i \neq j$.

(3) Το A λέγεται **άνω τριγωνικό** μητρώο (upper triangular) αν $a_{ij} = 0$ όταν $i > j$, δηλαδή τα στοιχεία κάτω από την κύρια διαγώνιο είναι μηδενικά.

(4) Το A λέγεται **κάτω τριγωνικό** μητρώο (lower triangular) αν $a_{ij} = 0$ όταν $i < j$, δηλαδή τα στοιχεία πάνω από την κύρια διαγώνιο είναι μηδενικά.

► **Άσκηση 2.10.** Δείξτε ότι το γινόμενο διαγωνίων μητρώων είναι επίσης διαγώνιο και στη συνέχεια διατυπώστε ένα κανόνα για τον πολλαπλασιασμό διαγωνίων μητρώων. Χρησιμοποιήστε αυτό το αποτέλεσμα για να δείξετε ότι αν

$$A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}, \qquad ab \neq 0,$$

τότε

$$A^{-1} = \begin{pmatrix} 1/a & 0 \\ 0 & 1/b \end{pmatrix}.$$

Συγκρίνετε με το αποτέλεσμα της Άσκησης 2.8.

Παράδειγμα 2.11. Θεωρούμε τα άνω τριγωνικά μητρώα

$$A = \begin{pmatrix} a_1 & a_2 & a_3 \\ 0 & a_4 & a_5 \\ 0 & 0 & a_6 \end{pmatrix} \qquad \text{και} \qquad B = \begin{pmatrix} b_1 & b_2 & b_3 \\ 0 & b_4 & b_5 \\ 0 & 0 & b_6 \end{pmatrix}.$$

Οι γραμμικοί συνδυασμοί όσο και το γινόμενο των A και B είναι άνω τριγωνικά μητρώα. Πράγματι αν r και s είναι πραγματικές ή μιγαδικές σταθερές, τότε

$$rA + sB = r\begin{pmatrix} a_1 & a_2 & a_3 \\ 0 & a_4 & a_5 \\ 0 & 0 & a_6 \end{pmatrix} + s\begin{pmatrix} b_1 & b_2 & b_3 \\ 0 & b_4 & b_5 \\ 0 & 0 & b_6 \end{pmatrix} = \begin{pmatrix} ra_1 + sb_1 & ra_2 + sb_2 & ra_3 + sb_3 \\ 0 & ra_4 + sb_4 & ra_5 + sb_5 \\ 0 & 0 & ra_6 + sb_6 \end{pmatrix},$$

το οποίο είναι άνω τριγωνικό. Όμοια

$$AB = \begin{pmatrix} a_1 & a_2 & a_3 \\ 0 & a_4 & a_5 \\ 0 & 0 & a_6 \end{pmatrix}\begin{pmatrix} b_1 & b_2 & b_3 \\ 0 & b_4 & b_5 \\ 0 & 0 & b_6 \end{pmatrix} = \begin{pmatrix} a_1b_1 & a_1b_2 + a_2b_4 & a_1b_3 + a_2b_5 + a_3b_6 \\ 0 & a_4b_4 & a_4b_5 + a_5b_6 \\ 0 & 0 & a_6b_6 \end{pmatrix},$$

είναι επίσης άνω τριγωνικό.

Ορισμός 2.10. Ενα $n \times n$ διαγώνιο μητρώο A συμβολίζουμε ως

$$A = \mathrm{diag}\begin{pmatrix} a_{11} & a_{22} & \cdots & a_{nn} \end{pmatrix}.$$

2.3.1 Μητρώα μετάθεσης

Ας θεωρήσουμε ένα 3×3 μητρώο το οποίο προκύπτει από το ταυτοτικό μητρώο I μεταθέτοντας κάποιες γραμμές του. Ένα τέτοιο μητρώο θα το λέμε **μητρώο μετάθεσης**. Για παράδειγμα το μητρώο

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

προκύπτει μεταθέτοντας τις δύο πρώτες γραμμές του I. Στη συνέχεια ας δούμε το αποτέλεσμα του πολλαπλασιασμού ενός τέτοιου μητρώου με ένα διάνυσμα,

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} y \\ x \\ z \end{pmatrix}. \tag{2.25}$$

Το γινόμενο λοιπόν είναι το διάνυσμα που προκύπτει από το αρχικό αν μεταθέσουμε τις ίδιες γραμμές που μετατέθηκαν στο I για να προκύψει το μητρώο μετάθεσης. Από τον ορισμό του πολλαπλασιασμού μητρώων το ανάλογο αποτέλεσμα συμβαίνει σε κάθε στήλη ενός $3 \times m$ μητρώου αν πολλαπλασιαστεί από τα αριστερά με το μητρώο μετάθεσης (γιατί;). Τα 3×3 μητρώα μετάθεσης είναι $3! = 6$, τα

$$M_0 = I = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \qquad M_2 = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix},$$
$$M_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}, \qquad M_4 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \qquad M_5 = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}.$$

Παρατηρούμε ότι $M_1^{\mathrm{T}} = M_1$, $M_2^{\mathrm{T}} = M_2$, $M_3^{\mathrm{T}} = M_4$, $M_4^{\mathrm{T}} = M_3$, και $M_5^{\mathrm{T}} = M_5$. Το γινόμενο δύο μητρώων μετάθεσης M_iM_j έχει ως αποτέλεσμα την μετάθεση δύο γραμμών του M_j, κατά συνέπεια το γινόμενο M_iM_j είναι μητρώο μετάθεσης. Ενδεικτικά υπολογίζουμε

$$M_1M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I$$
$$M_3M_3 = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix} = M_4$$
$$M_4M_3 = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = I.$$

Παρατήρηση 2.12. Αν M είναι ένα 3×3 μητρώο μετάθεσης, τότε

$$M = \begin{pmatrix} \boldsymbol{e}_i & \boldsymbol{e}_j & \boldsymbol{e}_k \end{pmatrix} = \begin{pmatrix} \boldsymbol{e}_r^{\mathrm{T}} \\ \boldsymbol{e}_s^{\mathrm{T}} \\ \boldsymbol{e}_t^{\mathrm{T}} \end{pmatrix}$$

με $i, j, k \in \{1,2,3\}$ ανά δύο διαφορετικά μεταξύ τους και $r, s, t \in \{1,2,3\}$ επίσης ανά δύο διαφορετικά μεταξύ τους. Συνέπεια αυτού του αποτελέσματος είναι ότι

$$M^{\mathrm{T}} M = M M^{\mathrm{T}} = \begin{pmatrix} \boldsymbol{e}_i^{\mathrm{T}} \boldsymbol{e}_j \end{pmatrix} = I, \qquad i = 1,2,3 \quad j = 1,2,3$$

κατά συνέπεια $M_k^{-1} = M_k^{\mathrm{T}}$, για $k = 0, 1, \ldots, 5$. Τα μητρώα μετάθεσης αποτελούν μια ειδική κατηγορία των **ορθογωνίων** μητρώων που θα μελετήσουμε στην παράγραφο 5.6.

▶ Άσκηση 2.11. Για τα μητρώα μετάθεσης δείξτε ότι

(1) $M_k^{-1} = M_k^{\mathrm{T}}$, για $k = 0, 1, \ldots, 5$.

(2) Εφοδιασμένα με την πράξη του πολλαπλασιασμού αποτελούν ομάδα.

2.4 Διαμερισμένα μητρώα

Πολλές φορές ένα μητρώο μπορεί να θεωρηθεί ότι απαρτίζεται από άλλα μητρώα μικρότερου μεγέθους, ή ισοδύναμα ότι μπορεί να διαχωριστεί ή διαμεριστεί σε υπομητρώα όπως τα λέμε. Σχηματικά η διαμέριση αυτή επιτυγχάνεται με χρήση οριζόντιων και/ή κάθετων ευθειών που διατρέχουν ολόκληρο το μητρώο, και όχι ένα μόνο μέρος αυτού, διαχωρίζοντάς το σε τμήματα. Για παράδειγμα το

$$M = \begin{pmatrix} 3 & -2 & 1 & 0 \\ 2 & -1 & 0 & 1 \\ 1 & 10 & 0 & 2 \end{pmatrix}$$

μπορεί να ειδωθεί ως

$$M = \left(\begin{array}{cc|cc} 3 & -2 & 1 & 0 \\ 2 & -1 & 0 & 1 \\ \hline 1 & 10 & 0 & 2 \end{array}\right) = \begin{pmatrix} A & I \\ B & C \end{pmatrix} \tag{2.26}$$

όπου

$$A = \begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix}, \qquad B = \begin{pmatrix} 1 & 10 \end{pmatrix}, \qquad C = \begin{pmatrix} 0 & 2 \end{pmatrix}$$

και I το ταυτοτικό 2×2 μητρώο. Η διαμέριση ενός μητρώου σε υπομητρώα δεν είναι μοναδική, μπορούμε για παράδειγμα να γράψουμε

$$M = \left(\begin{array}{c|ccc} 3 & -2 & 1 & 0 \\ \hline 2 & -1 & 0 & 1 \\ 1 & 10 & 0 & 2 \end{array}\right) = \begin{pmatrix} E & F \\ G & H \end{pmatrix}$$

όπου

$$E = \begin{pmatrix} 3 \end{pmatrix}, \qquad F = \begin{pmatrix} -2 & 1 & 0 \end{pmatrix}, \qquad G = \begin{pmatrix} 2 \\ 1 \end{pmatrix}, \qquad H = \begin{pmatrix} -1 & 0 & 1 \\ 10 & 0 & 2 \end{pmatrix}.$$

Η επιλογή της διαμέρισης υπαγορεύεται από τον σκοπό για τον οποίο την επιβάλλουμε σ' ένα μητρώο. Στη περίπτωση που ένα μητρώο είναι διαμερισμένο, ή προκύπτει από παράθεση άλλων μητρώων συμβατών διαστάσεων θα λέμε το αρχικό μητρώο **σύνθετο**. Στην Παρατήρηση 2.2 είδαμε ότι ένα $n \times m$ μητρώο A μπορεί να διαμεριστεί κατά γραμμές ή στήλες ως

$$A = \left(\begin{array}{cccc} a_{11} & a_{12} & \cdots & a_{1m} \\ \hline a_{21} & a_{22} & \cdots & a_{2m} \\ \hline \vdots & \vdots & & \vdots \\ \hline a_{n1} & a_{n2} & \cdots & a_{nm} \end{array}\right) = \begin{pmatrix} A^1 \\ A^2 \\ \vdots \\ A^n \end{pmatrix} \qquad \text{και} \qquad A = \left(\begin{array}{c|c|c|c} a_{11} & a_{12} & \cdots & a_{1m} \\ a_{21} & a_{22} & \cdots & a_{2m} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nm} \end{array}\right) = \begin{pmatrix} A_1 & A_2 & \cdots & A_m \end{pmatrix}$$

όπου

$$A^i = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{im} \end{pmatrix} \qquad \text{και} \qquad A_j = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix}$$

με $i = 1, 2, \dots, n$ και $j = 1, 2, \dots, m$.

Πράξεις με διαμερισμένα μητρώα

• **Πρόσθεση και πολλαπλασιασμός με σταθερά**. Εάν A και B είναι $n \times m$ μητρώα διαμερισμένα ως $A = \left(A_{ij}\right)$ και $B = \left(B_{ij}\right)$ με $i = 1, 2, \dots, r$ και $j = 1, 2, \dots, s$, όπου $r \leq n$ και $s \leq m$, ώστε το A_{ij} και B_{ij} είναι του ιδίου μεγέθους, τότε όπως θα περιμέναμε

$$A + B = \left(A_{ij} + B_{ij}\right)$$

και για σταθερές λ και μ

$$\lambda A + \mu B = \left(\lambda A_{ij} + \mu B_{ij}\right).$$

• **Πολλαπλασιασμός**. Εάν C είναι ένα $m \times l$ μητρώο, τότε το γινόμενο, όπως γνωρίζουμε, κωδικοποιείται ως

$$AC = \left(A^i C_j\right), \qquad i = 1, 2, \dots, n \quad j = 1, 2, \dots, l$$

όπου A^i είναι τα υπομητρώα-γραμμές του A, και C_j τα υπομητρώα-στήλες του B.

Παράδειγμα 2.12. Ας θεωρήσουμε το διαμερισμένο μητρώο

$$N = \left(\begin{array}{cc|c} 3 & -2 & 1 \\ 2 & -1 & 0 \\ \hline 1 & 10 & 2 \end{array}\right) = \begin{pmatrix} A & B \\ C & D \end{pmatrix}.$$

Το μητρώο N^2 ορίζεται. Ας δούμε αν μέσω της διαμέρισης μπορούμε να πάρουμε, όπως θα φαινόταν λογικό, ότι

$$N^2 = \begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A^2 + BC & AB + BD \\ CA + DC & CB + D^2 \end{pmatrix}. \tag{2.27}$$

Για να ισχύει το αποτέλεσμα αυτό θα πρέπει αρχικά οι πράξεις μεταξύ των σχετικών υπομητρώων στο τελευταίο μητρώο να ορίζονται. Πράγματι

$$A^2 + BC = \begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix}^2 + \begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 1 & 10 \end{pmatrix} = \begin{pmatrix} 5 & -4 \\ 4 & -3 \end{pmatrix} + \begin{pmatrix} 1 & 10 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 6 & 6 \\ 4 & -3 \end{pmatrix}$$

$$AB + BD = \begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix}\begin{pmatrix} 2 \end{pmatrix} = \begin{pmatrix} 3 \\ 2 \end{pmatrix} + \begin{pmatrix} 2 \\ 0 \end{pmatrix} = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$$

$$CA + DC = \begin{pmatrix} 1 & 10 \end{pmatrix}\begin{pmatrix} 3 & -2 \\ 2 & -1 \end{pmatrix} + \begin{pmatrix} 2 \end{pmatrix}\begin{pmatrix} 1 & 10 \end{pmatrix} = \begin{pmatrix} 23 & -12 \end{pmatrix} + \begin{pmatrix} 2 & 20 \end{pmatrix} = \begin{pmatrix} 25 & 8 \end{pmatrix}$$

$$CB + D^2 = \begin{pmatrix} 1 & 10 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} + \begin{pmatrix} 2 \end{pmatrix}^2 = \begin{pmatrix} 1 \end{pmatrix} + \begin{pmatrix} 4 \end{pmatrix} = \begin{pmatrix} 5 \end{pmatrix}.$$

Αν με N_1, N_2, N_3 συμβολίσουμε τις στήλες του $N = \begin{pmatrix} n_{ij} \end{pmatrix}$ και με N_1', N_2', N_3' τις στήλες του N^2, από την (2.13) βρίσκουμε

$$N_1' = N_1 n_{11} + N_2 n_{21} + N_3 n_{31} = \begin{pmatrix} a_{11} \\ a_{21} \\ c_{11} \end{pmatrix} a_{11} + \begin{pmatrix} a_{12} \\ a_{22} \\ c_{12} \end{pmatrix} a_{21} + \begin{pmatrix} b_{11} \\ b_{21} \\ d_{11} \end{pmatrix} c_{11} = \begin{pmatrix} A_1 \\ C_1 \end{pmatrix} a_{11} + \begin{pmatrix} A_2 \\ C_2 \end{pmatrix} a_{21} + \begin{pmatrix} B_1 \\ D_1 \end{pmatrix} c_{11} \quad (2.28)$$

Θέλουμε να συγκρίνουμε την N_1' με την πρ΄τη στήλη του μητρώου στα δεξιά στην (2.27). Συμβολίζοντας με A_1', A_2' τις στήλες του A^2 η πρώτη στήλη του $A^2 + BC$ είναι

$$A_1 a_{11} + A_2 a_{21} + B_1 c_{11}.$$

Όμοια η πρώτη στήλη του $CA + DC$ είναι

$$C_1 a_{11} + C_2 a_{21} + D_1 c_{11},$$

κατά συνέπεια η ζητούμενη πρώτη στήλη είναι η

$$\begin{pmatrix} A^2 + BC \\ CA + DC \end{pmatrix}_1 = \begin{pmatrix} A_1 a_{11} + A_2 a_{21} + B_1 c_{11} \\ C_1 a_{11} + C_2 a_{21} + D_1 c_{11} \end{pmatrix} = \begin{pmatrix} A_1 \\ C_1 \end{pmatrix} a_{11} + \begin{pmatrix} A_2 \\ C_2 \end{pmatrix} a_{21} + \begin{pmatrix} B_1 \\ D_1 \end{pmatrix} c_{11}$$

αποτέλεσμα το οποίο συμφωνεί με αυτό στην (2.28). Όμοια αποδεικνύεται το ανάλογο αποτέλεσμα για τις υπόλοιπες στήλες. Συνεπώς η ισότητα στην (2.27) ισχύει, επομένως

$$N^2 = \left(\begin{array}{cc|c} 6 & 6 & 5 \\ 4 & -3 & 2 \\ \hline 25 & 8 & 5 \end{array}\right).$$

▶ **Άσκηση 2.12.** Εάν τα υπομητρώα A, B, C, D, P, Q, R, S είναι διαστάσεων

A	B	C	D	P	Q	R	S
$n_1 \times m_1$	$n_1 \times m_2$	$n_2 \times m_1$	$n_2 \times m_2$	$m_1 \times k_1$	$m_1 \times k_2$	$m_2 \times k_1$	$m_2 \times k_2$

αποδείξτε ότι η πράξη του πολλαπλασιασμού υλοποιείται όπως στο σχήμα

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}\begin{pmatrix} P & Q \\ R & S \end{pmatrix} = \begin{pmatrix} AP + BR & AQ + BS \\ CP + DR & CQ + DS \end{pmatrix}. \quad (2.29)$$

► **Άσκηση 2.13.** Δίνονται τα μητρώα

$$A=\left(\begin{array}{c|cc} a_{11} & a_{12} & a_{13}\\ \cline{2-3} a_{21} & a_{22} & a_{23}\\ a_{31} & a_{32} & a_{33}\end{array}\right),\quad B=\left(\begin{array}{cc|c} b_{11} & b_{12} & b_{13}\\ \hline b_{21} & b_{22} & b_{23}\\ b_{31} & b_{32} & b_{33}\end{array}\right),\quad C=\left(\begin{array}{c|cc} c_{11} & c_{12} & c_{13}\\ c_{21} & c_{22} & c_{23}\\ \hline c_{31} & c_{32} & c_{33}\end{array}\right),\quad D=\left(\begin{array}{cc|c} d_{11} & d_{12} & d_{13}\\ d_{21} & d_{22} & d_{23}\\ \hline d_{31} & d_{32} & d_{33}\end{array}\right).$$

Ενώ το γινόμενο οποιωνδήποτε δύο από αυτά ορίζεται, εξετάστε ποια από τα γινόμενα AA, AB, AC, AD, BA, BB, BC, BD, CA, CB, CC, CD, DA, DB, DC, DD μπορούν να εκφραστούν μέσω των σχετικών πράξεων των επιμέρους υπομητρώων όπως στην (2.29).

- **Αναστροφή**. Αν M είναι όπως στην (2.26), τότε

$$M^{\mathrm{T}}=\begin{pmatrix} A & I\\ B & C\end{pmatrix}^{\mathrm{T}}=\left(\begin{array}{cc|cc} 3 & -2 & 1 & 0\\ 2 & -1 & 0 & 1\\ \hline 1 & 10 & 0 & 2\end{array}\right)^{\mathrm{T}}=\left(\begin{array}{cc|c} 3 & 2 & 1\\ -2 & -1 & 10\\ \hline 1 & 0 & 0\\ 0 & 1 & 2\end{array}\right)=\begin{pmatrix} A^{\mathrm{T}} & B^{\mathrm{T}}\\ I^{\mathrm{T}} & C^{\mathrm{T}}\end{pmatrix}.$$

Το αποτέλεσμα γενικεύεται, έτσι

$$A^{\mathrm{T}}=\begin{pmatrix} A_{11} & A_{12} & \cdots & A_{1m}\\ A_{21} & A_{22} & \cdots & A_{2m}\\ \vdots & \vdots & & \vdots\\ A_{n1} & A_{n2} & \cdots & A_{nm}\end{pmatrix}^{\mathrm{T}}=\begin{pmatrix} A_{11}^{\mathrm{T}} & A_{21}^{\mathrm{T}} & \cdots & A_{n1}^{\mathrm{T}}\\ A_{12}^{\mathrm{T}} & A_{22}^{\mathrm{T}} & \cdots & A_{n2}^{\mathrm{T}}\\ \vdots & \vdots & & \vdots\\ A_{1m}^{\mathrm{T}} & A_{2m}^{\mathrm{T}} & \cdots & A_{nm}^{\mathrm{T}}\end{pmatrix}.$$

- **Αντιστροφή**. Αν για κάποιο τυπικό μητρώο υπάρχει το αντίστροφο, δεν υπάρχει κάποια κατάλληλη διαμέριση και κάποιος απλός τύπος που να εκφράζει το αντίστροφο μέσω των υπομητρώων που συνθέτουν τη διαμέριση και των αντιστόφων τους[4]. Όμως σε κάποιες περιπτώσεις η διαμέριση σε τετραγωνικά μητρώα ειδικής μορφής επιτρέπει την εύρεση του αντιστρόφου. Ας θεωρήσουμε την περίπτωση ενός τριγωνικού τύπου διαμερισμένου μητρώου

$$T=\begin{pmatrix} A & B\\ O & C\end{pmatrix},$$

όπου τα A, B, C είναι τετραγωνικά μητρώα και O είναι το μηδενικό μητρώο. Αναζητούμε μητρώο S ώστε $TS=I$. Επειδή το γινόμενο είναι διαγώνιο, άρα τριγωνικό και το γινόμενο τριγωνικών μητρώων είναι τριγωνικό, αναζητούμε το S σε ανάλογη μορφή και θεωρούμε την εξίσωση

$$\begin{aligned} TS=I &\Leftrightarrow \begin{pmatrix} A & B\\ O & C\end{pmatrix}\begin{pmatrix} A' & B'\\ O & C'\end{pmatrix}=\begin{pmatrix} I & O\\ O & I\end{pmatrix}\\ &\Leftrightarrow \begin{pmatrix} AA' & AB'+BC'\\ O & CC'\end{pmatrix}=\begin{pmatrix} I & O\\ O & I\end{pmatrix}\end{aligned}$$

[4] Μπορεί να αποδειχθεί ότι εάν τα μητρώα A και D είναι $n\times n$ και $m\times m$ αντίστοιχα, το B είναι $n\times k$, το C είναι $m\times k$ και το $S=D-CA^{-1}B$, ή το $T=A-BD^{-1}C$ υπάρχει, τότε

$$\begin{pmatrix} A & B\\ C & D\end{pmatrix}^{-1}=\begin{pmatrix} A^{-1}+A^{-1}BS^{-1}CA^{-1} & -A^{-1}BS^{-1}\\ -S^{-1}CA^{-1} & S^{-1}\end{pmatrix},\qquad \text{ή}\qquad \begin{pmatrix} A & B\\ C & D\end{pmatrix}^{-1}=\begin{pmatrix} T^{-1} & -T^{-1}BD^{-1}\\ -D^{-1}CT^{-1}C)^{-1} & D^{-1}+D^{-1}CT^{-1}BD^{-1}\end{pmatrix},$$

οποτεδήποτε τα μητρώα που εμφανίζονται στον τύπο αντεστραμένα υπάρχουν. Τα μητρώα S και T λέγονται **συμπληρώματα του Schur**. Βλέπε σχετικά [8] απ΄ όπου επιλύοντας βρίσκουμε

$$\begin{aligned} AA' = I &\Rightarrow A' = A^{-1} \\ AB' + BC' = O &\Rightarrow B' = -A^{-1}BC^{-1} \\ CC' = I &\Rightarrow C' = C^{-1} \end{aligned}$$

υποθέτοντας βεβαίως ότι τα A και C αντιστρέφονται. Συνεπώς σε αυτή την περίπτωση

$$\begin{pmatrix} A & B \\ O & C \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & -A^{-1}BC^{-1} \\ O & C^{-1} \end{pmatrix}.$$

Διαμερισμένο διαγώνιο μητρώο. Ειδικά αν τα μητρώα A και B είναι αντιστρέψιμα, τότε

$$\begin{pmatrix} A & O \\ O & B \end{pmatrix}^{-1} = \begin{pmatrix} A^{-1} & O \\ O & B^{-1} \end{pmatrix}.$$

2.5 Ασκήσεις

1. Εάν

$$A = \begin{pmatrix} 1 & -2 \\ 2 & -1 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & 1 & 3 \\ 3 & -1 & 2 \end{pmatrix}, \quad C = \begin{pmatrix} 1 & -1 & 3 \\ -2 & 0 & -2 \end{pmatrix}, \quad D = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$$

να υπολογισθούν εφόσον ορίζονται τα μητρώα:

(α΄) $A(2B - 3C)$.
(β΄) CD.
(γ΄) $(2B - 3C)D$.
(δ΄) $(AB)D$.
(ε΄) $A(BD)$.
(ϛ΄) $C^{\mathrm{T}}A$.
(ζ΄) $D^{\mathrm{T}}C^{\mathrm{T}}$.
(η΄) $B^{\mathrm{T}}C$.
(θ΄) DD^{T}.

2. Εάν

$$A = \begin{pmatrix} 1 & -2 \\ -2 & 3 \\ 1 & -3 \end{pmatrix}, \quad B = \begin{pmatrix} 2 & -1 & 2 \\ 3 & 0 & 1 \end{pmatrix}, \quad C = \begin{pmatrix} -5 & -8 & 9 \\ 4 & 13 & -16 \\ -11 & -11 & 11 \end{pmatrix}$$

να βρεθεί μητρώο M ώστε $AMB = C$.

3. Αν

$$A = \begin{pmatrix} 0 & a_1 \\ 0 & 0 \end{pmatrix}, \quad \text{και} \quad B = \begin{pmatrix} 0 & b_1 & b_2 \\ 0 & 0 & b_3 \\ 0 & 0 & 0 \end{pmatrix},$$

δείξτε ότι $A^2 = O$, και $B^3 = O$.

4. Να βρεθούν όλα τα μητρώα τα οποία μετατίθενται με το μητρώο A σε κάθε μια από τις περιπτώσεις:

(α΄) $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}$.
(β΄) $A = \begin{pmatrix} a & 0 \\ 0 & b \end{pmatrix}$, $a \neq b$.
(γ΄) $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}$.
(δ΄) Το A είναι ένα τυχαίο $n \times n$ μητρώο.

5. Η ακολουθία των αριθμών **Fibonacci** ορίζεται αναδρομικά από τις σχέσεις

$$f_1 = 1, \quad f_2 = 1, \quad f_n = f_{n-1} + f_{n-2}, \quad n \geq 3,$$

έτσι οι πρώτοι μερικοί όροι της ακολουθίας είναι

$$1, 1, 2, 3, 5, 8, 13, 21, 34, 55, 89, 144, \dots$$

Έστω

$$F = \begin{pmatrix} f_2 & f_1 \\ f_1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 0 \end{pmatrix}.$$

Θέτοντας $f_0 = 0$, δείξτε ότι για κάθε $n = 1, 2, 3, \dots$ ισχύει

$$F^n = \begin{pmatrix} f_{n+1} & f_n \\ f_n & f_{n-1} \end{pmatrix}.$$

6. Να βρεθούν όλα τα 2×2 άνω τριγωνικά μητρώα

$$M = \begin{pmatrix} a & b \\ 0 & c \end{pmatrix}, \qquad b \neq 0$$

τέτοια ώστε $M^2 = I$.

7. Έστω $F = \left(f_{ij}\right)$ ένα $n \times m$ μητρώο του οποίου τα στοιχεία f_{ij} είναι παραγωγίσιμες συναρτήσεις σε κάποιο διάστημα. Ένα τέτοιο μητρώο θα το λέμε παραγωγίσιμο. Ορίζουμε την παράγωγο του F με τη σχέση

$$\frac{\mathrm{d}F}{\mathrm{d}x} = \left(f'_{ij}(x)\right).$$

(α΄) Εάν F και G είναι $n \times m$ παραγωγίσιμα μητρώα και λ, μ είναι σταθερές δείξτε ότι

$$\frac{\mathrm{d}}{\mathrm{d}x}(\lambda F + \mu G) = \lambda \frac{\mathrm{d}F}{\mathrm{d}x} + \mu \frac{\mathrm{d}G}{\mathrm{d}x}.$$

(β΄) Εάν F και G είναι $n \times m$ και $m \times k$ αντίστοιχα, παραγωγίσιμα μητρώα δείξτε ότι

$$\frac{\mathrm{d}}{\mathrm{d}x}(FG) = \left(\frac{\mathrm{d}F}{\mathrm{d}x}\right)G + F\left(\frac{\mathrm{d}G}{\mathrm{d}x}\right).$$

(γ΄) Εάν F είναι ένα παραγωγίσιμο τετραγωνικό μητρώο για το οποίο υπάρχει το αντίστροφο μητρώο F^{-1}, μπορεί να αποδειχθεί ότι το F^{-1} είναι παραγωγίσιμο[5]. Δείξτε ότι

$$\frac{\mathrm{d}}{\mathrm{d}x}F^{-1} = F^{-1}\left(\frac{\mathrm{d}F}{\mathrm{d}x}\right)F^{-1}.$$

[5]Στη περίπτωση ενός 2×2 αντιστρέψιμου μητρώου

$$F = \begin{pmatrix} f_{11} & f_{12} \\ f_{21} & f_{22} \end{pmatrix}$$

είναι

$$F^{-1} = \frac{1}{f_{11}f_{22} - f_{12}f_{21}} \begin{pmatrix} f_{22} & -f_{12} \\ -f_{21} & f_{11} \end{pmatrix}$$

με $f_{11}(x)f_{22}(x) - f_{12}(x)f_{21}(x) \neq 0$ για όλα τα x στο διάστημα ορισμού. Έτσι αν το F είναι παραφωγίσιμο τα στοιχεία του F^{-1} είναι παραγωγίσιμες συναρτήσεις, κατά συνέπεια το F^{-1} είναι παραγωγίσιμο.

8. Οι έγχρωμοι εκτυπωτές χρησιμοποιούν τρία χρώματα για έγχρωμες εκτυπώσεις τα: βιολετί ή φούξια (magenta), κίτρινο (yellow) και κυανό (cyan). Τα χρώματα αυτά m, y και c συνδέονται με τα βασικά χρώματα r, g και b του χρωματικού μοντέλου rgb με τη σχέση, βλέπε Σχήμα 1.6
$$\begin{pmatrix} m \\ y \\ c \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix} \begin{pmatrix} r \\ g \\ b \end{pmatrix}$$
Χρησιμοποιείστε αντίστροφα μητρώα για ν εκφράσετε τα r, g και b μέσω των m, y και c.

9. Έστω ότι το τετραγωνικό μητρώο A ικανοποιεί τη σχέση $A^2 - 4A + I = O$, όπου O είναι το μηδενικό μητρώο. Δείξτε ότι το A αντιστρέφεται και ότι $A^{-1} = 4I - A$.

10. Έστω A ένα τετραγωνικό μητρώο ώστε $A^2 = O$. Ένα τέτοιο μητρώο λέγεται **ταυτοδύναμο**. Δείξτε ότι

    (α΄) $(I - A)^n = I - nA$, για κάθε φυσικό αριθμό n.

    (β΄) Το μητρώο $I - A$ είναι αντιστρέψιμο και $(I - A)^{-1} = I + A$.

    (γ΄) Το μητρώο $I - nA$ είναι αντιστρέψιμο και βρείτε το αντίστροφο.

11. Έστω A τετραγωνικό μητρώο ώστε $A^{k+1} = O$, για κάποιο θετικό ακέραιο k. Δείξτε ότι
$$(I - A)^{-1} = I + A + A^2 + \cdots + A^k.$$

12. Έστω J_n το $n \times n$ μητρώο του οποίου όλα τα στοιχεία είναι 1. Δείξτε ότι

    (α΄) $J_n^2 = nJ_n$

    (β΄) $(I - J_n)^{-1} = I - \dfrac{1}{n-1} J_n$.

13. Θεωρούμε το διαγώνιο μητρώο
$$A = \begin{pmatrix} a_1 & 0 & \cdots & 0 \\ 0 & a_2 & \cdots & 0 \\ \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & a_n \end{pmatrix}$$
με $a_i \neq 0$ για $i = 1, 2, \ldots, n$. Δείξτε ότι το A είναι αντιστρέψιμο και βρείτε το A^{-1}.

14. Έστω ότι
$$M = \begin{pmatrix} \alpha & \beta \\ \gamma & \delta \end{pmatrix}, \qquad \text{με} \quad \begin{cases} \alpha^2 + \gamma^2 = 1 \\ \beta^2 + \delta^2 = 1 \\ \alpha\beta + \gamma\delta = 0 \end{cases}$$

    (α΄) Δείξτε ότι για το M^{T} ισχύουν οι ανάλογες σχέσεις με αυτές που ισχύουν για το M, δηλαδή οι στήλες του είναι μοναδιαία διανύσματα κάθετα μεταξύ τους.

    (β΄) Δείξτε ότι $M^{-1} = M^{\mathrm{T}}$.

15. Εάν
$$C = \begin{pmatrix} 1 & c_1 & c_2 \\ 0 & 1 & c_3 \\ 0 & 0 & 1 \end{pmatrix},$$

ορίζουμε $N = C - I$. Δείξτε ότι $N^3 = O$. Δείξτε επίσης ότι το C αντιστρέφεται και ότι

$$\begin{aligned} C^{-1} &= I - N + N^2 \\ &= 3I - 3C + C^2. \end{aligned}$$

Υπόδειξη: $C = I + N$.

16. Εάν A και B είναι μητρώα κατάλληλων διαστάσεων ώστε $AB = O$ και το A είναι αντιστρέψιμο, δείξτε ότι $B = O$.

17. Έστω ότι το A είναι ένα τετραγωνικό μητρώο.

   (α΄) Αν $A^2 = O$, δείξτε ότι το A δεν αντιστρέφεται.

   (β΄) Αν $AB = O$ για κάποιο μη μηδενικό μητρώο B εξετάστε αν το A έχει αντίστροφο.

18. Εάν A και B είναι $n \times n$ μητρώα και τέτοια ώστε τα μητρώα που εμφανίζονται στις παρακάτω σχέσεις ορίζονται, αποδείξτε τις ταυτότητες.

   (α΄) $(A^{-1} + B^{-1})^{-1} = A(A + B)^{-1}B$.

   (β΄) $(I + AB)^{-1}A = A(I + BA)^{-1}$.

   (γ΄) $(A + BB^{\mathrm{T}})^{-1}B = A^{-1}B(I + B^{\mathrm{T}}A^{-1}B)^{-1}$.

2.6 Παράρτημα III: Ομάδες

Ορισμός 2.11. Έστω $\mathcal{G}$ ένα μη κενό σύνολο στο οποίο έχει ορισθεί μια πράξη $\mathcal{G} \times \mathcal{G} \ni (x, y) \to x \cdot y \in \mathcal{G}$. Θα λέμε ότι η δομή $(\mathcal{G}, \cdot)$ είναι **ομάδα** (group) εάν ικανοποιούνται οι νόμοι:

(G1) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$, για κάθε $x, y, z \in \mathcal{G}$.

(G2) Υπάρχει $e \in \mathcal{G}$, ώστε $e \cdot x = x \cdot e = x$ για κάθε $x \in \mathcal{G}$.

(G3) Για κάθε $x \in \mathcal{G}$, υπάρχει $x^{-1} \in \mathcal{G}$, ώστε $x^{-1} \cdot x = x \cdot x^{-1} = e$.

Ορισμός 2.12. Μια ομάδα $(\mathcal{G}, \cdot)$ λέγεται **αβελιανή** (abelian) ή **μεταθετική** (commutative) εάν η πράξη είναι μεταθετική ισχύει δηλαδή

(GA) $x \cdot y = y \cdot x$, για κάθε $x, y \in \mathcal{G}$.

Παράδειγμα 2.13. Αν $\mathbb{Q}$ είναι το σύνολο των ρητών αριθμών, $\mathbb{Q}^* = \mathbb{Q} \setminus \{0\}$, και $\mathbb{R}^* = \mathbb{R} \setminus \{0\}$, είναι προφανές ότι οι $(\mathbb{Q}, +)$, $(\mathbb{Q}^*, \cdot)$, $(\mathbb{R}, +)$, $(\mathbb{R}^*, \cdot)$, όπου “$+$” και “$\cdot$” είναι οι συνήθεις πράξεις πρόσθεση και πολλαπλασιασμός, είναι ομάδες. Όμοια οι $(\mathbb{C}, +)$ και $(\mathbb{C}^*, \cdot)$ είναι ομάδες. Επίσης η δομή $(\mathbb{R}^n, +)$, όπου η πρόσθεση είναι η πρόσθεση διανυσμάτων, είναι ομάδα. Επίσης το σύνολο των ακεραίων $\mathbb{Z}$ με την πράξη της πρόσθεσης, $(\mathbb{Z}, +)$) είναι ομάδα.

▶ **Άσκηση 2.14.** Δείξτε ότι το σύνολο

$$\left\{\frac{1+2m}{1-2n} : m, n \in \mathbb{Z}\right\}$$

με πράξη τον πολλαπλασιασμό, είναι ομάδα.

Παράδειγμα 2.14. Έστω A ένα μη κενό πεπερασμένο σύνολο. Κάθε ένα-προς-ένα απεικόνιση του A επί του A λέγεται **μετάθεση** (permutation) του A. Αν $A = \{a_1, a_2, \ldots, a_n\}$, ο τρόπος που συνήθως συμβολίζουμε μια μετάθεση, έστω s, του A είναι

$$s = \begin{pmatrix} a_1 & a_2 & \ldots & a_n \\ s(a_1) & s(a_2) & \ldots & s(a_n) \end{pmatrix} \tag{2.30}$$

Το σύνολο των μεταθέσεων του A συμβολίζουμε με S_A.

- Έστω $s_1, s_2 \in S_A$ και ας θεωρήσουμε τη σύνθεση $s_1 \circ s_2$, όπου $(s_1 \circ s_2)(a_i) = s_1(s_2(a_i))$. Επειδή οι s_1 και s_2 είναι ένα-προς-ένα έχουμε

$$s_1(s_2(a_i)) = s_1(s_2(a_j)) \Rightarrow s_2(a_i) = s_2(a_j) \Rightarrow a_i = a_j,$$

  δηλαδή η $s_1 \circ s_2$ είναι ένα-προς-ένα. Επίσης $s_1(s_2(A)) = s_1(A) = A$, δηλαδή η $s_1 \circ s_2$ είναι επί του A κατά συνέπεια $s_1 \circ s_2 \in S_A$.

- Από τις ιδιότητες της σύνθεσης συναρτήσεων έπεται ότι αν $s_1, s_2, s_3 \in S_A$, τότε

$$(s_1 \circ s_2) \circ s_3 = s_1 \circ (s_2 \circ s_3).$$

- Αν με s_0 συμβολίσουμε την ταυτοτική απεικόνιση, δηλαδή $s_0(a_i) = a_i$, για $i = 1, 2, \ldots, n$, τότε $s_0 \in S_A$ και για κάθε $s \in S_A$

$$(s_0 \circ s)(a_i) = s_0(s(a_i)) = s(a_i) = s(s_0(a_i)) = (s \circ s_0)(a_i)$$

  για κάθε $i = 1, 2, \ldots, n$.

- Αν $s \in S_A$ και s^{-1} είναι η αντίστροφη απεικόνιση της s, τότε $s^{-1} \in S_A$ (γιατί;) και $s \circ s^{-1} = s^{-1} \circ s = s_0$. Αν η s είναι όπως στην (2.30), τότε

$$s^{-1} = \begin{pmatrix} s(a_1) & s(a_2) & \ldots & s(a_n) \\ a_1 & a_2 & \ldots & a_n \end{pmatrix}. \tag{2.31}$$

Αν $s_1, s_2 \in S_A$ ορίζουμε το γινόμενο, γράφουμε απλά, $s_1 s_2$ να είναι η σύνθεση $s_1 \circ s_2$, έτσι

$$\begin{aligned} s_1 s_2 &= \begin{pmatrix} a_1 & a_2 & \ldots & a_n \\ s_1(a_1) & s_1(a_2) & \ldots & s_1(a_n) \end{pmatrix} \begin{pmatrix} a_1 & a_2 & \ldots & a_n \\ s_2(a_1) & s_2(a_2) & \ldots & s_2(a_n) \end{pmatrix} \\ &= \begin{pmatrix} a_1 & a_2 & \ldots & a_n \\ s_1(s_2(a_1)) & s_1(s_2(a_2)) & \ldots & s_1(s_2(a_n)) \end{pmatrix}, \end{aligned}$$

τότε όπως δείξαμε η $(S_A, \cdot)$ είναι ομάδα.
Το πλήθος των μεταθέσεων του A είναι $n!$, όσο και το πλήθος των λέξεων του αλφαβήτου $\{a_1, a_2, \ldots, a_n\}$ με n γράμματα χωρίς επαναλήψεις. Αν $A = \{1, 2, \ldots, n\}$, τότε το σύνολο

των μεταθέσεων του A συμβολίζουμε απλά με S_n. Για παράδειγμα $S_3 = \{s_0, s_1, \cdots, s_5\}$, όπου

$$s_0 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \end{pmatrix}, \qquad s_1 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}, \qquad s_2 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix},$$
$$s_3 = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}, \qquad s_4 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}, \qquad s_5 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix}.$$

Για να υπολογίσουμε το γινόμενο s_1s_4, για παράδειγμα, σκεφτόμαστε ότι το 1 μέσω της s_4 απεικονίζεται στο 3, το οποίο, με τη σειρά του, μέσω της s_1 απεικονίζεται στο 2, έτσι $s_1s_4(1) = 2$, όμοια, σχηματικά,

$$2 \xrightarrow{s_4} 1 \xrightarrow{s_1} 1 \Rightarrow 2 \xrightarrow{s_1s_4} 1.$$

Με αυτόν τον τρόπο βρίσκουμε

$$s_1s_4 = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 3 & 1 & 2 \\ 2 & 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 1 & 3 \end{pmatrix} = s_2.$$

Επίσης, από τον ορισμό του αντίστροφου,

$$s_3^{-1} = \begin{pmatrix} 2 & 3 & 1 \\ 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix} = s_4.$$

Οι ομάδες S_n λέγονται **συμμετρικές ομάδες** και παίζουν σπουδαίο ρόλο στη θεωρία ομάδων. Προσοχή

$$s_4s_1 = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 2 & 1 \end{pmatrix} = s_5,$$

δηλαδή $s_1s_4 \neq s_4s_1$, κατά συνέπεια οι ομάδα S_3 **δεν** είναι αβελιανή.

Παρατήρηση 2.13. Αν $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ είναι η κανονική βάση του $\mathbb{R}^3$ και $S_3 = \{s_0, s_1, \ldots, s_5\}$ είναι η συμμετρική ομάδα, βλέπε Παράδειγμα 2.14, ένας τρόπος να ορισθούν τα μητρώα μετάθεσης είναι ο ακόλουθος:

$$M_k = \begin{pmatrix} \boldsymbol{e}_{s_k(1)} & \boldsymbol{e}_{s_k(2)} & \boldsymbol{e}_{s_k(3)} \end{pmatrix}, \qquad k = 0, 1, \ldots, 5. \tag{2.32}$$

Στη δική μας αρίθμηση των M_k υπάρχει αυτή η αντιστοιχία γιατί ακολουθήθηκε ο ίδιος τρόπος παράθεσης των μεταθέσεων. Στη γενική περίπτωση ισχύει ότι αν M_j είναι ένα μητρώο μετάθεσης, τότε υπάρχει μετάθεση $s_i \in S_3$ ώστε

$$M_j = \begin{pmatrix} \boldsymbol{e}_{s_i(1)} & \boldsymbol{e}_{s_i(2)} & \boldsymbol{e}_{s_i(3)} \end{pmatrix}. \tag{2.33}$$

Παράδειγμα 2.15. Λέγοντας **γραμμική κίνηση** (linear motion) στο $\mathbb{R}^2$ εννοούμε μια απεικόνιση $T : \mathbb{R}^2 \to \mathbb{R}^2$ η οποία "μετακινεί" ή μετασχηματίζει το σημείο (x, y) στο σημείο (x', y') ώστε

$$\begin{aligned} x' &= ax + by, \\ y' &= cx + dy, \end{aligned} \qquad \text{με} \qquad ad - bc \neq 0. \tag{2.34}$$

Την ποσότητα $ad - bc$ λέμε **ορίζουσα** (determinant) του μετασχηματισμού T και τη δηλώνουμε ως

$$J_T := \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc. \tag{2.35}$$

Η σχέση $J_T \neq 0$ εξασφαλίζει ότι η αντιστοιχία $(x, y) \leftrightarrow (x', y')$ είναι ένα-προς-ένα, κατά συνέπεια η αντίστροφη απεικόνιση T^{-1} υπάρχει. Πράγματι, λύνοντας το σύστημα βρίσκουμε ότι

$$\begin{aligned} x &= \frac{d}{ad - bc}x' + \frac{-b}{ad - bc}y', \\ y &= \frac{-c}{ad - bc}x' + \frac{a}{ad - bc}y'. \end{aligned} \tag{2.36}$$

Επιπλέον υπολογίζοντας την ορίζουσα του αντίστροφου μετασχηματισμού βρίσκουμε

$$J_{T^{-1}} = \frac{da - bc}{(ad - bc)^2} = \frac{1}{ad - bc} \neq 0,$$

κατά συνέπεια η αντίστροφη απεικόνιση T^{-1} είναι επίσης γραμμική κίνηση. Εάν $T(x, y) = (x', y')$ και $T'(x', y') = (x'', y'')$, όπου η T' δίνεται όπως στην (2.34) με a', b', c', d' στη θέση των a, b, c, d αντίστοιχα, και με $T'T$ συμβολίσουμε τη σύνθεση $T' \circ T$, τότε

$$\begin{aligned} x'' &= a'x' + b'y' = a'(ax + by) + b'(cx + dy) \\ &= (a'a + b'c)x + (a'b + b'd)y \\ y'' &= c'x' + d'y' = c'(ax + by) + d'(cx + dy) \\ &= (c'a + d'c)x + (c'b + d'd)y \end{aligned}$$

και για την ορίζουσα της σύνθεσης υπολογίζουμε

$$(a'a + b'c)(c'b + d'd) - (a'b + b'd)(c'a + d'c) = (ad - bc)(a'd' - b'c') \neq 0,$$

ισοδύναμα

$$J_{T \circ T'} = J_T J_{T'} \neq 0,$$

κατά συνέπεια η $T'T$ είναι και αυτή μια γραμμική κίνηση. Επιπλέον η ταυτοτική απεικόνιση $I(x, y) = (x, y)$ είναι επίσης γραμμική κίνηση αφού η (2.34) ικανοποιείται

$$\begin{aligned} x &= 1x + 0y \\ y &= 0x + 1y. \end{aligned}$$

Δείχνεται εύκολα (και αφήνεται σαν άσκηση) ότι αν T, T', T'' είναι γραμμικές κινήσεις, τότε $(TT')T'' = T(T'T'')$. Επομένως οι κινήσεις αυτές με την πράξη του πολλαπλασιασμού-σύνθεσης αποτελούν ομάδα. Η ομάδα αυτή λέγεται διδιάστατη **γενική γραμμική ομάδα** (general linear group) και συμβολίζεται με $GL(2, \mathbb{R})$.
Είδαμε ότι ο μετασχηματισμός T είναι ένα-προς-ένα, ισοδύναμα αντιστρέψιμος, αν και μόνο αν $J_T \neq 0$, δηλαδή η ορίζουσά του είναι διάφορη του μηδενός.
Κάθε τέτοια απεικόνιση μπορεί να εκφραστεί ως το γινόμενο ενός 2×2 μητρώου με διάνυσμα στο $\mathbb{R}^2$, συγκεκριμένα

$$T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}. \tag{2.37}$$

Είδαμε ότι η σχέση $ad-bc \neq 0$ εξασφαλίζει ότι υπάρχει ο αντίστροφος μετασχηματισμός και δίνεται από τη σχέση

$$T^{-1}\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} \tag{2.38}$$

βλέπε (2.36), έτσι ώστε αν $\boldsymbol{y} = T(\boldsymbol{x}) = A\boldsymbol{x}$, τότε $\boldsymbol{x} = T^{-1}(\boldsymbol{y}) = A^{-1}\boldsymbol{y}$ αφού

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix}^{-1} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

(Άσκηση 2.8). Εάν $T(x,y) = (x',y')$ και $T'(x',y') = (x'',y'')$, όπου η T' δίνεται όπως στην (2.37) με a',b',c',d' στη θέση των a,b,c,d αντίστοιχα, και με $T'T$ συμβολίσουμε τη σύνθεση $T' \circ T$ δείξαμε ότι

$$T'T\begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} a'a+b'c & a'b+b'd \\ c'a+d'c & c'b+d'd \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}$$

το οποίο όμως είναι

$$= \begin{pmatrix} a' & b' \\ c' & d' \end{pmatrix}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix}.$$

Μπορούμε λοιπόν να ορίσουμε, ισοδύναμα, τη διδιάστατη γενική γραμμική ομάδα $GL(2,\mathbb{R})$ ως το σύνολο των 2×2 μητρώων με μη μηδενική ορίζουσα (ισοδύναμα των αντιστρέψιμων μητρώων), με πράξη τον πολλαπλασιασμό.

Σημείωση 2.3. Με ανάλογο τρόπο μπορούμε, γενικεύοντας, να ορίσουμε την ομάδα $GL(n,\mathbb{R})$. Αυτούς τους μετασχηματισμούς μελετώντας ο Arthur Cayley κατανόησε ότι τα μητρώα που ορίζουν τους μετασχηματισμούς είναι αλγεβρικές οντότητες αφού ορίζοντας ως γινόμενο δύο τέτοιων το μητρώο που αντιστοιχεί στη σύνθεση των μετασχηματισμών προκύπτει ένα λογικό αποτέλεσμα, η δε πράξη του πολλαπλασιασμού έχει φυσιολογικές ιδιότητες. Στη περίπτωση μάλιστα όπου τα μητρώα είναι 1×1 το γινόμενο αντιστοιχεί στο σύνηθες γινόμενο αριθμών. Έτσι ενώ ο Sylvester είναι ο γεννήτορας της έννοιας του μητρώου ο Cayley είναι ο δημιουργός της Άλγεβρας των μητρώων.

2.7 Παράρτημα IV: Σώματα

Θυμίζουμε τον ορισμό του σώματος

Ορισμός 2.13. Έστω $\mathcal{F}$ ένα μη κενό σύνολο στο οποίο έχουν ορισθεί δύο πράξεις, πρόσθεση $\mathcal{F}\times\mathcal{F} \ni (x,y) \to x+y \in \mathcal{F}$ και πολλαπλασιασμός $\mathcal{F}\times\mathcal{F} \ni (x,y) \to x\cdot y \in \mathcal{F}$. Η δομή $(\mathcal{F},+,\cdot)$ λέγεται **σώμα** (field) εάν ικανοποιούνται οι νόμοι:

(F1) $x+y = y+x$, για κάθε $x,y \in \mathcal{F}$.

(F2) $(x+y)+z = x+(y+z)$, για κάθε $x,y,z \in \mathcal{F}$.

(F3) Υπάρχει $0 \in \mathcal{F}$, ώστε $x+0 = x$, για κάθε $x \in \mathcal{F}$.

(F4) Για κάθε $x \in \mathcal{F}$ υπάρχει $-x \in \mathcal{F}$, ώστε $x + (-x) = 0$.

(F5) $x \cdot y = y \cdot x$, για κάθε $x, y \in \mathcal{F}$.

(F6) $(x \cdot y) \cdot z = x \cdot (y \cdot z)$, για κάθε $x, y, z \in \mathcal{F}$.

(F7) Υπάρχει $1 \in \mathcal{F}$, $1 \neq 0$ ώστε $x \cdot 1 = x$, για κάθε $x \in \mathcal{F}$.

(F8) Για κάθε $x \neq 0$ υπάρχει $x^{-1} \in \mathcal{F}$, ώστε $x \cdot x^{-1} = 1$.

(F9) $x \cdot (y + z) = x \cdot y + x \cdot z$, για κάθε $x, y, z \in \mathcal{F}$.

Παράδειγμα 2.16. Είναι προφανές ότι οι $(\mathbb{Q}, +, \cdot)$, $(\mathbb{R}, +, \cdot)$, $(\mathbb{C}, +, \cdot)$ όπου "+" και "·" είναι οι συνήθεις πράξεις πρόσθεση και πολλαπλασιασμός, είναι σώματα. Αν $\mathbb{Z}$ το σύνολο των ακεραίων η δομή $(\mathbb{Z}, +, \cdot)$) δεν είναι σώμα (γιατί;).

► **Άσκηση 2.15.** Με $\mathbb{Q}[\sqrt{2}]$ συμβολίζουμε το σύνολο των αριθμών της μορφής $x + y\sqrt{2}$, όπου x και y είναι ρητοί αριθμοί. Δείξτε ότι το $\mathbb{Q}[\sqrt{2}]$ με τη πρόσθεση και τον πολλαπλασιασμό είναι σώμα.

Παράδειγμα 2.17 (Οι αλγεβρικοί αριθμοί). Εάν r είναι ρίζα μιας πολυωνυμικής εξίσωσης

$$a_0 + a_1 x + a_2 x^2 + \cdots + a_n x^n = 0$$

με $a_j \in \mathbb{Q}$, $j = 0, 1, 2, \ldots, n$ ο r λέγεται **αλγεβρικός αριθμός**. Για παράδειγμα οι $\sqrt{2}$, $\sqrt[3]{7}$ είναι αλγεβρικοί αριθμοί. Οι e και π δεν είναι αλγεβρικοί αριθμοί. Μπορεί να δειχθεί ότι οι αλγεβρικοί αριθμοί με την πρόσθεση και τον πολλαπλασιασμό αποτελούν σώμα.

Κεφάλαιο 3

Επίλυση γραμμικών συστημάτων

3.1 Η έννοια της απαλοιφής

Ορισμός 3.1. Μια εξίσωση

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b,$$

με $a_1, a_2, \ldots, a_n, b$ πραγματικές ή μιγαδικές σταθερές και αγνώστους $x_1, x_2, \ldots, x_n$ λέγεται **γραμμική εξίσωση** (linear equation). Ένα σύστημα γραμμικών εξισώσεων λέγεται **γραμμικό σύστημα** (linear system). Ένα γραμμικό σύστημα

$$A\boldsymbol{x} = \boldsymbol{b},$$

όπου $A \in \mathbb{M}^{n,m}(\mathbb{K})$, $\boldsymbol{x} \in \mathbb{K}^m$ και $\boldsymbol{b} \in \mathbb{K}^n$, λέγεται **ομοιογενές** ή **ομογενές** (homogeneous) αν $\boldsymbol{b} = \mathbf{0}$, και **μη ομοιογενές** ή **μη ομογενές** (non homogeneous) αν $\boldsymbol{b} \neq \mathbf{0}$.

Ο βασικός τρόπος επίλυσης ενός γραμμικού συστήματος έγκειται στον μετασχηματισμό του δοσμένου συστήματος σε ένα νέο σύστημα το οποίο έχει την ίδια λύση με το αρχικό, ένα **ισοδύναμο** όπως λέμε σύστημα το οποίο είναι ευκολώτερο να λυθεί. Για παράδειγμα για το "τριγωνικό" σύστημα

$$\begin{aligned} x + 3y - 2z &= 3 \\ y + z &= 3 \\ 3z &= 6 \end{aligned} \tag{3.1}$$

η λύση προκύπτει σχεδόν αμέσως αφού με την "προς τα πίσω αντικατάσταση" διαδοχικά υπολογίζουμε

(α΄) Από την 3η εξίσωση: $z = 2$

(β΄) Από την 2η εξίσωση: $y = 3 - z = 3 - 2 = 1$ και τέλος

(γ΄) Από την 1η εξίσωση: $x = 3 - 3y + 2z = 3 - 3 + 4 = 4$.

Κατά συνέπεια αν δοθεί ένα σύστημα προσπαθούμε διαμορφώνοντας μια ακολουθία από ισοδύναμα συστήματα να καταλήξουμε σε ένα τριγωνικό όπως στην (3.1), ή "περίπου" τριγωνικό, όπως θα δούμε παρακάτω, σύστημα. Πώς όμως από ένα σύστημα προκύπτει ένα ισοδύναμο; Η $x_0 = 4$, $y_0 = 1$, $z_0 = 2$ είναι η λύση του (3.1) και αν με E_1, E_2, E_3 παραστήσουμε τις εξισώσεις του (3.1), τότε το σύστημα, για παράδειγμα, που προκύπτει αντικαθιστώντας την πρώτη εξίσωση E_1 με την $3E_1$, την δεύτερη εξίσωση E_2 με το άθροισμα $2E_1 + E_2$, και την τρίτη E_3 με το άθροισμα $3E_1 + E_3$ προκύπτει ένα σύστημα του οποίου η λύση είναι η (x_0, y_0, z_0) αφού

$$\begin{aligned} 3x_0 + 9y_0 - 6z_0 &= 9 \\ 2x_0 + 7y_0 - 3z_0 &= 9 \\ 3x_0 + 9y_0 - 3z_0 &= 15 \end{aligned} \tag{3.2}$$

Δύο συστήματα είναι ισοδύναμα αν το ένα προκύπτει από το άλλο εκτελώντας τις πράξεις:

1. Πολλαπλασιασμός μιας εξίσωσης με μια μη μηδενική σταθερά.
2. Μετάθεση δύο εξισώσεων.
3. Πρόσθεση ενός πολλαπλασίου μιας εξίσωσης σε άλλη εξίσωση.

Με την τρίτη πράξη επιτυγχάνεται η απαλοιφή κάποιου επιλεγμένου αγνώστου. Ας δούμε λοιπόν σε ένα παράδειγμα όπου εκτελώντας αυτές τις πράξεις σε διαδοχικά βήματα καταλήγουμε σε ένα ισοδύναμο τριγωνικό σύστημα. Εργαζόμενοι σε διπλό ταμπλώ αποτυπώνουμε όλα τα βήματα αυτού του μετασχηματισμού και στα αντίστοιχα μητρώα.

Παράδειγμα 3.1. Να επιλυθεί το σύστημα

$$\begin{aligned} x + 3y - 2z + w &= -1 \\ 2x + y + z + 4w &= -3 \\ -x + 2y + z + 6w &= -4 \end{aligned}$$

Σε κάθε βήμα γράφουμε E_1, E_2, E_3 για την πρώτη δεύτερη και τρίτη εξίσωση του συστήματος, αντίστοιχα. Γράφουμε το αρχικό σύστημα καθώς και το αντίστοιχο μητρώο

$$\begin{aligned} x + 3y - 2z + w &= -1 \\ 2x + y + z + 4w &= -3 \\ -x + 2y + z + 6w &= -4 \end{aligned} \qquad \begin{pmatrix} 1 & 3 & -2 & 1 & -1 \\ 2 & 1 & 1 & 4 & -3 \\ -1 & 2 & 1 & 6 & -4 \end{pmatrix}$$

αντικαθιστούμε την δεύτερη εξίσωση E_2 με την $-2E_1 + E_2$, η οποία γίνεται η νέα E_2

$$\begin{aligned} x + 3y - 2z + w &= -1 \\ -5y + 5z + 2w &= -1 \\ -x + 2y + z + 6w &= -4 \end{aligned} \qquad \begin{pmatrix} 1 & 3 & -2 & 1 & -1 \\ 0 & -5 & 5 & 2 & -1 \\ -1 & 2 & 1 & 6 & -4 \end{pmatrix}$$

αντικαθιστούμε την E_3 με την $E_1 + E_3$, η οποία γίνεται η νέα E_3

$$\begin{aligned} x + 3y - 2z + w &= -1 \\ -5y + 5z + 2w &= -1 \\ 5y - z + 7w &= -5 \end{aligned} \qquad \begin{pmatrix} 1 & 3 & -2 & 1 & -1 \\ 0 & -5 & 5 & 2 & -1 \\ 0 & 5 & -1 & 7 & -5 \end{pmatrix}$$

αντικαθιστούμε E_3 με την $E_2 + E_3$ και παίρνουμε

$$\begin{aligned} x + 3y - 2z + w &= -1 \\ -5y + 5z + 2w &= -1 \\ 4z + 9w &= -6 \end{aligned} \qquad \begin{pmatrix} 1 & 3 & -2 & 1 & -1 \\ 0 & -5 & 5 & 2 & -1 \\ 0 & 0 & 4 & 9 & -6 \end{pmatrix}$$

Παρατηρούμε ότι η διαδικασία "τριγωνοποίησης" του συστήματος τερματίζεται (γιατί;), επομένως με προς τα πίσω αντικατάσταση, από το τελευταίο σύστημα βρίσκουμε

$$\left.\begin{aligned} z &= -\frac{3}{2} - \frac{9}{4}w \\ y &= \frac{1}{5} + z + \frac{2}{5}w \\ x &= -1 - 3y + 2z - w \end{aligned}\right\} \Leftrightarrow \left\{\begin{aligned} w &= t \\ z &= -\frac{3}{2} - \frac{9}{4}t \\ y &= -\frac{13}{10} - \frac{37}{20}t \\ x &= -\frac{1}{10} + \frac{1}{20}t \end{aligned}\right.$$

όπου $t \in \mathbb{R}$, κατά συνέπεια το σύστημα έχει άπειρες λύσεις.

Παρατηρούμε ότι θα μπορούσαμε να πάρουμε τη λύση του συστήματος εκτελώντας τις πράξεις, που περιγράψαμε στη διαδικασία της επίλυσης, όχι στις εξισώσεις του συστήματος αλλά απευθείας στις γραμμές του αντίστοιχου μητρώου. Για παράδειγμα θα μπορούσαμε από το αρχικό να πάμε στο δεύτερο μητρώο στο οποίο η δεύτερη γραμμή αντικαθίσταται από το άθροισμα (διανυσματικό) του πολλαπλασίου της πρώτης γραμμής επί -2 και της δεύτερης γραμμής, σχηματικά

$$r_2 \to -2r_1 + r_2,$$

όπου με r_i δηλώνουμε την i γραμμή (row) του μητρώου. Όμοια για οικονομία στα βήματα μπορούμε να πάμε από το αρχικό στο τρίτο, στη διαδικασία επίλυσης, μητρώο με τις πράξεις

$$r_2 \to -2r_1 + r_2, \qquad \text{και} \qquad r_3 \to r_1 + r_3$$

στο πρώτο βήμα, και στη συνέχεια με την πράξη

$$r_3 \to r_2 + r_3$$

να καταλήξουμε στο τελευταίο μητρώο. Τη διαδικασία αυτή τη δηλώνουμε με

$$\begin{pmatrix} 1 & 3 & -2 & 1 & -1 \\ 2 & 1 & 1 & 4 & -3 \\ -1 & 2 & 1 & 6 & -4 \end{pmatrix} \xrightarrow[r_3 \to r_2 + r_3]{r_2 \to -2r_1 + r_2} \begin{pmatrix} 1 & 3 & -2 & 1 & -1 \\ 0 & -5 & 5 & 2 & -1 \\ 0 & 5 & -1 & 7 & -5 \end{pmatrix} \xrightarrow{r_3 \to r_2 + r_3} \begin{pmatrix} 1 & 3 & -2 & 1 & -1 \\ 0 & -5 & 5 & 2 & -1 \\ 0 & 0 & 4 & 9 & -6 \end{pmatrix}.$$

Σημαντικό ρόλο στη διαδικασία της απαλοιφής παίζουν τα πρώτα, από αριστερά, μη μηδενικά στοιχεία κάθε γραμμής τα οποία, τελικά, μηδενίζουν τα στοιχεία του μητρώου κάτω από αυτά. Στο Παράδειγμα 3.1 τα στοιχεία αυτά είναι τα κυκλωμένα

$$\begin{pmatrix} 1 & 3 & -2 & 1 & -1 \\ 2 & 1 & 1 & 4 & -3 \\ -1 & 2 & 1 & 6 & -4 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -2 & 1 & -1 \\ 0 & -5 & 5 & 2 & -1 \\ 0 & 5 & -1 & 7 & -5 \end{pmatrix} \to \begin{pmatrix} 1 & 3 & -2 & 1 & -1 \\ 0 & -5 & 5 & 2 & -1 \\ 0 & 0 & 4 & 9 & -6 \end{pmatrix}.$$

Ορισμός 3.2. Το πρώτο μη μηδενικό στοιχείο της γραμμής του μητρώου που εκτελεί την απαλοιφή λέγεται **οδηγός** (pivot). Η θέση που βρίσκεται ένας οδηγός λέγεται **θέση οδηγού** (pivot possition).

Οι οδηγοί του μητρώου του Παραδείγματος 3.1 είναι οι 1, -5 και 4. Οι οδηγοί ενός μητρώου δεν είναι όλοι εμφανείς από την αρχή αλλά εμφανίζονται κατά τη διαδικασία της απαλοιφής.

Θα δούμε αργότερα ότι οι οδηγοί αποκαλύπτουν σημαντική πληροφορία σχετικά με τη δομή του μητρώου.

Σε αντιστοιχία με τις πράξεις που διατηρούν την ισοδυναμία συστημάτων ορίζουμε τις **στοιχειώδεις πράξεις γραμμών** σε μητρώα να είναι

1. Πολλαπλασιασμός μιας γραμμής με μια μη μηδενική σταθερά.
2. Μετάθεση δύο γραμμών.
3. Πρόσθεση ενός πλλαπλασίου μιας γραμμής σε άλλη γραμμή.

Παρατηρώντας το τελευταίο μητρώο στην ακολουθία απλοποίησης, δηλαδή στη **διαδικασία της απαλοιφής**, βλέπουμε ότι εφαρμόζοντας τις στοιχειώδεις πράξεις γραμμών μπορούμε να πάρουμε ακόμα "απλούστερο" μητρώο που λογικά αντιστοιχεί σε απλούστερο τελικό σύστημα, για παράδειγμα πολαπλασιάζοντας την δεύτερη γραμμή με $-1/5$ και την τρίτη με $1/4$ παίρνουμε

$$\begin{pmatrix} 1 & 3 & -2 & 1 & -1 \\ 0 & 1 & -1 & -2/5 & 1/5 \\ 0 & 0 & 1 & 9/4 & -3/2 \end{pmatrix}$$

και συνεχίζοντας (με την πράξη $r_1 \to r_1 - 3r_2$) βρίσκουμε διαδοχικά

$$\begin{pmatrix} 1 & 0 & 1 & 11/5 & -8/5 \\ 0 & 1 & -1 & -2/5 & 1/5 \\ 0 & 0 & 1 & 9/4 & -3/2 \end{pmatrix} \xrightarrow[r_2 \to r_2 + r_3]{r_1 \to r_1 - r_3} \begin{pmatrix} 1 & 0 & 0 & -1/20 & -1/10 \\ 0 & 1 & 0 & 37/20 & -13/10 \\ 0 & 0 & 1 & 9/4 & -3/2 \end{pmatrix}$$

απ΄ όπου διαβάζουμε τη λύση του συστήματος

$$\left.\begin{aligned} x - \frac{1}{20}w &= -\frac{1}{10} \\ y + \frac{37}{20}w &= -\frac{13}{10} \\ z + \frac{9}{4}w &= -\frac{3}{2} \end{aligned}\right\} \Leftrightarrow \begin{cases} x = -\dfrac{1}{10} + \dfrac{1}{20}t \\ y = -\dfrac{13}{10} - \dfrac{37}{20}t \\ z = -\dfrac{3}{2} - \dfrac{9}{4}t \\ w = t \end{cases} \qquad t \in \mathbb{R}.$$

Παράδειγμα 3.2. Θεωρούμε το σύστημα

$$\begin{aligned} x + y + 2z &= a \\ x - z &= b \\ 2x + y + z &= c \end{aligned}$$

όπου a, b, c είναι πραγματικές παράμετροι. Δείξτε ότι το σύστημα έχει λύση αν και μόνο αν $c = a + b$.
Θεωρούμε το αντίστοιχο επαυξημένο μητρώο

$$\begin{pmatrix} 1 & 1 & 2 & a \\ 1 & 0 & -1 & b \\ 2 & 1 & 1 & c \end{pmatrix}$$

προσθέτοντας -1 φορές την πρώτη γραμμή στη δεύτερη παίρνουμε

$$\begin{pmatrix} 1 & 1 & 2 & a \\ 0 & -1 & -3 & b-a \\ 2 & 1 & 1 & c \end{pmatrix}$$

προσθέτοντας -2 φορές την πρώτη γραμμή στη τρίτη παίρνουμε

$$\begin{pmatrix} 1 & 1 & 2 & a \\ 0 & -1 & -3 & b-a \\ 0 & -1 & -3 & c-2a \end{pmatrix}$$

προσθέτοντας -1 φορές την δεύτερη γραμμή στη τρίτη παίρνουμε

$$\begin{pmatrix} 1 & 1 & 2 & a \\ 0 & -1 & -3 & b-a \\ 0 & 0 & 0 & c-a-b \end{pmatrix}.$$

Η τρίτη εξίσωση του συστήματος που αντιστοιχεί στο τελευταίο μητρώο είναι

$$0x + 0y + 0z = c - a - b$$

κατά συνέπεια το σύστημα έχει λύση, είναι **συμβατό** όπως λέμε, αν και μόνο αν $c-a-b = 0$, ισοδύναμα $c = a + b$.
Το πλήρες σύστημα που αντιστοιχεί στο τελευταίο μητρώο είναι

$$\begin{aligned} x + y + 2z &= a \\ -y - 3z &= b - a \\ 0 &= c - a - b \end{aligned}$$

και δεχόμενοι τη συνθήκη συμβατότητας $c = a + b$ με την προς τα πίσω αντικατάσταση βρίσκουμε

$$\left.\begin{aligned} x &= a - y - 2z \\ y &= a - b - 3z \end{aligned}\right\} \Leftrightarrow \left\{\begin{aligned} &x = a - (a - b - 3t) - 2t \\ &y = a - b - 3t \\ &z = t \end{aligned}\right\} \Leftrightarrow \left\{\begin{aligned} &x = b + t \\ &y = a - b - 3t \\ &z = t \end{aligned}\right. \qquad t \in \mathbb{R},$$

συνεπώς για $c = a + b$ το σύστημα έχει άπειρες λύσεις.

Παρατήρηση 3.1 (Θεμελιώδης παρατήρηση για ύπαρξη λύσης). Ας θεωρήσουμε το σύστημα $A\boldsymbol{x} = \boldsymbol{b}$, όπου A είναι ένα $n \times m$ μητρώο. Αν $\boldsymbol{a}_j$, $j = 1, 2, \ldots, m$ είναι οι στήλες του A, τότε από τον ορισμό της πράξης του πολλαπλασιασμού μητρώων έπεται ότι το σύστημα είναι ισοδύναμο, βλέπε (2.4) ή (2.9), με την εξίσωση

$$\boldsymbol{a}_1 x_1 + \boldsymbol{a}_2 x_2 + \cdots + \boldsymbol{a}_m x_m = \boldsymbol{b},$$

όπου $x_1, x_2, \ldots, x_m$ είναι οι συνιστώσες του $\boldsymbol{x}$, κατά συνέπεια: *Το σύστημα $A\boldsymbol{x} = \boldsymbol{b}$ έχει λύση αν και μόνον αν το $\boldsymbol{b}$ είναι γραμμικός συνδυασμός των στηλών του A.*

3.2 Η γεωμετρική προσέγγιση

Γνωρίζουμε ότι η γραμμική εξίσωση $ax+by=c$ παριστάνει μια ευθεία στο $\mathbb{R}^2$, ενώ μια $ax+by+cz=d$ παριστάνει ένα επίπεδο στο $\mathbb{R}^3$. Θα λέμε ότι η γραμμική εξίσωση $a_1x_1+a_2x_2+\cdots+a_nx_n=b$ παριστάνει ένα **υπερεπίπεδο** στο $\mathbb{R}^n$.

Ας θεωρήσουμε ένα απλό σύστημα δύο εξισώσεων με δύο αγνώστους,

$$\begin{aligned} a_1x + b_1y &= c_1 \\ a_2x + b_2y &= c_2 \end{aligned}$$

Στη γενική περίπτωση, κάθε εξίσωση παριστάνει μια ευθεία στο επίπεδο, κατά συνέπεια οι δύο ευθείες είτε είναι παράλληλες, είτε τέμνονται σε ένα σημείο, είτε τέμνονται σε περισσότερα του ενός σημεία, δηλαδή ταυτίζονται. Έτσι λοιπόν το σύστημα είτε δεν έχει λύση, είτε έχει μοναδική λύση, είτε έχει άπειρες λύσεις. Γράφοντας το σύστημα στην ισοδύναμη διανυσματική μορφή

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} x + \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} y = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} \Leftrightarrow \boldsymbol{a}x + \boldsymbol{b}y = \boldsymbol{c},$$

λογικά θα πρέπει να εξάγουμε τα ίδια συμπεράσματα. Πράγματι υπάρχουν οι εξής περιπτώσεις να θεωρήσουμε.

(α΄) Τα διανύσματα $\boldsymbol{a}$, $\boldsymbol{b}$ δεν περιέχονται στην ίδια ευθεία. Στην περίπτωση αυτή το διάνυσμα $\boldsymbol{c}$ εκφράζεται ως διαγώνιος παραλληλογράμμου με με πλευρές κατά μήκος των διανυσμάτων $\boldsymbol{a}$ και $\boldsymbol{b}$, ισοδύναμα ως γραμμικός συνδυασμός των $\boldsymbol{a}$ και $\boldsymbol{b}$, . Έτσι το αρχικό σύστημα έχει μοναδική λύση.

(β΄) Τα διανύσματα $\boldsymbol{a}$, $\boldsymbol{b}$ περιέχονται στην ίδια ευθεία. Τότε θα είναι $\boldsymbol{b} = \lambda\boldsymbol{a}$ για κάποιο $\lambda \in \mathbb{R}$ και η διανυσματική εξίσωση γίνεται

$$\begin{pmatrix} a_1 \\ a_2 \end{pmatrix} (x + \lambda y) = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix}$$

οπότε διακρίνουμε τις δύο υποπεριπτώσεις:

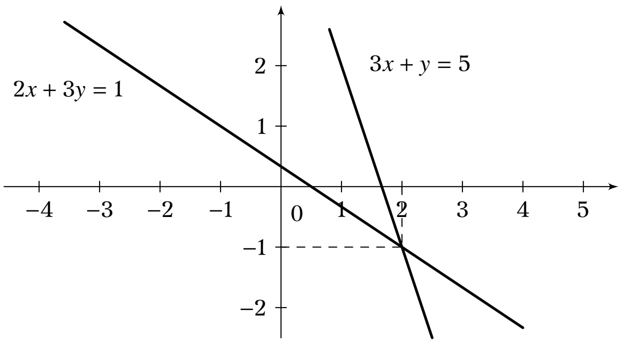
(β1΄) Το διάνυσμα $\boldsymbol{c}$ δεν περιέχεται στην ευθεία που ορίζει το $\boldsymbol{a}$, ισοδύναμα $\boldsymbol{c} \neq t\boldsymbol{a}$ για κάθε $t \in \mathbb{R}$, κατά συνέπεια το σύστημα δεν έχει λύση.

(β2΄) Το διάνυσμα $\boldsymbol{c}$ περιέχεται στην ευθεία που ορίζει το $\boldsymbol{a}$, ισοδύναμα $\boldsymbol{c} = \mu\boldsymbol{a}$ για κάποιο $\mu \in \mathbb{R}$, ισοδύναμα $x + \lambda y = \mu$. Τότε για κάθε $t \in \mathbb{R}$ η $x = \mu - \lambda t$, $y = t$ είναι λύση του συστήματος, οπότε το σύστημα έχει άπειρες λύσεις.

Ας επιλύσουμε γραφικά ένα απλό σύστημα, ακολουθώντας τις δύο θεωρήσεις, δηλαδή τη θεώρηση κατά γραμμές και τη θεώρηση κατά στήλες, δηλαδή αυτή μέσω της διανυσματικής εξίσωσης. Έστω λοιπόν το

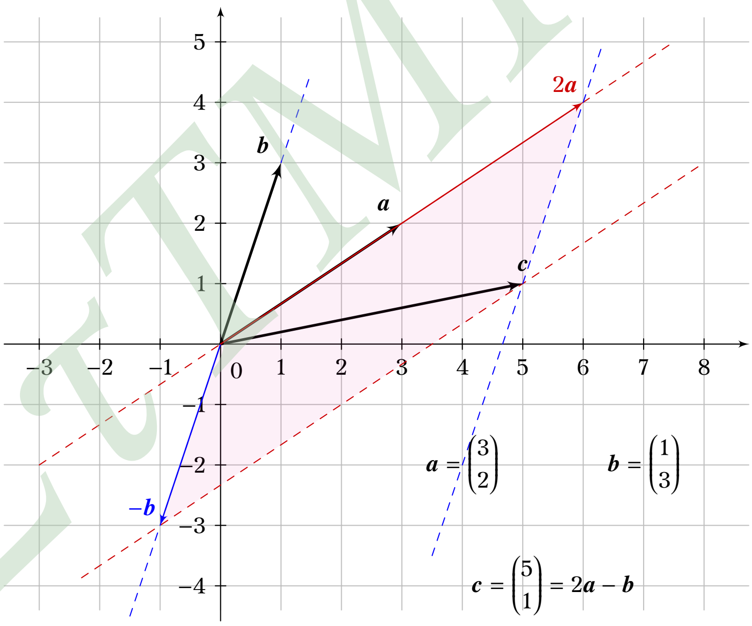
$$\begin{aligned} 3x + y &= 5 \\ 2x + 3y &= 1 \end{aligned} \qquad \Leftrightarrow \qquad \begin{pmatrix} 3 \\ 2 \end{pmatrix} x + \begin{pmatrix} 1 \\ 3 \end{pmatrix} y = \begin{pmatrix} 5 \\ 1 \end{pmatrix}.$$

Θεώρηση κατά γραμμές: Οι ευθείες έχουν διαφορετικές κλίσεις κατά συνέπεια τέμνονται σε ένα σημείο. Πράγματι στο Σχήμα 3.1 αποτυπώνεται αυτό το αποτέλεσμα και η λύση του συστήματος είναι η $x = 2$, $y = -1$.



Σχήμα 3.1: Γραφική επίλυση συστήματος Ι.

Θεώρηση κατά στήλες: Τα διανύσματα των συντελεστών, $\boldsymbol{a}$ και $\boldsymbol{b}$ δεν περιέχονται στην ίδια ευθεία, επομένως το διάνυσμα των σταθερών $\boldsymbol{c}$ εκφράζεται ως γραμμικός συνδυασμός αυτών των διανυσμάτων. Οι ευθείες που περιέχουν αντίστοιχα τα $\boldsymbol{a}$ και $\boldsymbol{b}$ και οι παράλληλες προς αυτές από το σημείο $(c_1, c_2) = (5, 1)$ ορίζουν μοναδικό παραλληλόγραμμο με "διαγώνιο" το διάνυσμα $\boldsymbol{c}$, βλέπε Σχήμα 3.2. Από το σχήμα βλέπουμε ότι $\boldsymbol{c} = 2\boldsymbol{a} - \boldsymbol{b}$ γεγονός το οποίο ισοδυναμεί με την ύπαρξη της μοναδικής λύσης $x = 3$, $y = -1$ του συστήματος.



Σχήμα 3.2: Γραφική επίλυση συστήματος ΙΙ.

Παρόμοια μπορούμε να μελετήσουμε ένα σύστημα τριών γραμμικών εξισώσεων με τρεις αγνώστους, στο $\mathbb{R}^3$, όπου γεωμετρικά κοιτάμε τις πιθανές θέσεις τριών επιπέδων, κατά συνέπεια των πιθανών τομών-ευθειών, καθώς δύο τεμνόμενα επίπεδα τέμνονται κατά μήκος ευθείας, και η μη ύπαρξη, ή η ύπαρξη και μοναδικότητα ή μη λύσης για το σύστημα σχετίζεται με τη θέση των σχετικών ευθειών-τομών.

3.3 Η απαλοιφή Gauss και Gauss-Jordan

Ας υποθέσουμε ότι ξεκινώντας από ένα τυχαίο, επαυξημένο 4×6 μητρώο, το οποίο αντιστοιχεί σε κάποιο τυπικό σύστημα τεσσάρων εξισώσεων με έξι αγνώστους, καταλήγουμε με τη διαδικασία της απαλοιφής στο

$$\begin{pmatrix} 1 & 2 & 0 & 4 & 2 & 2 & 5 \\ 0 & 0 & 1 & 5 & 1 & 1 & 3 \\ 0 & 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.3}$$

το οποίο δεν είναι τριγωνικό, και ούτε θα μπορούσε να είναι, αλλά έχει τη συγκεκριμένη κλιμακωτή μορφή[1]. Παρατηρούμε ότι τα κυκλωμένα 1 στο μητρώο είναι σε θέσεις οδηγών.

Ορισμός 3.3. Θα λέμε ότι ένα μητρώο είναι σε **κλιμακωτή μορφή** (row-echelon form) αν έχει τις ιδιότητες:

(1) Εάν μια γραμμή του δεν αποτελείται εξ ολοκλήρου από μηδενικά, τότε το πρώτο μη μηδενικό στοιχείο στη γραμμή είναι το 1.

(2) Οι γραμμές που αποτελούνται εξ ολοκλήρου από μηδενικά είναι και οι τελευταίες γραμμές του μητρώου.

(3) Αν δύο διαδοχικές γραμμές περιέχουν 1 σε θέση οδηγού, τότε το 1 της επάνω γραμμής βρίσκεται στα αριστερά του 1 της κάτω γραμμής.

Θα λέμε ότι ένα μητρώο είναι σε **ανηγμένη κλιμακωτή μορφή** (reduced row-echelon form) αν επιπλέον των (1)–(3) έχει την ιδιότητα

(4) Εάν μια στήλη περιέχει 1 σε θέση οδηγού, τότε περιέχει μηδενικά στις υπόλοιπες γραμμές της στήλης.

Παρατηρούμε ότι το μητρώο στην (3.3) είναι σε κλιμακωτή μορφή αλλά δεν είναι σε ανηγμένη κλιμακωτή μορφή, αφού πάνω από τον οδηγό στην τρίτη γραμμή περιέχονται μη μηδενικά στοιχεία. Οι στοιχειώδεις πράξεις γραμμών

$$r_1 \to -2r_3 + r_1 \qquad \text{και} \qquad r_2 \to -1r_3 + r_2$$

οδηγούν στο

$$\begin{pmatrix} 1 & 2 & 0 & 4 & 0 & 0 & 1 \\ 0 & 0 & 1 & 5 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \tag{3.4}$$

το οποίο είναι σε ανηγμένη κλιμακωρή μορφή.

Ορισμός 3.4. Τη διαδικασία απαλοιφής που οδηγεί σε μια κλιμακωτή μορφή ονομάζουμε **απαλοιφή Gauss** (Gauss elimination), ενώ εκείνη που οδηγεί στην ανηγμένη κλιμακωτή μορφή ονομάζουμε **απαλοιφή Gauss-Jordan** (Gauss-Jordan elimination).

[1]Σε αναλογία με τη γραφική παράσταση μιας φθίνουσας κλιμακωτής συνάρτησης.

Παράδειγμα 3.3. Να βρεθεί μια κλιμακωτή μορφή για το μητρώο

$$\begin{pmatrix} 0 & 0 & -2 & 8 & 12 \\ 2 & 4 & -8 & 12 & 28 \\ 2 & 4 & -5 & -6 & 22 \end{pmatrix}.$$

Βήμα 1. Εντοπίζουμε την πρώτη από αριστερά στήλη η οποία περιέχει μη μηδενικά στοιχεία. Εδώ είναι η πρώτη στήλη. Αντικαθιστούμε την πρώτη γραμμή, αν είναι απαραίτητο, με άλλη γραμμή η οποία περιέχει μη μηδενικό στοιχείο στη θέση του οδηγού. Εδώ μεταθέτουμε την πρώτη με την δεύτερη γραμμή.

$$\begin{pmatrix} 2 & 4 & -8 & 12 & 28 \\ 0 & 0 & -2 & 8 & 12 \\ 2 & 4 & -5 & -6 & 22 \end{pmatrix}$$

Βήμα 2. Προσθέτουμε κατάλληλα πολλαπλάσια της πρώτης γραμμής στις αποκάτω γραμμές ώστε τα στοιχεία κάτω από τον οδηγό να είναι μηδενικά. Εδώ πολλαπλασιάζουμε την πρώτη γραμμή με -1 και την προσθέτουμε στην τρίτη γραμμή.

$$\begin{pmatrix} 2 & 4 & -8 & 12 & 28 \\ 0 & 0 & -2 & 8 & 12 \\ 0 & 0 & 3 & -18 & -6 \end{pmatrix}$$

Βήμα 3. Συνεχίζουμε με τη δεύτερη γραμμή να εκτελεί την απαλοιφή. Πολλαπλασιάζουμε την δεύτερη γραμμή με $3/2$ και την προσθέτουμε στην τρίτη γραμμή.

$$\begin{pmatrix} 2 & 4 & -8 & 12 & 28 \\ 0 & 0 & -2 & 8 & 12 \\ 0 & 0 & 0 & -6 & 12 \end{pmatrix}$$

Βήμα 4. Πολλαπλασιάζουμε κάθε γραμμή με κατάλληλη σταθερά ώστε να εμφανιστεί το 1 στη θέση του οδηγού. Πολλαπλασιάζουμε την πρώτη γραμμή με $1/2$, την δεύτερη γραμμή με $-1/2$ και την τρίτη γραμμή με $-1/6$.

$$\begin{pmatrix} 1 & 2 & -4 & 6 & 14 \\ 0 & 0 & 1 & -4 & -6 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix}$$

Το τελευταίο μητρώο είναι η ζητούμενη μορφή.

Παράδειγμα 3.4. Να βρεθεί η ανηγμένη κλιμακωτή μορφή για το μητρώο

$$\begin{pmatrix} 0 & 0 & -2 & 8 & 12 \\ 2 & 4 & -8 & 12 & 28 \\ 2 & 4 & -5 & -6 & 22 \end{pmatrix}.$$

Στο Παράδειγμα 3.3 βρήκαμε μια κλιμακωτή μορφή του μητρώου και από αυτήν συνεχίζουμε τη διαδικασία Gauss-Jordan. Αρχίζοντας από την τελευταία μη μηδενική γραμμή και ανεβαίνοντας προσθέτουμε κατάλληλα πολλαπλάσια κάθε γραμμής με οδηγό στις αποπάνω γραμμές ώστε τα στοιχεία πάνω από κάθε οδηγό να είναι μηδενικά.

$$\begin{pmatrix} 1 & 2 & -4 & 6 & 14 \\ 0 & 0 & 1 & -4 & -6 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix}$$

Βήμα 5α. Πολλαπλασιάζουμε την τρίτη γραμμή με 4 και την προσθέτουμε στην δεύτερη.

$$\begin{pmatrix} 1 & 2 & -4 & 6 & 14 \\ 0 & 0 & 1 & 0 & -14 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix}$$

Βήμα 5β. Πολλαπλασιάζουμε την τρίτη γραμμή με −6 και την προσθέτουμε στην πρώτη.

$$\begin{pmatrix} 1 & 2 & -4 & 0 & 26 \\ 0 & 0 & 1 & 0 & -14 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix}$$

Βήμα 6. Πολλαπλασιάζουμε την δεύτερη γραμμή με 4 και την προσθέτουμε στην πρώτη.

$$\begin{pmatrix} 1 & 2 & 0 & 0 & -30 \\ 0 & 0 & 1 & 0 & -14 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix}$$

Το τελευταίο μητρώο είναι η ζητούμενη μορφή.

Παρατήρηση 3.2. Παρατηρούμε ότι για ένα γραμμικό ομοιογενές σύστημα $A\boldsymbol{x} = \mathbf{0}$, n εξισώσεων με $n+1$ αγνώστους στην καλύτερη περίπτωση, δηλαδή αυτή στην οποία υπάρχουν n οδηγοί, η ανηγμένη κλιμακωτή μορφή R_0 του επαυξημένου $n \times (n+2)$ μητρώου $\begin{pmatrix} A & \mathbf{0} \end{pmatrix}$ είναι

$$R_0 = \begin{pmatrix} 1 & 0 & 0 & \cdots & 0 & a_1 & 0 \\ 0 & 1 & 0 & \cdots & 0 & a_2 & 0 \\ 0 & 0 & 1 & \cdots & 0 & a_3 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & a_n & 0 \end{pmatrix},$$

απ΄ όπου φαίνεται (εδώ μπορούμε να σκεφτούμε την ομοιογενή περίπτωση του Παραδείγματος 3.1), ότι υπάρχουν άπειρες το πλήθος μη μηδενικές λύσεις.

Παρατήρηση 3.3. Η διαδικασία της απαλοιφής που παράγει την ανηγμένη κλιμακωτή μορφή ενός μητρώου $m \times n$ μπορεί να περιγραφεί αλγοριθμικά ως

Βήμα 1 Εντοπίζουμε την πρώτη από αριστερά στήλη του μητρώου η οποία περιέχει μη μηδενικά στοιχεία και αντικαθιστούμε την πρώτη γραμμή, αν είναι απαραίτητο, με άλλη γραμμή η οποία περιέχει μη μηδενικό στοιχείο στη θέση του οδηγού.

Βήμα 2 Προσθέτουμε κατάλληλα πολλαπλάσια της πρώτης γραμμής στις αποκάτω γραμμές ώστε τα στοιχεία κάτω από τον οδηγό να γίνουν μηδέν.

Βήμα 3α Εκτελούμε τα Βήματα 1 και 2 στο μητρώο που προέκυψε μετά την ολοκλήρωση του Βήματος 2 αγνοώντας την πρώτη γραμμή και την πρώτη στήλη.

Βήμα 3β Εκτελούμε το Βήμα 3α στο μητρώο που προέκυψε μετά την ολοκλήρωση του Βήματος 3α.

Συνεχίζουμε μέχρις ότου αποκαλυφθούν όλοι οι οδηγοί του μητρώου.

Βήμα 4 Πολλαπλασιάζουμε κάθε γραμμή που περιέχει οδηγό με τον αντίστροφο του οδηγού ώστε στη θέση των οδηγών να εμφανιστεί το 1.

Βήμα 5 Προσθέτουμε κατάλληλα πολλαπλάσια κάθε γραμμής που περιέχει οδηγό στις αποπάνω γραμμές ώστε και τα στοιχεία πάνω από τον οδηγό να γίνουν μηδενικά.

Σημειώνουμε ότι η κλιμακωτή μορφή μητρώου δεν είναι μοναδική.

► **Άσκηση 3.1.** Επαναλάβατε και ολοκληρώστε την διαδικασία απαλοιφής στο μητρώο του Παραδείγματος 3.3 όπου τώρα στο Βήμα 2 μεταθέστε την πρώτη και την τρίτη γραμμή. Τι παρατηρείτε;

3.3.1 Μοναδικότητα της ανηγμένης κλιμακωτής μορφής

Σε αντίθεση με την κλιμακωτή μορφή μητρώου η ανηγμένη κλιμακωτή μορφή είναι μοναδική. Ως προετοιμασία για την απόδειξη αυτού του αποτελέσματος ας ξεκινήσουμε θεωρώντας τα δύο μητρώα τα οποία είναι σε ανηγμένη κλιμακωτή μορφή

$$R_{01} = \begin{pmatrix} 1 & 2 & 0 & 4 & 0 & 0 \\ 0 & 0 & 1 & 5 & 0 & 2 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \qquad R_{02} = \begin{pmatrix} 1 & 2 & 0 & 4 & 0 & 2 \\ 0 & 0 & 1 & 5 & 0 & 3 \\ 0 & 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

και διαφέρουν κατά την έκτη στήλη. Στη συνέχεια διαμορφώνουμε τα μητρώα με τελευταία στήλη την στήλη κατά την οποία διαφέρουν και με προηγούμενες τις κοινές στήλες οδηγών. Αυτά είναι, αντίστοιχα τα

$$R'_{01} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}, \qquad R'_{02} = \begin{pmatrix} 1 & 0 & 0 & 2 \\ 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Σημειώνουμε ότι η πρώτη μη μηδενική στήλη ενός $m \times n$ μητρώου σε ανηγμένη κλιμακωτή μορφή είναι το διάνυσμα του $\mathbb{R}^m$ με 1 στην πρώτη γραμμή και 0 στις υπόλοιπες.

Θεώρημα 3.1. *Η ανηγμένη κλιμακωτή μορφή μητρώου είναι μοναδική.*

Απόδειξη. Αν $R_{01} \neq R_{02}$ είναι ανηγμένες κλιμακωτές μορφές ενός μητρώου, επιλέγουμε την πρώτη από αριστερά στήλη κατά την οποία τα δύο μητρώα διαφέρουν και τις προηγούμενες αυτής στήλες οδηγών και διαμορφώνουμε, όπως στην εισαγωγή της παραγράφου, τα μητρώα R'_{01} και R'_{02}. Τότε $R'_{01} \neq R'_{02}$. Στη συνέχεια θεωρούμε το καθένα από αυτά τα μητρώα ως το επαυξημένο μητρώο κάποιου συστήματος. Τότε θα έχουμε ότι

$$R'_{0i} = \left(\begin{array}{c|c} I & \boldsymbol{u}_i \\ \hline O & \mathbf{0} \end{array}\right), \qquad \text{ή} \qquad R'_{0i} = \left(\begin{array}{c|c} I & \mathbf{0} \\ \hline O & \boldsymbol{e}_1 \end{array}\right), \qquad i = 1, 2 \qquad \text{και} \qquad \boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

όπου I είναι το ταυτοτικό μητρώο και O το μηδενικό μητρώο, κατάλληλων διαστάσεων και τα δύο. Επειδή τα R_{01} και R_{02} είναι γραμμοϊσοδύναμα, άρα και τα R'_{01} και R'_{02}, αφού η παράλειψη στηλών δεν επηρεάζει τους συνδυασμούς των γραμμών, τα δύο συστήματα είτε έχουν κοινή λύση, είτε είναι αδύνατα και τα δύο. Έτσι στη μεν πρώτη περίπτωση θα είναι $\boldsymbol{u}_1 = \boldsymbol{u}_2$, στη δε δεύτερη τα δύο μητρώα είναι του δεύτερου τύπου, κατά συνέπεια και στις δύο περιπτώσεις θα είναι $R'_{01} = R'_{02}$. Αυτό όμως είναι άτοπο και καταλήξαμε σε άτοπο γιατί υποθέσαμε ότι $R_{01} \neq R_{02}$. Επομένως δεν υπάρχει στήλη κατά την οποία να διαφέρουν τα R_{01} και R_{02}, επομένως $R_{01} = R_{02}$. □

3.3.2 Στοιχειώδη μητρώα

Το αποτέλεσμα της εφαρμογής μιας στοιχειώδους πράξης σ΄ ένα μητρώο μπορεί να προκύψει με πολλαπλασιασμό του μητρώου με κατάλληλο μητρώο. Ειδικά η διαδικασία της απαλοιφής μπορεί να να επιτευχθεί πολλαπλασιάζοντας διαδοχικά το αρχικό μητρώο με κατάλληλα μητρώα. Κάθε τέτοιο μητρώο είναι αντίστοιχο μιας από τις στοιχειώδεις πράξεις γραμμών και ονομάζεται στοιχειώδες μητρώο (elementary matrix).

Ορισμός 3.5. Κάθε τετραγωνικό μητρώο το οποίο προκύπτει από το μοναδιαίο μητρώο I εκτελώντας μία μόνο στοιχειώδη πράξη γραμμών λέγεται **στοιχειώδες μητρώο**, ή μητρώο Gauss.

Παράδειγμα 3.5. Τα παρακάτω μητρώα είναι στοιχειώδη.

$$(\alpha') \begin{pmatrix} 1 & 0 \\ 0 & 2 \end{pmatrix}, \qquad (\beta') \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad (\gamma') \begin{pmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \qquad (\delta') \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}.$$

Το μητρώο στο (α΄) προκύπτει από το ταυτοτικό μητρώο με πολλαπλασιασμό της δεύτερης γραμμής του ταυτοτικού με το 2. Ειδικά το μητρώο στο (β΄) είναι ένα μητρώο μετάθεσης το οποίο προέρχεται από το I με μία μόνο εναλλαγή γραμμών, ενώ το μητρώο στο (γ΄) προκύπτει αν στη δεύτερη γραμμή του I προστεθεί το τριπλάσιο της πρώτης γραμμής.

Ας υποθέσουμε ότι το στοιχειώδες μητρώο E προέρχεται από το I εκτελώντας μια στοιχειώδη πράξη γραμμών. Αν E' είναι το στοιχειώδες μητρώο το οποίο προέρχεται από το I εκτελώντας την αντίστροφη στοχειώδη πράξη γραμμών ισχυριζόμαστε ότι $E^{-1} = E'$. Για παράδειγμα το (α΄) στο Παράδειγμα 3.5 προέρχεται από το I αν πολλαπλασιάσουμε τη δεύτερη γραμμή με 2. Αν αντίστροφα πολλαπλασιάσουμε την δεύτερη γραμμή του I με 1/2 προκύπτει το

$$\begin{pmatrix} 1 & 0 \\ 0 & 1/2 \end{pmatrix}$$

για το οποίο ισχύει

$$\begin{pmatrix}1 & 0\\ 0 & 1/2\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 2\end{pmatrix}\begin{pmatrix}1 & 0\\ 0 & 1/2\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

Όμοια το (ε΄)

$$\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}$$

προέρχεται από το I αν πολλαπλασιάσουμε τη δεύτερη γραμμή με 2 και την προσθέσουμε στην πρώτη γραμμή. Αν αντίστροφα πολλαπλασιάσουμε τη δεύτερη γραμμή με -2 και την προσθέσουμε στην πρώτη προκύπτει το

$$\begin{pmatrix}1 & -2\\ 0 & 1\end{pmatrix}$$

και τα δύο μητρώα είναι το ένα αντίστροφο του άλλου αφού

$$\begin{pmatrix}1 & -2\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 2\\ 0 & 1\end{pmatrix}\begin{pmatrix}1 & -2\\ 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0\\ 0 & 1\end{pmatrix}.$$

Γενικότερα, αν με r_i, $i = 1, 2, \ldots, n$ συμβολίσουμε τις γραμμές ενός μητρώου A, τότε η πράξη

$$r_k \to r_k' = r_k + \lambda r_j \to r_k'' = r_k' - \lambda r_j = r_k + \lambda r_j - \lambda r_j = r_k$$

υλοποιείται μέσω δύο κατάλληλων μητρώων, το ένα αντίστροφο του άλλου. Για παράδειγμα η πράξη $r_3 \to r_3 + \lambda r_2$, για A ένα $4 \times m$ μητρώο, επιτυγχάνεται με τον πολλαπλασιασμό με το στοιχειώδες μητρώο $E(\lambda)$

$$E(\lambda)A = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & \lambda & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}r_1\\ r_2\\ r_3\\ r_4\end{pmatrix} = \begin{pmatrix}r_1\\ r_2\\ r_3 + \lambda r_2\\ r_4\end{pmatrix}$$

και

$$E(-\lambda)E(\lambda)A = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & -\lambda & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix}\begin{pmatrix}r_1\\ r_2\\ r_3 + \lambda r_2\\ r_4\end{pmatrix} = \begin{pmatrix}r_1\\ r_2\\ (r_3 + \lambda r_2) - \lambda r_2\\ r_4\end{pmatrix} = \begin{pmatrix}r_1\\ r_2\\ r_3\\ r_4\end{pmatrix} = A.$$

Αναλυτικά γράφοντας το στοιχειώδες μητρώο $E(-\lambda)$ ως

$$E(-\lambda) = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & -\lambda & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} = \begin{pmatrix}1 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 1 & 0\\ 0 & 0 & 0 & 1\end{pmatrix} - \lambda\begin{pmatrix}0 & 0 & 0 & 0\\ 0 & 0 & 0 & 0\\ 0 & 1 & 0 & 0\\ 0 & 0 & 0 & 0\end{pmatrix} = I - \lambda N,$$

όπου I είναι το 4×4 ταυτοτικό μητρώο και N το μητρώο δεξιά του λ, τότε όντως

$$E(-\lambda)E(\lambda) = (I - \lambda N)(I + \lambda N) = I + \lambda N - \lambda N + \lambda N - \lambda^2 N^2 = I,$$

αφού όπως εύκολα διαπιστώνεται $N^2 = O$.

Επίσης αν με M_{ij} συμβολίσουμε το μητρώο μετάθεσης που εναλλάσσει τις γραμμές i και j του I, τότε $M_{ij}^{-1} = M_{ji} = M_{ij}$, πράγματι για το (β΄) έχουμε

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

Διατυπώνουμε λοιπόν δίχως απόδειξη, αλλά χωρίς φόβο και πάθος, το

Θεώρημα 3.2. *Κάθε στοιχειώδες μητρώο E είναι αντιστρέψιμο και το E^{-1} είναι επίσης στοιχειώδες μητρώο.*

Παράδειγμα 3.6. Για το μητρώο A του Παραδείγματος 3.3 δείχνουμε πώς πολλαπλασιάζοντάς το διαδοχικά με το στοιχειωδώδες μητρώο που αντιστοιχεί στο κάθε βήμα της απαλοιφής Gauss καταλήγουμε στην κλιμακωτή μορφή, έστω R, που υπολογίσαμε. Έτσι υπολογίζουμε

$$E_1A = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 0 & 0 & -2 & 8 & 12 \\ 2 & 4 & -8 & 12 & 28 \\ 2 & 4 & -5 & -6 & 22 \end{pmatrix} = \begin{pmatrix} 2 & 4 & -8 & 12 & 28 \\ 0 & 0 & -2 & 8 & 12 \\ 2 & 4 & -5 & -6 & 22 \end{pmatrix} = A_1$$

$$E_2A_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 2 & 4 & -8 & 12 & 28 \\ 0 & 0 & -2 & 8 & 12 \\ 2 & 4 & -5 & -6 & 49 \end{pmatrix} = \begin{pmatrix} 2 & 4 & -8 & 12 & 28 \\ 0 & 0 & -2 & 8 & 12 \\ 0 & 0 & 3 & -18 & -6 \end{pmatrix} = A_2$$

$$E_3A_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 3/2 & 1 \end{pmatrix}\begin{pmatrix} 2 & 4 & -8 & 12 & 28 \\ 0 & 0 & -2 & 8 & 12 \\ 0 & 0 & 3 & -18 & -6 \end{pmatrix} = \begin{pmatrix} 2 & 4 & -8 & 12 & 28 \\ 0 & 0 & -2 & 8 & 12 \\ 0 & 0 & 0 & -6 & 12 \end{pmatrix} = A_3$$

$$E_4A_3 = \begin{pmatrix} 1/2 & 0 & 0 \\ 0 & -1/2 & 0 \\ 0 & 0 & -1/6 \end{pmatrix}\begin{pmatrix} 2 & 4 & -8 & 12 & 28 \\ 0 & 0 & -2 & 8 & 12 \\ 0 & 0 & 0 & -6 & 12 \end{pmatrix} = \begin{pmatrix} 1 & 2 & -4 & 6 & 14 \\ 0 & 0 & 1 & -4 & -6 \\ 0 & 0 & 0 & 1 & -2 \end{pmatrix} = R.$$

Παρατήρηση 3.4 (Σημαντική). Στο Παράδειγμα 3.6 δείξαμε ότι πολλαπλασιάζοντας διαδοχικά ένα μητρώο A με κατάλληλο στοιχειώδες μητρώο καταλήγουμε σε μια κλιμακωτή μορφή R του αρχικού μητρώου, συγκεκριμένα δείξαμε ότι

$$E_4E_3E_2E_1A = R,$$

κατά συνέπεια αφού κάθε στοιχειώδες μητρώο αντιστρέφεται

$$A = LR \tag{3.5}$$

όπου

$$L = E_1^{-1}E_2^{-1}E_3^{-1}E_4^{-1}.$$

Θα λέμε ότι δύο μητρώα $n \times m$ είναι **γραμμοϊσοδύναμα** (row equivelent) εάν το ένα προκύπτει από το άλλο με διαδοχική εφαρμογή ενός πεπερασμένου πλήθους στοιχειωδών πράξεων γραμμών. Επειδή για κάθε στοιχειώδες μητρώο L_k το L_k^{-1} είναι επίσης στοιχειώδες η σχέση (3.5) εκφράζει ακριβώς ότι κάθε μητρώο είναι γραμμοϊσοδύναμο με κάθε κλιμακωτή μορφή του, και κατ΄ επέκταση γραμμοϊσοδύναμο με την ανηγμένη κλιμακωτή μορφή του. Θα συμβολίζουμε μια κλιμακωτή μορφή μητρώου με R και με R_0 την ανηγμέμη κλιμακωτή μορφή του.

Παρατήρηση 3.5 (Κάτω-άνω τριγωνική παραγοντοποίηση, ή LU παραγοντοποίηση). Ας θεωρήσουμε το τετραγωνικό μητρώο

$$B = \begin{pmatrix} 2 & -8 & 12 \\ 0 & -2 & 8 \\ 2 & -5 & -6 \end{pmatrix}.$$

Είναι εκείνο που προκύπτει από το μητρώο A_1 στο Παράδειγμα 3.6 αν διαγράψουμε τη δεύτερη και τη πέμπτη στήλη. Η διαδικασία της απαλοιφής Gauss, **κατά την οποία δεν υπάρχει μετάθεση γραμμών** καταλήγει στο αποτέλεσμα

$$E_3E_2B = \begin{pmatrix} 2 & -8 & 12 \\ 0 & -2 & 8 \\ 0 & 0 & -6 \end{pmatrix} = U,$$

όπου το U είναι ένα άνω τριγωνικό μητρώο, ισοδύναμα $B = E_2^{-1}E_3^{-1}U$ όπου

$$E_2^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{pmatrix}, \qquad E_3^{-1} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 3/2 & 1 \end{pmatrix}.$$

Το γινόμενο στα αριστερά του U υπολογίζεται εύκολα αν αναλογιστούμε τον τρόπο δράσης ενός στοιχειώδους μητρώου στο μητρώο το οποίο πολλαπλασιάζει από αριστερά, έτσι βρίσκουμε

$$\begin{pmatrix} 2 & -8 & 12 \\ 0 & -2 & 8 \\ 2 & -5 & -6 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & -3/2 & 1 \end{pmatrix}\begin{pmatrix} 2 & -8 & 12 \\ 0 & -2 & 8 \\ 0 & 0 & -6 \end{pmatrix}, \qquad \text{ή} \qquad B = LU.$$

Παρατηρούμε ότι το μητρώο B εκφράζεται ως γινόμενο ένός κάτω τριγωνικού (lower triangular) μητρώου L και ενός άνω τριγωνικού (upper triangular) μητρώου U. Το αποτέλεσμα αυτό εκφράζεται ως **κάτω-άνω τριγωνική παραγοντοποίηση**, ή και **LU παραγοντοποίηση** (Lower-Upper decomposition or factorization).

Το αποτέλεσμα της Παρατήρησης 3.5 γενικεύεται. Το διατυπώνουμε υπό μορφή Θεωρήματος. Η απόδειξή του αναπαράγει τα βήματα του παραδείγματος της Παρατήρησης 3.5 και παραλείπεται.

Θεώρημα 3.3. *Εάν A είναι ένα τετραγωνικό μητρώο του οποίου η κλιμακωτή μορφή προκύπτει δίχως εναλλαγή γραμμών, τότε το μητρώο παραγοντοποιείται ως $A = LU$, όπου το L είναι ένα κάτω τριγωνικό μητρώο και το U ένα άνω τριγωνικό μητρώο.*

Στη γενική περίπτωση, αν για το τετραγωνικό μητρώο A κατά τη διαδικασία της απαλοιφής υπάρχει εναλλαγή γραμμών, τότε για κάποιο μητρώο μετάθεσης P ισχύει $PA = LU$ (γιατί;).

3.4 Το αντίστροφο μητρώο

Έστω A ένα τετραγωνικό μητρώο και ας θεωρήσουμε το σύστημα $A\boldsymbol{x} = \boldsymbol{b}$. Αν το A είναι αντιστρέψιμο, τότε πολλαπλασιάζοντας από αριστερά με A^{-1} παίρνουμε

$$A^{-1}(A\boldsymbol{x}) = A^{-1}\boldsymbol{b} \Rightarrow (A^{-1}A)\boldsymbol{x} = A^{-1}\boldsymbol{b} \Rightarrow \boldsymbol{x} = A^{-1}\boldsymbol{b}, \tag{3.6}$$

αφού $A^{-1}A = I$, κατά συνέπεια το σύστημα έχει μοναδική λύση για κάθε $\boldsymbol{b}$. Ειδικά αν $\boldsymbol{b} = \mathbf{0}$ και το A είναι αντιστρέψιμο, τότε το σύστημα $A\boldsymbol{x} = \mathbf{0}$ έχει μοναδική λύση την μηδενική $\boldsymbol{x} = \mathbf{0}$. Αποδεικνύεται ότι ισχύει και το αντίστροφο, δηλαδή αν για το τετραγωνικό μητρώο A το σύστημα $A\boldsymbol{x} = \mathbf{0}$ έχει μοναδική λύση, την $\boldsymbol{x} = \mathbf{0}$, τότε το A είναι αντστρέψιμο.

Θεώρημα 3.4. *Αν A είναι ένα τετραγωνικό μητρώο, τα παρακάτω είναι ισοδύναμα*

(1) *Το A είναι αντιστρέψιμο.*

(2) *Το σύστημα $A\boldsymbol{x} = \mathbf{0}$ έχει μοναδική λύση την μηδενική.*

(3) *Το A είναι γραμμοϊσοδύναμο με το ταυτοτικό μητρώο I.*

(4) *Αν R_0 είναι η ανηγμένη κλιμακωτή μορφή του A, τότε $R_0 = I$.*

Απόδειξη. Αποδεικνύουμε τις συνεπαγωγές (1) ⇒ (2) ⇒ (3) ⇒ (4) ⇒ (1).

(1) ⇒ (2) Αν το A^{-1} υπάρχει και το $\boldsymbol{x}$ είναι τέτοιο ώστε $A\boldsymbol{x} = \mathbf{0}$, τότε όπως στην (3.6) έπεται ότι $\boldsymbol{x} = A^{-1}\mathbf{0} = \mathbf{0}$, ισοδύναμα η μοναδική λύση του συστήματος $A\boldsymbol{x} = \mathbf{0}$ = είναι η μηδενική.

(2) ⇒ (3) Επιλύοντας το σύστημα $A\boldsymbol{x} = \mathbf{0}$ με απαλοιφή Gauss-Jordan καταλήγουμε, από την υπόθεση στο $\boldsymbol{x} = \mathbf{0}$, ή $I\boldsymbol{x} = \mathbf{0}$, κατά συνέπεια το A είναι γραμμοϊσοδύναμο με το ταυτοτικό μητρώο I.

(3) ⇒ (4) Από τον ορισμό της ανηγμένης κλιμακωτής μορφής έπεται ότι (3) ⇔ (4), κατά συνέπεια το ζητούμενο ισχύει.

(4) ⇒ (1) Αφού $R_0 = I$ υπάρχουν στοιχειώδη μητρώα $E_1, \ldots, E_k$ ώστε $E_k \cdots E_1 A = I$, κατά συνέπεια

$$A = E_1^{-1} \cdots E_k^{-1} I = (E_k \cdots E_1)^{-1}.$$

Έτσι το A είναι ίσο με ένα αντιστρέψιμο μητρώο, επομένως είναι και το ίδιο αντιστρέψιμο.

Συνεπώς τα (1), (2), (3), (4) είναι ισοδύναμα. □

3.4.1 Υπολογισμός του αντίστροφου

Το Θεώρημα 3.4 μας δίνει ένα συστηματικό τρόπο ελέγχου κατά πόσον το αντίστροφο μητρώου υπάρχει και την ίδια στιγμή και τρόπο υπολογισμού του. Αν λοιπόν για το τετραγωνικό μητρώο A υπάρχει το αντίστροφο, έστω X τότε $AX = I$, επιπλέον υπάρχουν στοιχειώδη μητρώα $E_1, \ldots, E_k$ τα οποία διεκπεραιώνουν την απαλοιφή Gauss-Jordan ώστε $E_k \cdots E_1 A = R_0 = I$. Έτσι θα έχουμε

$$\begin{aligned} AX = I &\Leftrightarrow E_k \cdots E_1 AX = E_k \cdots E_1 I \\ &\Leftrightarrow IX = E_k \cdots E_1 \\ &\Leftrightarrow X = L \end{aligned}$$

όπου $L = E_k \cdots E_1$. Έτσι ξεκινώντας από το μητρώο $\begin{pmatrix} A & I \end{pmatrix}$ με τη διαδικασία απαλοιφής παίρνουμε

$$E_k \cdots E_1 \begin{pmatrix} A & I \end{pmatrix} = \begin{pmatrix} E_k \cdots E_1 A & E_k \cdots E_1 I \end{pmatrix} = \begin{pmatrix} I & E_k \cdots E_1 \end{pmatrix}$$

ή σχηματικά

$$\begin{pmatrix} A & I \end{pmatrix} \xrightarrow{\text{Gauss-Jordan}} \begin{pmatrix} I & L \end{pmatrix}$$

και το μητρώο $L = E_k \cdots E_1$ είναι το A^{-1}. Κατά συνέπεια αν για κάποιο τετραγωνικό μητρώο A ξεκινώντας από το σύνθετο μητρώο $\begin{pmatrix} A & I \end{pmatrix}$ και εφαρμόζοντας στοιχειώδεις πράξεις γραμμών καταλήξουμε σ΄ ένα μητρώο $\begin{pmatrix} I & L \end{pmatrix}$, τότε το A αντιστρέφεται και $A^{-1} = L$.

Παράδειγμα 3.7. Σύμφωνα με την Άσκηση 2.8 έχουμε

$$\text{αν} \qquad A = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}, \qquad \text{τότε} \qquad A^{-1} = \begin{pmatrix} 2 & -1 \\ -1 & 1 \end{pmatrix}.$$

Δείχνουμε το αποτέλεσμα με τη διαδικασία που περιγράψαμε. Πράγματι διαμορφώνοντας το μητρώο $\begin{pmatrix} A & I \end{pmatrix}$ με τις κατάλληλες στοιχειώδεις πράξεις παίρνουμε

$$\begin{pmatrix} 1 & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \end{pmatrix} \xrightarrow[r_2 \to -r_1 + r_2]{r_1 \to r_1} \begin{pmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{pmatrix} \xrightarrow[r_2 \to r_2]{r_1 \to -r_2 + r_1} \begin{pmatrix} 1 & 0 & 2 & -1 \\ 0 & 1 & -1 & 1 \end{pmatrix}.$$

Στο αριστερό μπλοκ του τελευταίου μητρώου εμφανίζεται το μοναδιαίο μητρώο I και στο δεξιό μπλοκ βλέπουμε το αναμενόμενο A^{-1}.

Παρατήρηση 3.6. Αποδεικνύουμε τον τύπο του αντίστροφου μητρώου, για 2×2 μητρώα, με την διαδικασία που ακολουθήσαμε στο Παράδειγμα 3.7. Θυμίζουμε ότι στην Άσκηση 2.8 ζητήθηκε να δειχθεί ότι για το μητρώο

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

υπάρχει το αντίστροφο αν και μόνο αν η ορίζουσα του A, $J = ad - bc$, είναι διάφορη του μηδενός, και αν αυτό ισχύει, τότε

$$A^{-1} = \frac{1}{J} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Έστω $J \neq 0$ και έστω $a \neq 0$, τότε

$$\begin{aligned} \begin{pmatrix} a & b & 1 & 0 \\ c & d & 0 & 1 \end{pmatrix} &\xrightarrow[r_2 \to -(c/a) r_1 + r_2]{r_1 \to r_1} \begin{pmatrix} a & b & 1 & 0 \\ 0 & d - bc/a & -c/a & 1 \end{pmatrix} \\ &\xrightarrow[r_2 \to (a/J) r_2]{r_1 \to (1/a) r_1} \begin{pmatrix} 1 & b/a & 1/a & 0 \\ 0 & 1 & -c/J & a/J \end{pmatrix} \\ &\xrightarrow[r_2 \to r_2]{r_1 \to -(b/a) r_2 + r_1} \begin{pmatrix} 1 & 0 & 1/a + bc/aJ & -b/J \\ 0 & 1 & -c/J & a/J \end{pmatrix} = \begin{pmatrix} 1 & 0 & d/J & -b/J \\ 0 & 1 & -c/J & a/J \end{pmatrix} \end{aligned}$$

απ΄ όπου προκύπτει το ζητούμενο. Αν $a = 0$, τότε $cb \neq 0$, αφού $J \neq 0$, οπότε μεταθέτοντας τις γραμμές του A και ακολουθώντας την ίδια διαδικασία καταλήγουμε στο ανάλογο αποτέλεσμα.

Παράδειγμα 3.8. Να βρεθεί το αντίστροφο του μητρώου

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 5 & 3 \\ 1 & 0 & 8 \end{pmatrix}$$

εφόσον αυτό υπάρχει.

Διαμορφώνουμε το μητρώο $\begin{pmatrix} A & I \end{pmatrix}$ και εκτελούμε κατάλληλες στοιχειώδεις πράξεις γραμμών ώστε να καταλήξουμε σε ένα μητρώο $\begin{pmatrix} I & X \end{pmatrix}$. Έτσι έχουμε

$$\begin{aligned}
\begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 5 & 3 & 0 & 1 & 0 \\ 1 & 0 & 8 & 0 & 0 & 1 \end{pmatrix} &\xrightarrow[r_3 \to -r_1 + r_3]{r_2 \to -2r_1 + r_2} \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & -2 & 5 & -1 & 0 & 1 \end{pmatrix} \\
&\xrightarrow[r_3 \to 2r_2 + r_3]{r_1 \to -2r_2 + r_1} \begin{pmatrix} 1 & 0 & 9 & 5 & -2 & 0 \\ 0 & 1 & -3 & -2 & 1 & 0 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{pmatrix} \\
&\xrightarrow[r_2 \to -3r_3 + r_2]{r_1 \to 9r_3 + r_1} \begin{pmatrix} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & -1 & -5 & 2 & 1 \end{pmatrix} \\
&\xrightarrow[r_3 \to -r_3]{} \begin{pmatrix} 1 & 0 & 0 & -40 & 16 & 9 \\ 0 & 1 & 0 & 13 & -5 & -3 \\ 0 & 0 & 1 & 5 & -2 & -1 \end{pmatrix}
\end{aligned}$$

κατά συνέπεια, και ελέγχεται εύκολα ότι είναι όντως έτσι,

$$A^{-1} = \begin{pmatrix} -40 & 16 & 9 \\ 13 & -5 & -3 \\ 5 & -2 & -1 \end{pmatrix}.$$

Παράδειγμα 3.9. Να βρεθεί, εφόσον αυτό υπάρχει, το αντίστροφο του μητρώου

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 5 \\ -1 & 3 & 1 \end{pmatrix}$$

Ξεκινώντας από το μητρώο $\begin{pmatrix} A & I \end{pmatrix}$ με τη διαδικασία της απαλοιφής βρίσκουμε

$$\begin{pmatrix} 1 & 0 & 2 & 1 & 0 & 0 \\ 2 & 1 & 5 & 0 & 1 & 0 \\ -1 & 3 & 1 & 0 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 3 & 3 & 1 & 0 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 2 & 1 & 0 & 0 \\ 0 & 1 & 1 & -2 & 1 & 0 \\ 0 & 0 & 0 & 7 & 3 & 1 \end{pmatrix}.$$

Παρατηρούμε ότι η διαδικασία της απαλοιφής Gauss καταλήγει σε ένα μητρώο με μηδενική γραμμή στο αριστερό μπλοκ όπου αναμένεται το ταυτοτικό μητρώο. Κατά συνέπεια το μητρώο δεν αντιστρέφεται.

Παρατήρηση 3.7 (Σχετικά με το Παράδειγμα 3.9). Τι συμβαίνει και το μητρώο A του Παραδείγματος 3.9 δεν αντιστρέφεται; Ενα από αυτά που συμβαίνουν είναι ότι η ανηγμένη κλιμακωτή μορφή του A δεν είναι το ταυτοτικό μητρώο, βλέπε Θεώρημα 3.4. Τι άλλο μπορεί να συμβαίνει; Ας δοκιμάσουμε να λύσουμε το σύστημα $A\boldsymbol{x} = \boldsymbol{0}$.

Παράδειγμα 3.10. Εάν A είναι το μητρώο του Παραδείγματος 3.9 να λυθεί το σύστημα $A\boldsymbol{x} = \boldsymbol{0}$.

Εκτελώντας στοιχειώδεις πράξεις γραμμών στο μητρώο $\begin{pmatrix} A & \mathbf{0} \end{pmatrix}$ παίρνουμε

$$\begin{pmatrix} 1 & 0 & 2 & 0 \\ 2 & 1 & 5 & 0 \\ -1 & 3 & 1 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 3 & 3 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 2 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Έτσι αν $\boldsymbol{x} = \begin{pmatrix} x_1 & x_2 & x_3 \end{pmatrix}^{\mathrm{T}}$, το τελευταίο μητρώο αντιστοιχεί στο σύστημα

$$\left.\begin{aligned} x_1 + 2x_3 &= 0 \\ x_2 + x_3 &= 0 \end{aligned}\right\} \Leftrightarrow \begin{cases} x_1 = -2t \\ x_2 = -t \\ x_3 = t \end{cases} \qquad t \in \mathbb{R}.$$

Έτσι το σύστημα, όπως αναμενόταν εξάλλου υπό το πρίσμα του Θεωρήματος 3.4, έχει μη μηδενικές λύσεις, άπειρες το πλήθος.

Χρησιμοποιώντας τα αποτελέσματα του Θεωρήματος 3.4 δείχνουμε οτι για την ύπαρξη του αντιστρόφου τετραγωνικού μητρώου A αρκεί να δείξουμε ότι υπάρχει τετραγωνικό μητρώο B για το οποίο ισχύει ένα από τα δύο, $AB = I$, ή $BA = I$.

Θεώρημα 3.5. *Αν για το $n \times n$ μητρώο A υπάρχει $n \times n$ μητρώο B ώστε $AB = I$, τότε θα είναι $BA = I$, συνεπώς το A αντιστρέφεται και $A^{-1} = B$.*

Απόδειξη. Έστω ότι $BA = I$. Για να δείξουμε ότι το A είναι αντιστρέψιμο αρκεί να δείξουμε ότι το σύστημα $A\boldsymbol{x} = \mathbf{0}$ έχει μοναδική λύση την μηδενική. Θεωρούμε λοιπόν το σύστημα $A\boldsymbol{x} = \mathbf{0}$ και πολλαπλασιάζοντας με B παίρνουμε διαδοχικά

$$BA\boldsymbol{x} = B\mathbf{0} \Rightarrow I\boldsymbol{x} = \mathbf{0} \Rightarrow \boldsymbol{x} = \mathbf{0},$$

κατά συνέπεια το A είναι αντιστρέψιμο. Και πάλι από την υπόθεση $BA = I$ παίρνουμε

$$BAA^{-1} = IA^{-1} \Rightarrow BI = A^{-1} \Rightarrow B = A^{-1}.$$

Έστω τώρα ότι $AB = I$, τότε από το πρώτο μέρος της απόδειξης έπεται ότι η μοναδική λύση του συστήματος $B\boldsymbol{x} = \mathbf{0}$ είναι η μηδενική, κατά συνέπεια το B είναι αντιστρέψιμο και $A = B^{-1}$. Το A όντας ίσο με το αντιστρέψιμο μητρώο B^{-1} είναι αντιστρέψιμο και $A^{-1} = (B^{-1})^{-1} = B$. □

3.5 Ασκήσεις

1. Με τη μέθοδο της απαλοιφής να επιλυθεί καθένα από τα συστήματα βρίσκοντας την ανηγμένη κλιμακωτή μορφή του αντίστοιχου επαυξημένου μητρώου

(α΄) $$\begin{aligned} x + 3y - 2z &= -1 \\ 2x + y + z &= 3 \\ -x + 2y + z &= -4 \end{aligned}$$

(β΄) $$\begin{aligned} 3x + y - z &= 3 \\ x + 3y + z &= 2 \\ 5x - y - 3z &= 4 \end{aligned}$$

2. Δίνονται τα σημεία $P(1, 2)$, $Q(-1, 6)$, και $R(2, 3)$.

(α΄) Να βρεθεί πολυώνυμο $p(x) = ax^2+bx+c$ του οποίου το γράφημα περιέχει τα σημεία P, Q, και R.

(β΄) Υπάρχουν πολυώνυμα $q(x) = ax^3 + bx^2 + cx + d$ των οποίων το γράφημα περιέχει τα σημεία P, Q, και R;

3. Δίνεται το σύστημα

$$\begin{aligned} x + 2y + 3z &= 1 \\ 2x + 3y + 4z &= 2 \\ 3x + 5y + 7z &= a \end{aligned}$$

όπου a είναι πραγματική παράμετρος. Προσδιορίστε την παράμετρο a ώστε το σύστημα να έχει (i) καμία λύση, (ii) μία μόνο λύση, ή (iii) άπειρες λύσεις. Στην περίπτωση ύπαρξης λύσης ή λύσεων να βρεθούν αυτές.

4. Για κάθε ένα από τα συστήματα

(α΄) $$\begin{aligned} x + y + 2z &= a \\ 2x + 3y + 3z &= 2 \\ x + 2y + \mu z &= b \end{aligned}$$

(β΄) $$\begin{aligned} x + 2y + z &= a \\ 2x - 2y + 3z &= 2 \\ x + 2y + a^2 z &= b \end{aligned}$$

(γ΄) $$\begin{aligned} x + 4y - 2z &= 1 \\ 2y - 4z &= 5 \\ y + az &= \mu \end{aligned}$$

όπου a, b και μ είναι πραγματικές παράμετροι, να βρείτε συνθήκες για τις παραμέτρους ώστε το σύστημα να έχει (i) καμία λύση, (ii) μία μόνο λύση, ή (iii) άπειρες λύσεις.

5. Να βρεθεί μια άνω-κάτω τριγωνική παραγοντοποίηση για το μητρώο

$$\begin{pmatrix} 1 & 2 & 4 \\ 3 & 8 & 14 \\ 2 & 6 & 11 \end{pmatrix}.$$

6. Για ποιές τρεις τιμές της παραμέτρου λ είναι το μητρώο

$$A(\lambda) = \begin{pmatrix} 2 & \lambda & \lambda \\ \lambda & \lambda & \lambda \\ 8 & 7 & \lambda \end{pmatrix}$$

μη αντιστρέψιμο; Εξηγήστε σε κάθε περίπτωση γιατί συμβαίνει αυτό.

7. Έστω ότι το μητρώο

$$M(a, b) = \begin{pmatrix} a & 0 & b & 2 \\ a & a & 4 & 4 \\ 0 & a & 2 & b \end{pmatrix}$$

είναι το επαυξημένο μητρώο για ένα γραμμικό σύστημα τριών εξισώσεων με τρεις αγνώστους. Προσδιορίστε τις παραμέτρους a και b ώστε το σύστημα να έχει

(α΄) Καμία λύση.

(β΄) Μοναδική λύση.

(γ΄) Μια μονοπαραμετρική οικογένεια λύσεων.

(δ΄) Μια διπαραμετρική οικογένεια λύσεων.

8. Δείξτε ότι για $a \neq 0$ και $a \neq b$ το μητρώο

$$M = \begin{pmatrix} a & b & b \\ a & a & b \\ a & a & a \end{pmatrix}$$

είναι αντιστρέψιμο και βρείτε το αντίστροφο.

9. Να βρεθούν όλες οι τιμές της παραμέτρου λ ώστε το μητρώο

$$A(\lambda) = \begin{pmatrix} 3-\lambda & 2 & \lambda^2 \\ 0 & \lambda+2 & \lambda-2 \\ 0 & 0 & \lambda \end{pmatrix}$$

να είναι αντιστρέψιμο.

10. Έστω ότι τα A και B είναι τετραγωνικά μητρώα του ιδίου μεγέθους.

(α΄) Αν C είναι επίσης τετραγωνικό μητρώο ώστε

$$A(B-C) = I,$$

δείξτε ότι το A είναι αντιστρέψιμο και ότι $B = A^{-1} + C$.

(β΄) Εάν το $D = AB$ είναι αντιστρέψιμο, δείξτε ότι και το A είναι αντιστρέψιμο.

11. Τα στοιχεία κάθε γραμμής ενός 4×4 μητρώου M είναι μια διάταξη των αριθμών $-2, -1, 1, 2$. Δείξτε ότι ένα τέτοιο μητρώο δεν έχει αντίστροφο.
Υπόδειξη: Το άθροισμα των στοιχείων κάθε γραμμής είναι ίσο με μηδέν.

12. Εάν A και B είναι $n \times n$ μητρώα και το AB είναι αντιστρέψιμο, δείξτε ότι τα A και B είναι αντιστρέψιμα.[2]

13. Ας υποθέσουμε ότι τα A και B είναι $n \times n$ μητρώα ώστε $AB = I$. Από την Άσκηση 12 έπεται ότι υπάρχουν τα A^{-1} και B^{-1}. Δείξτε ότι $A = B^{-1}$, συνεπώς και $B = A^{-1}$.[3]

14. Εάν A και B είναι $n \times n$ μητρώα ώστε $A + B = AB$ δείξτε ότι τα μητρώα αντιμετατίθενται, δείξτε δηλαδή ότι $AB = BA$.

[2]Για μη τετραγωνικά μητρώα A και B είναι δυνατόν το AB να είναι αντιστρέψιμο δίχως ο ίδιος χαρακτηρισμός να έχει ένοια για το A ή το B.

[3]Η Άσκηση δείχνει ότι αν για τα τετραγωνικά μητρώα A και B ισχύει $AB = I$, τότε $BA = I$. Συγκρίνετε με το Ορισμό 1.20 και το Θεώρημα 3.5.

Κεφάλαιο 4

Διανυσματικοί χώροι

4.1 Διανυσματικοί χώροι

Είδαμε ότι εφοδιάζοντας το $\mathbb{R}^n$, ή το $\mathbb{M}^{n,m}$ με την πράξη της πρόσθεσης διανυσμάτων, ή μητρώων και του πολλαπλασιασμού διανύσματος επί σταθερά, ή μητρώου επί σταθερά διαμορφώσαμε τις δομές $(\mathbb{R}^n, +, \cdot)$ και $(\mathbb{M}^{n,m}, +, \cdot)$ για τις οποίες, γράφοντας X αντί για $\mathbb{R}^n$ ή $\mathbb{M}^{n,m}$ και δηλώνοντας τα διανύσματα ή τα μητρώα με συνήθεις χαρκτήρες, ισχύουν οι νόμοι

(V1) $x + y \in X$, για κάθε $x, y \in X$

(V2) $x + y = y + x$, για κάθε $x, y \in X$

(V3) $x + (y + z) = (x + y) + z$, για κάθε $x, y, z \in X$

(V4) υπάρχει $0 \in X$ τέτοιο ώστε $0 + x = x$ για κάθε $x \in X$

(V5) για κάθε $x \in X$ υπάρχει $-x \in X$ έτσι ώστε $-x + x = 0$

(V6) $\lambda x \in X$, για κάθε $x \in X$ και $\lambda, \in \mathbb{K}$

(V7) $\lambda(\mu x) = (\lambda\mu)x$, για κάθε $x \in X$ και $\lambda, \mu \in \mathbb{K}$

(V8) $1x = x$, για κάθε $x \in X$

(V9) $\lambda(x + y) = \lambda x + \lambda y$, για κάθε $x, y \in X$ και $\lambda \in \mathbb{K}$

(V10) $(\lambda + \mu)x = \lambda x + \mu x$, για κάθε $x \in X$ και $\lambda, \mu \in \mathbb{K}$

όπου με $\mathbb{K}$ δηλώνουμε το σώμα των πραγματικών, ή μιγαδικών αριθμών. Στη συνέχεια δίνουμε δύο επιπλέον παραδείγματα όπου γνωστά σύνολα με τις πράξεις της πρόσθεσης και του πολλαπλασιασμού με σταθερά αποκτούν δομή ανάλογη με αυτή των $(\mathbb{R}^n, +, \cdot)$ και $(\mathbb{M}^{n,m}, +, \cdot)$.

Παράδειγμα 4.1. Όπως γνωρίζουμε το άθροισμα δύο πραγματικών συναρτήσεων και το γινόμενο μιας συνάρτησης με πραγματική σταθερά ορίζονται με τις σχέσεις

$$(f + g)(x) = f(x) + g(x), \qquad \text{και} \qquad (\lambda f)(x) = \lambda f(x). \tag{4.1}$$

Αν $C[0,1]$ είναι το σύνολο των πραγματικών συνεχών συναρτήσεων ορισμένων στο διάστημα $[0,1]$ και $f, g \in C[0,1]$, τότε $f + g \in C[0,1]$ και $\lambda f \in C[0,1]$, για κάθε $\lambda \in \mathbb{R}$, αφού το άθροισμα συνεχών συναρτήσεων είναι συνεχής συνάρτηση στο κοινό πεδίο ορισμού τους καθώς και το γινόμενο συνεχούς συνάρτησης με σταθερά είναι επίσης συνεχής συνάρτηση. Επιπλέον οι νόμοι (V1)-(V10), με $\mathbf{0}$ να είναι η μηδενική συνάρτηση, δηλαδή $\mathbf{0}(t) = 0$ για κάθε $t \in [0,1]$, ικανοποιούνται σχεδόν αυτόματα αφού οι πράξεις μεταξύ συναρτήσεων μεταφέρονται μέσω της (4.1) στις πράξεις της πρόσθεσης και του πολλαπλασιασμού στους πραγματικούς αριθμούς. Κατά συνέπεια η δομή $(C[0,1],+,\cdot)$ συμπεριφέρεται όπως η $(\mathbb{R}^n,+,\cdot)$.

Παράδειγμα 4.2. Αν n είναι ένας θετικός ακέραιος, με $\mathbb{P}_n$ συμβολίζουμε το σύνολο των πολυωνύμων με πραγματικούς ή μιγαδικούς συντελεστές βαθμού το πολύ n. Αν $p, q \in \mathbb{P}_n$, τότε

$$p(x) = a_0 + a_1x + a_2x^2 + \cdots + a_nx^n \qquad \text{και} \qquad q(x) = b_0 + b_1x + b_2x^2 + \cdots + b_nx^n$$

οπότε

$$\begin{aligned}(p+q)(x) &= p(x) + q(x) = (a_0 + b_0) + (a_1 + b_1)x + (a_2 + b_2)x^2 + \cdots + (a_n + b_n)x^n \\ &= c_0 + c_1x + c_2x^2 + \cdots + c_nx^n\end{aligned}$$

και αν λ είναι ένας πραγματικός αριθμός, τότε

$$\begin{aligned}(\lambda p)(x) &= \lambda p(x) = \lambda a_0 + \lambda a_1x + \lambda a_2x^2 + \cdots + \lambda a_nx^n \\ &= d_0 + d_1x + d_2x^2 + \cdots + d_nx^n\end{aligned}$$

με $c_k, d_k \in \mathbb{R}$ για $k = 1, 2, \ldots, n$. Επομένως $p + q \in \mathbb{P}_n$ και $\lambda p \in \mathbb{P}_n$. Όπως και στο Παράδειγμα 4.1 η ισχύς των νόμων (V1)-(V10) επαληθεύεται χωρίς δυσκολία.

Ορισμός 4.1. Έστω $\mathcal{V}$ ένα μη κενό σύνολο. Αν "+" και "·" είναι δύο πράξεις, πρόσθεση και πολλαπλασιασμός για τις οποίες ικανοποιούνται οι νόμοι (V1)-(V10) θα λέμε ότι η τριάδα $(\mathcal{V},+,\cdot)$ είναι ένας **πραγματικός διανυσματικός χώρος** (real vector space), ή απλά **διανυσματικός χώρος**, αν $\mathbb{K} = \mathbb{R}$. Αν $\mathbb{K} = \mathbb{C}$, δηλαδή οι σταθερές στον πολλαπλασιασμό είναι μιγαδικές ο διανυσματικός χώρος λέγεται **μιγαδικός** . Την εξάρτηση του διανυσματικού χώρου από το σώμα $\mathbb{K}$, ή από ένα γενικό σώμα $\mathbb{F}$ τη δηλώνουμε γράφοντας $\mathcal{V}(\mathbb{K})$, ή $\mathcal{V}(\mathbb{F})$, αντίστοιχα. Στη γενική περίπτωση $\mathcal{V}(\mathbb{F})$ λέμε ότι ο διανυσματικός χώρος $\mathcal{V}$ είναι ορισμένος επάνω στο σώμα $\mathbb{F}$. Στη συνέχεια όταν λέμε διανυσματικό χώρο θα εννοούμε, συνήθως, πραγματικό διανυσματικό χώρο, εκτός αν αναφέρεται διαφορετικά. Τα στοιχεία του διανυσματικού χώρου θα τα λέμε **διανύσματα** (vectors).

Έτσι οι $\mathbb{R}^n$, $\mathbb{M}^{n,m}(\mathbb{K})$, $C[0,1]$, και $\mathbb{P}_n(\mathbb{K})$ με πράξεις την πρόσθεση διανυσμάτων, μητρώων, ή συναρτήσεων-πολυωνύμων και τον πολλαπλασιασμό με πραγματικές ή μιγαδικές σταθερές, ανάλογα, είναι τυπικά παραδείγματα διανυσματικών χώρων. Στη συνέχεια τον τυχαίο διανυσματικό χώρο θα συμβολίζουμε με απλά κεφαλαία γράμματα του λατινικού αλφαβήτου, συνήθως, X, Y, Z.

► **Άσκηση 4.1.** Έστω $\mathbb{S}$ το σύνολο των ακολουθιών πραγματικών αριθμών. Μια τέτοια ακολουθία είναι μια συνάρτηση $a : \mathbb{N} \to \mathbb{R}$, την οποία συμβολίζουμε με (a_n), όπου αντί για

$a(n)$ γράφουμε a_n. Αν (a_n) και (b_n) είναι ακολουθίες, ορίζουμε το άθροισμά τους να είναι η ακολουθία $(a_n + b_n)$ και αν λ είναι ένας πραγματικός αριθμός ορίζουμε το γινόμενο της (a_n) με το λ ως την ακολουθία (λa_n), έχουμε δηλαδή τις (συνήθεις μεταξύ συναρτήσεων) πράξεις

$$(a_n) + (b_n) = (a_n + b_n), \qquad \lambda(a_n) = (\lambda a_n).$$

Δείξτε ότι το $\mathbb{S}$ με αυτές τις πράξεις είναι διανυσματικός χώρος.

Παρατήρηση 4.1. Αν $\mathbb{F}$ είναι ένα σώμα, τότε όπως στην περίπτωση του $\mathbb{R}^n$ ή του $\mathbb{C}^n$ η δομή των διανυσμάτων $(\mathbb{F}^n, +, \cdot)$ όπου για $\xi, \zeta \in \mathbb{F}^n$ και $\lambda \in \mathbb{F}$ κατ΄ αναλογία ορίζουμε

$$\boldsymbol{\xi} + \boldsymbol{\zeta} = \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} + \begin{pmatrix} \zeta_1 \\ \zeta_2 \\ \vdots \\ \zeta_n \end{pmatrix} = \begin{pmatrix} \xi_1 + \zeta_1 \\ \xi_2 + \zeta_2 \\ \vdots \\ \xi_n + \zeta_n \end{pmatrix}, \qquad \lambda\boldsymbol{\xi} = la \begin{pmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{pmatrix} = \begin{pmatrix} \lambda\xi_1 \\ \lambda\xi_2 \\ \vdots \\ \lambda\xi_n \end{pmatrix},$$

είναι διανυσματικός χώρος επάνω στο σώμα $\mathbb{F}$. Ειδικά το $\mathbb{F}$ μπορεί να θεωρηθεί ως διανυσματικός χώρος επί του εαυτού του.

▲ **Πρόταση 4.1.** *Σε κάθε διανυσματικό χώρο $(X, +, \cdot)$ ισχύουν τα*

(1) *Το μηδενικό διάνυσμα είναι μοναδικό.*

(2) *Το αντίθετο διάνυσμα είναι μοναδικό.*

(3) *Για κάθε $\boldsymbol{x} \in X$ είναι $0\boldsymbol{x} = \boldsymbol{0}$.*

(4) *Για κάθε $\boldsymbol{x} \in X$ είναι $(-1)\boldsymbol{x} = -\boldsymbol{x}$.*

Απόδειξη. Έστω $\boldsymbol{x}$ τυχαίο διάνυσμα του X.

(1) Αν $\boldsymbol{0}$ και $\boldsymbol{0}'$ είναι μηδενικά στοιχεία στον X, τότε

$$\boldsymbol{0} = \boldsymbol{0} + \boldsymbol{0}' = \boldsymbol{0}'.$$

(2) Αν $-\boldsymbol{x}$ και $\boldsymbol{x}'$ είναι αντίθετα διανύσματα του $\boldsymbol{x}$, τότε από τους νόμους του διανυσματικού χώρου έπεται

$$-\boldsymbol{x} = -\boldsymbol{x} + \boldsymbol{0} = -\boldsymbol{x} + (\boldsymbol{x} + \boldsymbol{x}') = (-\boldsymbol{x} + \boldsymbol{x}) + \boldsymbol{x}' = \boldsymbol{0} + \boldsymbol{x}' = \boldsymbol{x}'.$$

(3) Και πάλι από τους νόμους του διανυσματικού χώρου έπεται

$$0\boldsymbol{x} + \boldsymbol{x} = (0 + 1)\boldsymbol{x} = \boldsymbol{x}$$

κατά συνέπεια από την μοναδικότητα του μηδενικού διανύσματος έπεται το ζητούμενο.

(4) Παρόμοια

$$(-1)\boldsymbol{x} + \boldsymbol{x} = (-1 + 1)\boldsymbol{x} = 0\boldsymbol{x} = \boldsymbol{0}$$

και από την μοναδικότητα του αντιθέτου διανύσματος έπεται ότι $(-1)\boldsymbol{x} = -\boldsymbol{x}$. □

Παράδειγμα 4.3. Στο $\mathbb{R}^3$ θεωρούμε τα διανύσματα

$$\boldsymbol{u} = \begin{pmatrix} 1 & 1 & 0 \end{pmatrix}^{\mathrm{T}} \qquad \text{και} \qquad \boldsymbol{v} = \begin{pmatrix} 0 & 1 & 1 \end{pmatrix}^{\mathrm{T}}.$$

Δείχνουμε ότι το σύνολο όλων των γραμμικών συνδυασμών των $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανυσμα-

τικός χώρος.
Αν a και b είναι πραγματικοί αριθμοί τότε

$$a\boldsymbol{u} + b\boldsymbol{v} = a\begin{pmatrix}1\\1\\0\end{pmatrix} + b\begin{pmatrix}0\\1\\1\end{pmatrix} = \begin{pmatrix}a\\a+b\\b\end{pmatrix},$$

έτσι αν ονομάσουμε W το σύνολο των γραμμικών συνδυασμών τότε

$$W = \left\{\begin{pmatrix}x\\x+y\\y\end{pmatrix} : x, y \in \mathbb{R}\right\} = \left\{\begin{pmatrix}x\\y\\z\end{pmatrix} : x, z \in \mathbb{R} \quad \text{και} \quad y = x + z\right\}.$$

Αν $\boldsymbol{w}$ και $\boldsymbol{z}$ είναι διανύσματα στο W, τότε

$$\boldsymbol{w} = \begin{pmatrix}p\\p+q\\q\end{pmatrix} \qquad \text{και} \qquad \boldsymbol{z} = \begin{pmatrix}r\\r+s\\s\end{pmatrix}$$

οπότε για $\lambda \in \mathbb{R}$ έχουμε

$$\boldsymbol{w} + \boldsymbol{z} = \begin{pmatrix}p+r\\(p+q)+(r+s)\\q+s\end{pmatrix} = \begin{pmatrix}p+r\\(p+r)+(q+s)\\q+s\end{pmatrix}, \qquad \lambda\boldsymbol{w} = \begin{pmatrix}\lambda p\\\lambda(p+q)\\\lambda q\end{pmatrix} = \begin{pmatrix}\lambda p\\\lambda p+\lambda q\\\lambda q\end{pmatrix},$$

κατά συνέπεια το άθροισμα στοιχείων του W ανήκει στο W, όπως και κάθε πραγματικό πολλαπλάσιο στοιχείου του W ανήκει στο W. Το μηδενικό διάνυσμα $\mathbf{0}$ περιέχεται στο W αφού $\mathbf{0} = 0\boldsymbol{u} + 0\boldsymbol{v}$, ή αφού η δεύτερη συνιστώσα του $\mathbf{0}$ είναι ίση με το άθροισμα της πρώτης και της τρίτης $(0 = 0 + 0)$. Επίσης αν $\boldsymbol{w} \in W$, τότε για κάποια $a, b \in R$ είναι $\boldsymbol{w} = a\boldsymbol{u}+b\boldsymbol{v}$, οπότε $-\boldsymbol{w} = (-a)\boldsymbol{u}+(-b)\boldsymbol{v} \in W$. Οι υπόλοιποι νόμοι (V2), (V3) και (V7)-(V10) ικανοποιούνται αυτόματα αφού αφορούν σε ιδιότητες των πράξεων μεταξύ στοιχείων του $\mathbb{R}^3$, ή του $\mathbb{R}$ και του $\mathbb{R}^3$ και $W \subset \mathbb{R}^3$. Έτσι η δομή $(W, +, \cdot)$ είναι διανυσματικός χώρος.

4.2 Διανυσματικοί υπόχωροι

Από το Παράδειγμα 4.3 εξάγονται μερικά σημαντικά συμπεράσματα:

(Σ1) Ένα υποσύνολο ενός διανυσματικού χώρου μπορεί να είναι διανυσματικός χώρος με πράξεις τον περιορισμό των πράξεων του αρχικού χώρου στο υποσύνολο.

(Σ2) Το σύνολο όλων των γραμμικών συνδυασμών κάποιας συλλογής διανυσμάτων ενός διανυσματικού χώρου με πράξεις τον περιορισμό των πράξεων του αρχικού χώρου έχει τη δομή διανυσματικού χώρου.

Ορισμός 4.2. Εάν ο X είναι διανυσματικός χώρος και το W είναι υποσύνολο του X, $W \subseteq X$, θα λέμε ότι το W είναι **διανυσματικός υπόχωρος** (vector subspace) ή απλά **υπόχωρος** του X εάν ο περιορισμός των πράξεων του X στο W προσδίδει στο W τη δομή διανυσματικού χώρου.

Παράδειγμα 4.4. Το σύνολο $\mathbb{P}_n[0,1]$ των πολυωνύμων με πραγματικούς συντελεστές βαθμού το πολύ n στο $[0,1]$ με τις συνήθεις πράξεις είναι διανυσματικός χώρος και επειδή $\mathbb{P}_n[0,1] \subset C[0,1]$ έπεται ότι ο χώρος $\mathbb{P}_n[0,1]$ είναι διανυσματικός υπόχωρος του $C[0,1]$.

Θεώρημα 4.1. *Εάν ο X είναι διανυσματικός χώρος και $W \subseteq X$, ο W είναι υπόχωρος του X αν και μόνο αν για κάθε $\boldsymbol{u}$ και $\boldsymbol{v}$ στο W και για σταθερές λ και μ*

$$\lambda\boldsymbol{u} + \mu\boldsymbol{v} \in W.$$

Απόδειξη. Ας υποθέσουμε ότι ο W είναι υπόχωρος του X. Αν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα του W και λ και μ είναι πραγματικοί αριθμοί, τότε από τον ορισμό του διανυσματικού υπόχωρου έπεται ότι

$$\lambda\boldsymbol{u} \in W \quad \text{και} \quad \mu\boldsymbol{v} \in Wquad\text{κατά σθνέπεια} \quad \lambda\boldsymbol{u} + \mu\boldsymbol{v} \in W.$$

Ας υποθέσουμε τώρα ότι $W \subseteq X$, και ότι για κάθε $\boldsymbol{u}$ και $\boldsymbol{v}$ στο W και για κάθε λ και μ στο $\mathbb{R}$ είναι $\lambda\boldsymbol{u} + \mu\boldsymbol{v} \in W$. Έστω $\boldsymbol{x} \in W$ και έστω $\boldsymbol{y} \in W$, τότε

1. $\boldsymbol{x} + \boldsymbol{y} \in W$. Έπεται από την υπόθεση για $\lambda = \mu = 1$. (V1)
2. $\lambda\boldsymbol{x} \in W$ για κάθε $\lambda \in \mathbb{R}$. Έπεται από την υπόθεση για $\mu = 0$. (V6)
3. $\boldsymbol{0} \in W$. Έπεται από την υπόθεση για $\lambda = \mu = 0$. (V4)
4. $-\boldsymbol{x} \in W$. Έπεται από την υπόθεση για $\lambda = -1$ και $\mu = 0$. (V5)

Ο νόμος (V2) ικανοποιείται αφού $x + y \in W$ και $x + y = y + x$ στο X. Οι υπόλοιποι νόμοι αφορούν ιδιότητες των πράξεων κατά συνέπεια ικανοποιούνται, επομένως η δομή $(W, +, \cdot)$ είναι διανυσματικός χώρος, άρα υπόχωρος του X. □

Παράδειγμα 4.5. Εξετάστε αν το

$$U = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x, y \in \mathbb{R} \quad \text{και} \quad x + y = 0 \right\} = \left\{ \begin{pmatrix} x \\ -x \end{pmatrix} : x \in \mathbb{R} \right\}$$

είναι διανυσματικός υπόχωρος του $\mathbb{R}^2$.
Επικαλούμαστε το Θεώρημα 4.1. Αν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα του U, τότε

$$\boldsymbol{u} = \begin{pmatrix} x \\ -x \end{pmatrix}, \qquad \boldsymbol{v} = \begin{pmatrix} y \\ -y \end{pmatrix}, \qquad x, y \in \mathbb{R}$$

οπότε για $\lambda, \mu \in \mathbb{R}$ υπολογίζουμε

$$\lambda\boldsymbol{u} + \mu\boldsymbol{v} = \lambda \begin{pmatrix} x \\ -x \end{pmatrix} + \mu \begin{pmatrix} y \\ -y \end{pmatrix} = \begin{pmatrix} \lambda x + \mu y \\ -\lambda x - \mu y \end{pmatrix} = \begin{pmatrix} \lambda x + \mu y \\ -(\lambda x + \mu y) \end{pmatrix} = \begin{pmatrix} s \\ -s \end{pmatrix}$$

επομένως $\lambda\boldsymbol{u} + \mu\boldsymbol{v} \in U$, κατά συνέπεια το U είναι υπόχωρος του $\mathbb{R}^2$.

Παρατήρηση 4.2. Ένας διανυσματικός υπόχωρος περιέχει, αναγκαστικά, το μηδενικό διάνυσμα. Έτσι το υποσύνολο

$$V = \left\{ \begin{pmatrix} x \\ y \end{pmatrix} : x, y \in \mathbb{R} \quad \text{και} \quad x + y = 1 \right\} = \left\{ \begin{pmatrix} x \\ 1 - x \end{pmatrix} : x \in \mathbb{R} \right\}$$

δεν μπορεί να είναι υπόχωρος του $\mathbb{R}^2$, αφού

$$\begin{pmatrix} x \\ 1 - x \end{pmatrix} \neq \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

για όλα τα $x \in \mathbb{R}$. Επίσης αν $\boldsymbol{v} \in V$, τότε $\lambda \boldsymbol{v} \notin V$ για $\lambda \in \mathbb{R}$.

Ορισμός 4.3. Εάν ο X είναι διανυσματικός χώρος και $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ είναι διανύσματα του X, το σύνολο όλων των γραμμικών συνδυασμών των $\boldsymbol{u}_k$, $k = 1, 2, \ldots, n$ συμβολίζουμε με $\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$, έτσι

$$\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\} = \{\boldsymbol{v} : \boldsymbol{v} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_n\boldsymbol{u}_n,\ c_1, c_2, \ldots, c_n \in \mathbb{K}\}$$

και ονομάζουμε **διάνοιγμα** (span) των $\boldsymbol{u}_k$, $k = 1, 2, \ldots, n$. Αν $S \subseteq X$ με $\operatorname{span}(S)$ συμβολίζουμε το σύνολο όλων των γραμμικών συνδυασμών πεπερασμένου πλήθους στοιχείων του S.

Θεώρημα 4.2. *Έστω ότι τα $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ είναι διανύσματα του διανυσματικού χώρου X, τότε*

(1) *Το $\operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ είναι υπόχωρος του X.*

(2) *Εάν V είναι υπόχωρος του X και $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\} \subseteq V$, τότε ο $W = \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ είναι υπόχωρος του V, δηλαδή ο $\operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ είναι ο μικρότερος υπόχωρος ο οποίος περιέχει τα διανύσματα $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$.*

Απόδειξη. (1) Αν τα $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι στοιχεία του $\operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$, τότε

$$\boldsymbol{u} = c_1\boldsymbol{w}_1 + c_2\boldsymbol{w}_2 + \cdots + c_n\boldsymbol{w}_n \qquad \text{και} \qquad \boldsymbol{v} = c_1'\boldsymbol{w}_1 + c_2'\boldsymbol{w}_2 + \cdots + c_n'\boldsymbol{w}_n$$

για κάποιες σταθερές c_k, c_k', $k = 1, 2, \ldots, n$. Έτσι αν $\lambda, \mu \in \mathbb{R}$, τότε

$$\begin{aligned} \lambda\boldsymbol{u} + \mu\boldsymbol{v} &= \lambda(c_1\boldsymbol{w}_1 + c_2\boldsymbol{w}_2 + \cdots + c_n\boldsymbol{w}_n) + \mu(c_1'\boldsymbol{w}_1 + c_2'\boldsymbol{w}_2 + \cdots + c_n'\boldsymbol{w}_n) \\ &= (\lambda c_1 + \mu c_1')\boldsymbol{w}_1 + (\lambda c_2 + \mu c_2')\boldsymbol{w}_2 + \cdots + (\lambda c_n + \mu c_n')\boldsymbol{w}_n \\ &= c_1''\boldsymbol{w}_1 + c_2''\boldsymbol{w}_2 + \cdots + c_n''\boldsymbol{w}_n, \end{aligned}$$

όπου $c_k'' = \lambda c_k + \mu c_k'$, $k = 1, 2, \ldots, n$, κατά συνέπεια $\lambda\boldsymbol{u} + \mu\boldsymbol{v} \in \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ και το συμπέρασμα έπεται από το Θεώρημα 4.1.

(2) Αφού, από την υπόθεση, τα $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n$ είναι διανύσματα του διανυσματικού χώρου V το ζητούμενο έπεται από το (1). □

Παράδειγμα 4.6 (Η ευθεία ως διανυσματικός υπόχωρος). Εάν $\boldsymbol{u}$ είναι ένα μη μηδενικό διάνυσμα στο $\mathbb{R}^n$, τότε το $\mathrm{span}\{\boldsymbol{u}\}$ είναι υπόχωρος του $\mathbb{R}^n$ και αποτελείται απ΄ όλα τα διανύσματα της μορφής $\lambda\boldsymbol{u}$ με $\lambda \in \mathbb{R}$. Κατά συνέπεια είναι η ευθεία που περιέχει το $\boldsymbol{u}$, βλέπε Παρατήρηση 1.2, ας την πούμε $L_{\boldsymbol{u}}$. Η $L_{\boldsymbol{u}}$ μπορεί να παρασταθεί με την **διανυσματική** εξίσωση

$$\boldsymbol{r}(t) = t\boldsymbol{u}, \qquad t \in \mathbb{R}.$$

Ειδικά στο $\mathbb{R}^2$ αν $\boldsymbol{u} = \begin{pmatrix} p & q \end{pmatrix}^{\mathrm{T}}$ το διάνυσμα $\begin{pmatrix} x & y \end{pmatrix}^{\mathrm{T}}$ ανήκει στον $L_{\boldsymbol{u}}$ αν και μόνον αν

$$\begin{pmatrix} x \\ y \end{pmatrix} = t_0 \begin{pmatrix} p \\ q \end{pmatrix}$$

για κάποιο $t_0 \in \mathbb{R}$, ιδοδύναμα

$$x = pt_0 \qquad \text{και} \qquad y = qt_0$$

απ΄ όπου έπεται ότι $qx = py$, ή $qx - py = 0$. Καταλήγουμε λοιπόν στο συμπέρασμα ότι οι ευθείες που είναι υπόχωροι στο $\mathbb{R}^2$ έχουν **αλγεβρική** εξίσωση

$$ax + by = 0.$$

Σημειώνουμε ότι μια ευθεία με εξίσωση $ax + by + c = 0$, με $c \neq 0$, δεν μπορεί να είναι υπόχωρος του $\mathbb{R}^2$ γιατί δεν περιέχει το διάνυσμα $\boldsymbol{0}$.

4.3 Γραμμική ανεξαρτησία διανυσμάτων

Αν $\boldsymbol{a}$, και $\boldsymbol{b}$ είναι διανύσματα στο $\mathbb{R}^n$ και $\boldsymbol{a} = \beta\boldsymbol{b}$, για κάποια σταθερά β θα λέμε ότι τα δύο διανύσματα είναι γραμμικά εξαρτημένα. Στην περίπτωση αυτή μπορούμε να γράψουμε

$$(-1)\boldsymbol{a} + \beta\boldsymbol{b} = \boldsymbol{0}.$$

Ας υποθέσουμε στη συνέχεια ότι τα $\boldsymbol{a}$, και $\boldsymbol{b}$ είναι διανύσματα στο $\mathbb{R}^n$ και υπάρχουν σταθερές α και β, όχι και οι δύο μηδέν ώστε

$$\alpha\boldsymbol{a} + \beta\boldsymbol{b} = \boldsymbol{0}.$$

Ας υποθέσουμε ότι $\alpha \neq 0$, τότε θα είναι $\boldsymbol{a} = -(\beta/\alpha)\boldsymbol{b} = \lambda\boldsymbol{b}$, δηλαδή το ένα είναι πολλαπλάσιο του άλλου, δηλαδή τα δύο διανύσματα είναι γραμμικά εξαρτημένα.

Ορισμός 4.4. Θα λέμε ότι τα διανύσματα $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ του διανυσματικού χώρου $X(\mathbb{F})$ είναι **γραμμικά εξαρτημένα** (linearly dependent) εάν υπάρχουν σταθερές $c_1, c_2, \ldots, c_n \in \mathbb{F}$ όχι όλες ίσες με μηδέν ώστε

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_n\boldsymbol{u}_n = \boldsymbol{0}.$$

Ορισμός 4.5. Θα λέμε ότι τα διανύσματα $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ του διανυσματικού χώρου $X(\mathbb{F})$ είναι **γραμμικά ανεξάρτητα** (linearly independent) εάν οποτεδήποτε για σταθερές $c_1, c_2, \ldots, c_n \in \mathbb{F}$ ισχύει

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_n\boldsymbol{u}_n = \boldsymbol{0},$$

τότε $c_1 = c_2 = \cdots = c_n = 0$.

Παρατήρηση 4.3. Ας θεωρήσουμε τα διανύσματα $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ του διανυσματικού χώρου $X(\mathbb{F})$.

(1) Εάν τα διανύσματα είναι γραμμικά εξαρτημένα, τότε κάποιο ή κάποια από αυτά είναι γραμμικός συνδυασμός των υπολοίπων. Πράγματι υπάρχουν σταθερές $c_1, c_2, \ldots, c_n$ μία τουλάχιστον από τις οποίες είναι διάφορη του μηδενός ώστε

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_n\boldsymbol{u}_n = \mathbf{0}.$$

Έστω, για παράδειγμα, ότι $c_2 \neq 0$, τότε

$$\begin{aligned} \boldsymbol{u}_2 &= c_2^{-1}(-c_1\boldsymbol{u}_1 - c_3\boldsymbol{u}_3 - \cdots - c_n\boldsymbol{u}_n) \\ &= c_1'\boldsymbol{u}_1 + c_3'\boldsymbol{u}_3 + \cdots + c_n'\boldsymbol{u}_n, \end{aligned}$$

όπου $c_j' = -c_j c_2^{-1}$, $j = 1, 3, \ldots n$.

(2) Ισχύει και το αντίστροφο, δηλαδή αν κάποιο από τα διανύσματα είναι γραμμικός συνδυασμός των υπολοίπων, τότε τα διανύσματα είναι γραμμικά εξαρτημένα. Πράγματι αν, για παράδειγμα, υποθέσουμε ότι το διάνυσμα $\boldsymbol{u}_2$ είναι γραμμικός συνδυασμός των υπολοίπων, τότε

$$\boldsymbol{u}_2 = c_1\boldsymbol{u}_1 + c_3\boldsymbol{u}_3 + \cdots + c_n\boldsymbol{u}_n \Rightarrow c_1\boldsymbol{u}_1 + (-1)\boldsymbol{u}_2 + c_3\boldsymbol{u}_3 + \cdots + c_n\boldsymbol{u}_n = \mathbf{0},$$

δηλαδή τα διανύσματα είναι γραμμικά εξαρτημένα.

(3) Εάν κάποιο από τα διανύσματα είναι το μηδενικό, τότε τα διανύσματα είναι γραμμικά εξαρτημένα. Πράγματι υποθέτοντας, για παράδειγμα, ότι $\boldsymbol{u}_2 = \mathbf{0}$, τότε για $c_1 = c_3 = \cdots = c_n = 0$ και $c_2 \neq 0$ έχουμε

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_n\boldsymbol{u}_n = \mathbf{0}.$$

Παράδειγμα 4.7. Οι συναρτήσεις sin και cos είναι γραμμικά ανεξάρτητες στον χώρο $C[-\pi, \pi]$ ή $C(\mathbb{R})$.
Πράγματι αν ήταν γραμμικά εξαρτημένες τότε η μια θα ήταν πολλαπλάσιο της άλλης, ισοδύναμα για κάποια πραγματική σταθερά λ θα ήταν $\sin x = \lambda \cos x$, για κάθε x. Τότε όμως θα έπρεπε

$$\sin(\pi/2) = \lambda \cos(\pi/2) \Leftrightarrow 1 = \lambda 0$$

που είναι αδύνατο, άρα οι sin και cos είναι γραμμικά ανεξάρτητες.

Παράδειγμα 4.8. Τρία διανύσματα

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}, \qquad \boldsymbol{c} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix},$$

στο $\mathbb{R}^2$ είναι γραμμικά εξαρτημένα.
Περίπτωση 1: Τα $\boldsymbol{a}$ και $\boldsymbol{b}$ είναι γραμμικά εξαρτημένα. Τότε υπάρχουν σταθερές α, β με $\alpha \neq 0$, ή $\beta \neq 0$, ώστε $\alpha\boldsymbol{a} + \beta\boldsymbol{b} = \mathbf{0}$. Για $\gamma = 0$ είναι

$$\alpha\boldsymbol{a} + \beta\boldsymbol{b} + \gamma\boldsymbol{c} = \mathbf{0}$$

με $\alpha \neq 0$, ή $\beta \neq 0$, ισοδύναμα τα $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ είναι γραμμικά εξαρτημένα.
Περίπτωση 2: Τα $\boldsymbol{a}$ και $\boldsymbol{b}$ είναι γραμμικά ανεξάρτητα. Δείχνουμε ότι στη περίπτωση αυτή

$$a_1b_2 - a_2b_1 \neq 0. \tag{4.2}$$

Επειδή $\boldsymbol{a} \neq \boldsymbol{0}$ και $\boldsymbol{b} \neq \boldsymbol{0}$ θα πρέπει $b_1 \neq 0$, ή $b_2 \neq 0$ και $a_1 \neq 0$, ή $a_2 \neq 0$. Έστω $b_1 \neq 0$, τότε αφού $\boldsymbol{a} \neq \lambda \boldsymbol{b}$ για κάθε σταθερά λ,

$$\text{και} \qquad a_1 = \frac{a_1}{b_1} b_1 \qquad \text{θα πρέπει} \qquad a_2 \neq \frac{a_1}{b_1} b_2,$$

κατά συνέπεια $a_2b_1 \neq a_1b_2$. Στη συνέχεια δείχνουμε ότι υπάρχουν σταθερές α και β ώστε $\boldsymbol{c} = \alpha \boldsymbol{a} + \beta \boldsymbol{b}$, ή ισοδύναμα ότι το σύστημα

$$\begin{aligned} a_1\alpha + b_1\beta &= c_1 \\ a_2\alpha + b_2\beta &= c_2 \end{aligned} \tag{4.3}$$

με αγνώστους τα α και β έχει λύση. Πολλαπλασιάζοντας την πρώτη εξίσωση με b_2 τη δεύτερη με b_1 και αφαιρώντας βρίσκουμε

$$(a_1b_2 - a_2b_1)\alpha = b_2c_1 - b_1c_2 \Rightarrow \alpha = \frac{b_2c_1 - b_1c_2}{a_1b_2 - a_2b_1}, \tag{4.4}$$

αφού ο παρονομαστής είναι διάφορος του μηδενός. Όμοια πολλαπλασιάζοντας την πρώτη εξίσωση με a_2 τη δεύτερη με a_1 και αφαιρώντας βρίσκουμε

$$(a_1b_2 - a_2b_1)\beta = a_1c_2 - a_2c_1 \Rightarrow \beta = \frac{a_1c_2 - a_2c_1}{a_1b_2 - a_2b_1}. \tag{4.5}$$

Έτσι γι αυτά τα α και β έχουμε

$$\alpha \boldsymbol{a} + \beta \boldsymbol{b} + (-1)\boldsymbol{c} = \boldsymbol{0}$$

συνεπώς τα $\boldsymbol{a}, \boldsymbol{b}, \boldsymbol{c}$ είναι γραμμικά εξαρτημένα.

Παρατήρηση 4.4. Στο Παράδειγμα 4.8 δείξαμε ότι αν τα διανύσματα

$$\boldsymbol{a} = \begin{pmatrix} a_1 \\ a_2 \end{pmatrix}, \qquad \boldsymbol{b} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix},$$

είναι γραμμικά ανεξάρτητα τότε ισχύει η (βλέπε Παράδειγμα 2.15)

$$\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1b_2 - a_2b_1 \neq 0, \tag{4.6}$$

δηλαδή η ορίζουσα των συνιστωσών των διανυσμάτων είναι διαφορετική του μηδενός. Ισχύει και το αντίστροφο, δλαδή αν η ορίζουσα των συνιστωσών είναι διαφορετική του μηδενός, τότε τα διανύσματα είναι γραμμικά ανεξάρτητα. Πράγματι υποθέτοντας ότι ισχύει η (4.6) και θεωρώντας την εξίσωση

$$\alpha \boldsymbol{a} + \beta \boldsymbol{b} = \boldsymbol{0},$$

δηλαδή με $\boldsymbol{c} = \boldsymbol{0}$, από τις (4.4), (4.5) βρίσκουμε $\alpha = \beta = 0$ γεγονός που αποδεικνύει τον ισχυρισμό μας.

Σημείωση 4.1. Ένα από τα αποτελέσματα που δείξαμε στο Παράδειγμα 4.8 μπορεί να διατυπωθεί σε "οριζουσιακή" διάλεκτο ώς εξής: Ας θεωρήσουμε το γραμμικό σύστημα

$$\begin{aligned} a_1 x + b_1 y &= c_1 \\ a_2 x + b_2 y &= c_2 \end{aligned} \tag{4.7}$$

Εάν η ορίζουσα των συντελεστών

$$\det(\boldsymbol{a}\ \boldsymbol{b}) := \begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix} = a_1 b_2 - a_2 b_1 \tag{4.8}$$

είναι διάφορη του μηδενός, τότε το σύστημα (4.7) έχει μοναδική λύση η οποία δίνεται από τις σχέσεις

$$x = \frac{\det(\boldsymbol{c}\ \boldsymbol{b})}{\det(\boldsymbol{a}\ \boldsymbol{b})} = \frac{\begin{vmatrix} c_1 & b_1 \\ c_2 & b_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}, \qquad y = \frac{\det(\boldsymbol{a}\ \boldsymbol{c})}{\det(\boldsymbol{a}\ \boldsymbol{b})} = \frac{\begin{vmatrix} a_1 & c_1 \\ a_2 & c_2 \end{vmatrix}}{\begin{vmatrix} a_1 & b_1 \\ a_2 & b_2 \end{vmatrix}}. \tag{4.9}$$

Το αποτέλεσμα γενικεύεται και σε περισσότερες διαστάσεις (σύστημα n γραμμικών εξισώσεων με n αγνώστους). Ο τύπος στην (4.9) λέγεται κανόνας του Cramer. Θα επανέλθουμε σε επόμενο κεφάλαιο.

Το αποτέλεσμα του Παραδείγματος 4.8 γενικεύεται σε κάθε χώρο $\mathbb{R}^n$.

Θεώρημα 4.3. *Εάν $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{n+1}$, είναι διανύσματα στο $\mathbb{R}^n$, τότε είναι γραμμικά εξαρτημένα.*

Για την απόδειξη του θεωρήματος πρέπει να δείξουμε ότι υπάρχουν σταθερές $c_1, c_2, \ldots, c_{n+1}$ όχι όλες ίσες με μηδέν ώστε

$$c_1 \boldsymbol{u}_1 + c_2 \boldsymbol{u}_2 + \cdots + c_{n+1} \boldsymbol{u}_{n+1} = \boldsymbol{0}. \tag{4.10}$$

Γράφοντας

$$\boldsymbol{u}_j = \begin{pmatrix} a_1^j & a_2^j & \ldots & a_n^j \end{pmatrix}^{\mathrm{T}}, \qquad j = 1, 2, \ldots, n+1,$$

(γράφουμε a_i^j αντί για a_{ij}), βλέπουμε ότι η εξίσωση (4.10) είναι ισοδύναμη με το γραμμικό σύστημα

$$\begin{aligned} a_1^1 c_1 + a_1^2 c_2 + \cdots + a_1^{n+1} c_{n+1} &= 0 \\ a_2^1 c_1 + a_2^2 c_2 + \cdots + a_2^{n+1} c_{n+1} &= 0 \\ &\vdots \\ a_n^1 c_1 + a_n^2 c_2 + \cdots + a_n^{n+1} c_{n+1} &= 0 \end{aligned} \tag{4.11}$$

n εξισώσεων με $n+1$ αγνώστους, τις σταθερές c_j, $j = 1, 2, \ldots, n+1$. Κατά συνέπεια αρκεί, ισοδύναμα, να δείξουμε ότι το σύστημα (4.11) έχει μη μηδενική λύση. Αποδεικνύουμε λοιπόν το ισοδύναμο

Θεώρημα 4.4. *Κάθε ομοιογενές γραμμικό σύστημα n εξισώσεων με $n+1$ αγνώστους έχει μη μηδενική λύση.*

Απόδειξη. Το επαυξημένο μητρώο για το σύστημα (4.11) περιέχει το πολύ n, οδηγούς, βλέπε Παρατήρηση 3.2, κατά συνέπεια έχει άπειρες το πλήθος λύσεις, συνεπώς και μη μηδενική. □

▼ **Πόρισμα 4.1.** *Εάν $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n$, είναι διανύσματα στο $\mathbb{R}^m$, και $n > m$ τότε τα διανύσματα είναι γραμμικά εξαρτημένα.*

Απόδειξη. Επειδή $n \geq m+1$, από το Θεώρημα 4.3 έπεται ότι υπάρχουν σταθερές $c_1, c_2, \dots, c_{m+1}$ με τουλάχιστον μία $c_k \neq 0$ ώστε $c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_{m+1}\boldsymbol{u}_{m+1} = \boldsymbol{0}$. Τότε η εξίσωση

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_{m+1}\boldsymbol{u}_{m+1} + 0\boldsymbol{u}_{m+2} + \cdots + 0\boldsymbol{u}_n = \boldsymbol{0}$$

έχει μη μηδενική λύση. □

4.4 Βάση και διάσταση διανυσματικού χώρου

Έχουμε δει, βλέπε Παρατήρηση 1.3, ότι κάθε διάνυσμα $\boldsymbol{u} \in \mathbb{R}^3$ εκφράζεται ως γραμμικός συνδυασμός των διανυσμάτων

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

έτσι ώστε αν u_1, u_2, u_3 είναι οι συνιστώσες του $\boldsymbol{u}$, τότε

$$\boldsymbol{u} = u_1\boldsymbol{e}_1 + u_2\boldsymbol{e}_2 + u_3\boldsymbol{e}_3.$$

Από την τελευταία σχέση έπεται ότι αν $c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2 + c_3\boldsymbol{e}_3 = \boldsymbol{0}$, τότε $c_1 = c_2 = c_3 = 0$ (γιατί;) δηλαδή τα διανύσματα $\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3$ είναι γραμμικά ανεξάρτητα.

Ορισμός 4.6. Θα λέμε ότι τα διανύσματα $\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n$ ενός διανυσματικού χώρου $X(\mathbb{F})$ αποτελούν μια **βάση** (basis) του διανυσματικού χώρου αν

(1) Τα διανύσματα παράγουν το χώρο, δηλαδή $\operatorname{span}\{\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n\} = X$, και

(2) Τα διανύσματα $\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n$ είναι γραμμικά ανεξάρτητα.

Παράδειγμα 4.9. Ας θεωρήσουμε τον χώρο $\mathbb{M}^{2,2}$. Όπως στη περίπτωση των διανυσμάτων του $\mathbb{R}^n$ γράφουμε

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \begin{pmatrix} a & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & b \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ c & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & d \end{pmatrix} \\ &= a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}. \end{aligned}$$

Έπεται λοιπόν ότι τα μητρώα

$$\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

παράγουν το χώρο των 2×2 μητρώων, δηλαδή

$$\mathbb{M}^{2,2} = \operatorname{span}\left\{\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}\right\}.$$

Επιπλέον τα μητρώα αυτά είναι γραμμικά ανεξάρτητα αφού

$$a\begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix} + b\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} + c\begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix} + d\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

ισοδύναμα $a = b = c = d = 0$, κατά συνέπεια αποτελούν μια βάση για τον $\mathbb{M}^{2,2}$.

► **Άσκηση 4.2 (Τριγωνικά μητρώα).** Ας θεωρήσουμε το χώρο $\mathbb{M}^{3,3}$.

(α΄) Δείξτε ότι τόσο τα άνω τριγωνικά 3×3 μητρώα, όσο και τα κάτω τριγωνικά 3×3 μητρώα αποτελούν υπόχωρους του $\mathbb{M}^{3,3}$.

(β΄) Δείξτε ότι τα μητρώα

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

αποτελούν μια βάση για των υπόχωρο των 3×3 άνω τριγωνικών μητρώων.

(γ΄) Να βρεθεί μια βάση του χώρου των 3×3 κάτω τριγωνικών μητρώων.

► **Άσκηση 4.3.** Να βρεθεί μια βάση του χώρου των 2×2 συμμετρικών μητρώων.

Παράδειγμα 4.10. Τα διανύσματα

$$\boldsymbol{b}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{b}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{b}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

αποτελούν μια βάση του $\mathbb{R}^3$.
Πράγματι γράφοντας

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = a\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + b\begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} + c\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} a+b+c \\ b+c \\ c \end{pmatrix}$$

βλέπουμε ότι για $c = z$, $b = y - z$, $a = x - y$ για το τυχαίο διάνυσμα $\boldsymbol{u} = \begin{pmatrix} x & y & z \end{pmatrix}^{\mathrm{T}}$ έχουμε

$$\boldsymbol{u} = (x-y)\boldsymbol{b}_1 + (y-z)\boldsymbol{b}_2 + z\boldsymbol{b}_3,$$

δηλαδή $\operatorname{span}\{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\} = \mathbb{R}^3$. Επιπλέον από την εξίσωση $a\boldsymbol{b}_1 + b\boldsymbol{b}_2 + c\boldsymbol{b}_3 = \boldsymbol{0}$ βρίσκουμε

$$\begin{pmatrix} a+b+c \\ b+c \\ c \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{cases} a+b+c = 0 \\ b+c = 0 \\ c = 0 \end{cases}$$

απ΄ όπου με "προς τα πίσω αντικατάσταση" προκύπτει $a = b = c = 0$, δηλαδή τα διανύσματα $\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3$ είναι γραμμικά ανεξάρτητα, κατά συνέπεια αποτελούν μια βάση για τον $\mathbb{R}^3$.

Παρατήρηση 4.5. Από το Παράδειγμα 4.10 και την παρατήρηση στην εισαγωγή της παραγράφου συμπεραίνουμε ότι η βάση ενός διανυσματικού χώρου δεν είναι μοναδική. Για παράδειγμα οι

$$\mathcal{B}_1 = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}, \qquad \mathcal{B}_2 = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$$

είναι βάσεις του $\mathbb{R}^3$. Η ιδιαίτερη $\mathcal{B}_1$ λέγεται **συνήθης** βάση. Γενικά η $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n\}$ με $\boldsymbol{e}_k$ όπως στην (1.4) λέγεται συνήθης βάση για τον $\mathbb{R}^n$.

Το γεγονός ότι οι βάσεις $\mathcal{B}_1$ και $\mathcal{B}_2$ του $\mathbb{R}^3$ περιέχουν τον ίδιο αριθμό διανυσμάτων είναι καθόλου τυχαίο.

Θεώρημα 4.5. *Εάν $\{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m\}$ και $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ είναι βάσεις ενός διανυσματικού χώρου X, τότε $m = n$.*

Απόδειξη. Δίνουμε την απόδειξη με την εις άτοπο απαγωγή. Ας υποθέσουμε ότι $n > m$. Θα αποδείξουμε ότι τα $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ είναι γραμμικά εξαρτημένα. Αφού τα $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$ αποτελούν βάση υπάρχουν σταθερές ώστε

$$\begin{aligned}
\boldsymbol{v}_1 &= a_{11}\boldsymbol{b}_1 + a_{12}\boldsymbol{b}_2 + \cdots + a_{1m}\boldsymbol{b}_m \\
\boldsymbol{v}_2 &= a_{21}\boldsymbol{b}_1 + a_{22}\boldsymbol{b}_2 + \cdots + a_{2m}\boldsymbol{b}_m \\
&\vdots \\
\boldsymbol{v}_n &= a_{n1}\boldsymbol{b}_1 + a_{n2}\boldsymbol{b}_2 + \cdots + a_{nm}\boldsymbol{b}_m.
\end{aligned}$$

Αν $c_1, c_2, \ldots, c_n$ είναι σταθερές ώστε

$$c_1\boldsymbol{v}_1 + c_2\boldsymbol{v}_2 + \cdots + c_n\boldsymbol{v}_n = \boldsymbol{0}, \tag{4.12}$$

τότε από τις σχέσεις που εκφράζουν τα $\boldsymbol{v}$ μέσω των $\boldsymbol{b}$ παίρνουμε

$$\begin{aligned}
&(c_1a_{11} + c_2a_{21} + \cdots + c_na_{n1})\boldsymbol{b}_1 + (c_1a_{12} + c_2a_{22} + \cdots + c_na_{n2})\boldsymbol{b}_2 + \\
&\qquad \cdots + (c_1a_{1m} + c_2a_{2m} + \cdots + c_na_{nm})\boldsymbol{b}_m = \boldsymbol{0}.
\end{aligned}$$

Επειδή τα $\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots, \boldsymbol{b}_m$ είναι γραμμικά ανεξάρτητα έπεται ότι κάθε συντελεστής στην τελευταία εξίσωση είναι ίσος με μηδέν, δηλαδή

$$\left.\begin{aligned}
c_1a_{11} + c_2a_{21} + \cdots + c_na_{n1} &= 0 \\
c_1a_{12} + c_2a_{22} + \cdots + c_na_{n2} &= 0 \\
&\vdots \\
c_1a_{1m} + c_2a_{2m} + \cdots + c_na_{nm} &= 0
\end{aligned}\right\} \Leftrightarrow c_1\begin{pmatrix} a_{11} \\ a_{12} \\ \vdots \\ a_{1m} \end{pmatrix} + c_2\begin{pmatrix} a_{21} \\ a_{22} \\ \vdots \\ a_{2m} \end{pmatrix} + \cdots + c_n\begin{pmatrix} a_{n1} \\ a_{n2} \\ \vdots \\ a_{nm} \end{pmatrix} = \boldsymbol{0}.$$

Η τελευταιά εξίσωση εκφράζει ένα γραμμικό συνδυασμό n διανυσμάτων του $\mathbb{R}^m$ με $n > m$. Από το Πόρισμα 4.1 έπεται ότι τα n αυτά διανύσματα είναι γραμμικά εξαρτημένα, συνεπώς κάποιες από τις σταθερές $c_1, c_2, \ldots, c_n$ είναι διάφορες του μηδενός, γεγονός το οποίο αποδεικνύει μέσω της (4.12), ότι τα $\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n$ είναι γραμμικά εξαρτημένα. Αυτό όμως είναι άτοπο αφού τα διανύσματα αυτά αποτελούν βάση. Καταλήξαμε σε άτοπο γιατί υποθέσαμε ότι $n > m$, άρα θα πρέπει να είναι $n \leq m$. Υποθέτοντας στη συνέχεια ότι $n < m$ και εναλλάσσοντας τους ρόλους των $\boldsymbol{b}$ και $\boldsymbol{v}$ στην απόδειξη καταλήγουμε πάλι σε άτοπο. Επομένως συμπεραίνουμε ότι $n = m$. □

Ορισμός 4.7. Το πλήθος των στοιχείων μιας βάσης $\mathcal{B}$ ενός διανυσματικού χώρου X λέγεται **διάσταση** (dimension) του χώρου. Αν το $\mathcal{B}$ είναι πεπερασμένο σύνολο θα λέμε ότι ο X έχει πεπερασμένη διάσταση, διαφορετικά ο X θα λέγεται απειροδιάστατος χώρος. Αν η διάσταση του X είναι n, γράφουμε $\dim X = n$. Αν ο X είναι πεπερασμένης διάστασης γράφουμε $\dim X < \infty$.

Παράδειγμα 4.11. Τα διανύσματα

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

αποτελούν μια βάση για τον $\mathbb{R}^3$, κατά συνέπεια $\dim \mathbb{R}^3 = 3$.

Παράδειγμα 4.12. Τα μητρώα

$$E_1 = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \quad E_2 = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad E_3 = \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \quad E_4 = \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}$$

αποτελούν μια βάση για το χώρο των 2×2 μητρώων, κατά συνέπεια $\dim \mathbb{M}^{2,2} = 2^2$. Γενικότερα $\dim \mathbb{M}^{n,n} = n^2$, και $\dim \mathbb{M}^{n,m} = nm$ (γιατί;).

Παράδειγμα 4.13. Δείξτε ότι το σύνολο των μητρώων

$$\mathcal{D} = \left\{ \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix} : a \in \mathbb{R} \right\}$$

με την πράξη της πρόσθεσης μητρώων και πολλαπλασιασμό με σταθερά αποτελούν διανυσματικό χώρο. Βρείτε επίσης μια βάση για τον χώρο αυτό.

Το $\mathcal{D}$ είναι υποσύνολο του $\mathbb{R}^{3\times 3}$, αλλά και του χώρου των 3×3 τριγωνικών μητρώων. Έτσι αν a, b, λ, μ είναι πραγματικοί αριθμοί έχουμε

$$\lambda \begin{pmatrix} a & 0 & 0 \\ 0 & a & 0 \\ 0 & 0 & a \end{pmatrix} + \mu \begin{pmatrix} b & 0 & 0 \\ 0 & b & 0 \\ 0 & 0 & b \end{pmatrix} = \begin{pmatrix} \lambda a + \mu b & 0 & 0 \\ 0 & \lambda a + \mu b & 0 \\ 0 & 0 & \lambda a + \mu b \end{pmatrix} \in \mathcal{D}$$

κατά συνέπεια το $\mathcal{D}$ είναι υπόχωρος του $\mathbb{R}^{3\times 3}$, άρα είναι διανυσματικός χώρος.

Παρατηρούμε ότι $\mathcal{D} = \operatorname{span}\{I_3\}$ όπου I_3 είναι το 3×3 ταυτοτικό μητρώο. Αυτός είναι ένας άλλος τρόπος απόδειξης ότι το $\mathcal{D}$ είναι διανυσματικός χώρος, αφού κάθε διάνυγμα διανυσμάτων είναι διανυσματικός υπόχωρος. Επιπλέον, από αυτό το γεγονός, έπεται ότι το $\mathcal{B} = \{I_3\}$ είναι μια βάση για τον $\mathcal{D}$, έτσι $\dim \mathcal{D} = 1$.

Παράδειγμα 4.14. Σε κάθε διανυσματικό χώρο X ο υπόχωρος $\{\mathbf{0}\}$ δεν περιέχει γραμμικά ανεξάρτητα διανύσματα, κατά συνέπεια $\dim\{\mathbf{0}\} = 0$.

Κατά την απόδειξη του Θεωρήματος 4.5 δείξαμε ότι αν τα διανύσματα $\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_m$ αποτελούν βάση για ένα διανυσματικό χώρο X, τότε n το πλήθος διανύσματα του X, με $n > m$, είναι γραμμικά εξαρτημένα. Κατά συνέπεια έχουμε αποδείξει το αποτέλεσμα που ακολουθεί και είναι γενίκευση του Πορίσματος 4.1

Θεώρημα 4.6. *Εάν $\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_n$, είναι διανύσματα του m-διάστατου χώρου X, δηλαδή $\dim X = m$, και $n > m$, τότε τα διανύσματα είναι γραμμικά εξαρτημμένα.*

Παρατήρηση 4.6. Έστω A ένα $n \times n$ μητρώο. Επειδή η διάσταση του $\mathbb{M}^{n,n}$ είναι n^2 μια άμεση συνέπεια του Θεωρήματος 4.6 είναι ότι για $N \geq n^2$ τα μητρώα

$$I,\ A,\ A^2,\ \dots,\ A^N$$

είναι γραμμικά εξαρτημένα, κατά συνέπεια υπάρχουν σταθερές $a_0, a_1, a_2 \dots, a_N$, όχι όλες ίσες με μηδέν, ώστε

$$a_0 I + a_1 A + a_2 A^2 + \cdots + a_N A^N = O, \tag{4.13}$$

ισοδύναμα υπάρχει μη μηδενικό πολυώνυμο $p(t) = a_0 + a_1 t + a_2 t^2 + \cdots + a_N t^N$ ώστε $p(A) = O$. Σε επόμενο κεφάλαιο θα υπολογίσουμε για κάθε $n \times n$ μητρώο το αντίστοιχο πολυώνυμο p_A βαθμού ακριβώς n το οποίο ικανοποιεί την $p_A(A) = O$ και σχετίζεται με χαρακτηριστικές ιδιότητες του μητρώου.

Θεώρημα 4.7. *Εάν $\dim X < \infty$ και ο W είναι υπόχωρος του X, τότε*

(1) *Ο υπόχωρος W έχει πεπερασμένη διάσταση.*

(2) $\dim W \leq \dim X$.

(3) *Εάν $\dim W = \dim X$, τότε $W = X$.*

Απόδειξη. Έστω ότι $\dim X = n$ και έστω $\mathcal{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \dots, \boldsymbol{b}_n\}$ μια βάση του X.

(1) Ας υποθέσουμε ότι ο W έχει άπειρη διάσταση, τότε για κάθε πεπερασμένο σύνολο γραμμικά ανεξαρτήτων διανυσμάτων του X, έστω $B = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_m\}$, θα υπάρχει $w \in W$ ώστε $w \notin \operatorname{span} B$. Για $B = \mathcal{B}$ θα υπάρχει λοιπόν w στο W και $w \notin \operatorname{span} \mathcal{B}$. Αυτό όμως είναι άτοπο αφού $w \in X$ και $\mathcal{B}$ είναι βάση του X. Καταλήξαμε σε άτοπο γιατί υποθέσαμε ότι ο W είναι απειροδιάστατος, κατά συνέπεια ο W έχει πεπερασμένη διάσταση.

(2) Από το (1) θα είναι $\dim W = m$ για κάποιο θετικό ακέραιο m και από το Θεώρημα 4.6 έπεται ότι $m \leq n$.

(3) Είναι $W \subseteq X$ και ας υποθέσουμε ότι υπάρχει $\boldsymbol{x} \in X$ ώστε $\boldsymbol{x} \notin W$. Τότε αφενός $\boldsymbol{x} \neq \mathbf{0}$ (γιατί;) και αφετέρου αν $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$ είναι μια βάση για τον W, τότε

$$\boldsymbol{x} \neq c_1 \boldsymbol{w}_1 + c_2 \boldsymbol{w}_2 + \cdots + c_n \boldsymbol{w}_n$$

για κάθε n-άδα πραγματικών αριθμών $c_1, c_2, \dots, c_n$. Συμπεραίνουμε λοιπόν ότι τα διανύσματα $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n, \boldsymbol{x}\}$ είναι γραμμικά ανεξάρτητα[1]. Το γεγονός αυτό έρχεται σε αντίθεση με το αποτέλεσμα είτε του Θεωρήματος 4.6, είτε του (2) αφού

$$\dim \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n, \boldsymbol{x}\} \leq \dim X \Leftrightarrow n + 1 \leq n.$$

Καταλήξαμε σε άτοπο γιατί υποθέσαμε ότι $X \setminus W \neq \emptyset$, κατά συνέπεια $W = X$. □

Παράδειγμα 4.15. Τα διανύσματα

$$\boldsymbol{a}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{a}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix},$$

είναι γραμμικά ανεξάρτητα. Πράγματι

$$c_1\boldsymbol{a}_1 + c_2\boldsymbol{a}_2 = \boldsymbol{0} \Leftrightarrow \begin{pmatrix} c_1 \\ c_1 \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ c_2 \\ c_2 \end{pmatrix} = \begin{pmatrix} c_1 \\ c_1 + c_2 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

απ᾽ όπου βλέπουμε ότι $c_1 = c_2 = 0$. Παράγουν κατά συνέπεια έναν διδιάστατο υπόχωρο του $\mathbb{R}^3$, τον

$$W = \operatorname{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2\} = \left\{ \begin{pmatrix} x \\ y \\ z \end{pmatrix} : y = x + z, \quad \text{με} \quad x, z \in \mathbb{R} \right\}$$

για τον οποίο αποτελούν μια βάση. Ισχυριζόμαστε ότι εμπλουτίζοντας τη βάση $\{\boldsymbol{a}_1, \boldsymbol{a}_2\}$ του W με ένα διάνυσμα $\boldsymbol{a}_3$ του $\mathbb{R}^3$ ώστε τα $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ να είναι γραμμικά ανεξάρτητα, τότε παίρνουμε μια βάση του $\mathbb{R}^3$. Επιλέγοντας ένα διάνυσμα που "καταστρέφει" τη δομή των στοιχείων του W για παράδειγμα

$$\boldsymbol{a}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}$$

βλέπουμε ότι τα $\boldsymbol{a}_1, \boldsymbol{a}_2, \boldsymbol{a}_3$ είναι όντως γραμμικά ανεξάρτητα (γιατί;) και επιπλέον παράγουν τον χώρο $\mathbb{R}^3$, αφού

$$\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x \\ x \\ 0 \end{pmatrix} + \begin{pmatrix} 0 \\ z \\ z \end{pmatrix} + \begin{pmatrix} 0 \\ y - x - z \\ 0 \end{pmatrix} = x\boldsymbol{a}_1 + z\boldsymbol{a}_2 + (y - x - z)\boldsymbol{a}_3.$$

Σε σχέση με το Παράδειγμα 4.15 σημειώνουμε ότι οι συνιστώσες των διανυσμάτων του W ικανοποιούν την εξίσωση $x - y + z = 0$ η οποία είναι η εξίσωση του επιπέδου W, βλέπε Παράρτημα Α: Η εξίσωση του επιπέδου. Γενικά οι υπόχωροι διάστασης δύο στο $\mathbb{R}^3$ είναι

[1]Πράγματι αν

$$\lambda_1\boldsymbol{w}_1 + \lambda_2\boldsymbol{w}_2 + \cdots + \lambda_n\boldsymbol{w}_n + \lambda\boldsymbol{x} = \boldsymbol{0},$$

τότε $\lambda = 0$, γιατί σε διαφορετική περίπτωση το $\boldsymbol{x}$ θα εκφραζόταν ως γραμμικός συνδυασμός των $\boldsymbol{w}_k$, κατά συνέπεια $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 0$, αφού τα $\boldsymbol{w}_k$ είναι γραμμικά ανεξάρτητα.

επίπεδα. Στο ερώτημα σχετικά με το ποια είναι η εξίσωση του επιπέδου που παράγεται από τα γραμμικά ανεξάρτητα διανύσματα

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \qquad \boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix}$$

σκεφτόμαστε ότι αναζητάμε σταθερές a, b, c, όχι όλες ίσες με μηδέν, ώστε για το τυχαίο διάνυσμα $\lambda\boldsymbol{u} + \mu\boldsymbol{v}$ στο $\operatorname{span}\{\boldsymbol{u}, \boldsymbol{v}\}$ θα πρέπει να ισχύει

$$a(\lambda u_1 + \mu v_1) + b(\lambda u_1 + \mu v_2) + c(\lambda u_3 + \mu v_3) = 0$$

ισοδύναμα

$$\lambda(u_1 a + u_2 b + u_3 c) + \mu(v_1 a + v_2 b + v_3 c) = 0$$

για κάθε $\lambda, \mu \in \mathbb{R}$, κατά συνέπεια θα πρέπει τα a, b, c να είναι μη μηδενική λύση του συστήματος

$$\begin{aligned} u_1 a + u_2 b + u_3 c &= 0 \\ v_1 a + v_2 b + v_3 c &= 0. \end{aligned}$$

Το Θεώρημα 4.4 εξασφαλίζει την ύπαρξη τέτοιων λύσεων, έτσι αν (a, b, c) είναι μια τέτοια, τότε η ζητούμενη εξίσωση είναι η $ax + y + cz = 0$.

► **Άσκηση 4.4.** Δείξτε ότι το σύνολο των διανυσμάτων του $\mathbb{P}_n$

$$\mathcal{B} = \{1, x, x^2, \ldots, x^n\} \tag{4.14}$$

είναι μια βάση για τον $\mathbb{P}_n$. Συμπεράνατε ότι $\dim \mathbb{P}_n = n + 1$.

Στη διατύπωση της Άσκησης 4.4 έγινε μια επιτρεπτή παρατυπία. Στην (4.14) γράψαμε τις τιμές των πολυωνύμων και όχι τα πολυώνυμα. Θα έπρεπε για παράδειγμα να γράφαμε

$$\mathcal{B} = \{p_0, p_1, p_2, \ldots, p_n\} \qquad \text{με} \qquad p_0(x) = 1 \quad \text{και} \quad p_k(x) = x^k, \quad k = 1, 2, \ldots, n.$$

Για πρακτικούς λόγους και όταν δεν αναμένεται δημιουργία σύγχυσης επιλέγουμε, σε αυτή και σε ανάλογες περιπτώσεις, την οικονομικότερη γραφή (4.14).

Παρατήρηση 4.7. Αξίζει στο σημείο αυτό να σημειώσουμε ότι αν p είναι ένα πολυώνυμο στον χώρο $\mathbb{P}_n$ και $p(x) = a_0 + a_1 x + a_2 x^2 \cdots + a_n x^n$, τότε η απεικόνιση $L : \mathbb{P}_n \to \mathbb{R}^{n+1}$ που ορίζεται με τη σχέση

$$L(p) = \begin{pmatrix} a_0 \\ a_1 \\ \vdots \\ a_n \end{pmatrix} \in \mathbb{R}^{n+1}$$

είναι ένα-προς-ένα και επί και επιπλέον για $p, q \in \mathbb{P}_n$ και $\lambda \in \mathbb{R}$ ικανοποιεί τις σχέσεις

$$L(p + q) = L(p) + L(q), \qquad \text{και} \qquad L(\lambda p) = \lambda L(p). \tag{4.15}$$

Μια ένα-προς-ένα και επί απεικόνιση μεταξύ διανυσματικών χώρων η οποία ικανοποιεί τις σχέσεις στην (4.15) λέγεται **ισομορφισμός** . Η ύπαρξη ενός ισομορφισμού μεταξύ δύο διανυσματικών χώρων δηλώνει ότι οι δύο χώροι, πρακτικά, ταυτίζονται. Στη συγκεκριμένη περίπτωση δεχόμαστε ότι η k-συνιστώσα ενός διανύσματος στο $\mathbb{R}^{n+1}$ είναι ο συντελεστής του x^{k-1},

$k = 1, 2, \dots n + 1$. Έτσι, για παράδειγμα, έχουμε την αντιστοιχία

$$\mathbb{R}^3 \ni \begin{pmatrix} 2 \\ -3 \\ 1 \end{pmatrix} \quad \longleftrightarrow \quad 2 - 3x + x^2 \in \mathbb{P}_2.$$

Επίσης το άθροισμα διανυσμάτων ή το γινόμενο διανύσματος με σταθερά αντιστοιχεί και δίνει το ίδιο αποτέλεσμα με το άθροισμα των αντιστοίχων πολυωνύμων ή το γινόμενο του αντίστοιχου πολυωνύμου με την σταθερά.

Ορισμός 4.8. Αν $(X, +, \cdot)$ και $(Y, +, \cdot)$ είναι πραγματικοί ή μιγαδικοί διανυσματικοί χώροι μια απεικόνιση $F : X \to Y$ η οποία ικανοποιεί τις σχέσεις

$$F(x + y) = F(x) + F(y), \qquad \text{και} \qquad F(\lambda x) = \lambda F(x), \tag{4.16}$$

για κάθε x και y στο X και για κάθε πραγματική ή μιγαδική σταθερά λ, λέγεται **γραμμική απεικόνιση** (linear map) ή **γραμμικός μετασχηματισμός** (linear transformation). Πολλές φορές για μια τέτοια F, αν δεν υπάρχει σύγχυση, γράφουμε Fx αντί του $F(x)$.

Παρατήρηση 4.8. Ας αναρωτηθούμε ποιά είναι η διάσταση του χώρου των συνεχών συναρτήσεων, ας πούμε στο διάστημα $[0, 1]$, δηλαδή του $C[0, 1]$. Ας υποθέσουμε ότι $\dim C[0, 1] = n$ για κάποιο $n \in \mathbb{N}$. Τότε ενώ ο $\mathbb{P}_n[0, 1]$ είναι υπόχωρος του $C[0, 1]$ θα είναι

$$\dim \mathbb{P}_n[0, 1] = n + 1 > \dim C[0, 1]$$

γεγονός το οποίο, λόγω του Θεωρήματος 4.7, δεν μπορεί να συμβαίνει. Κατά συνέπεια η διάσταση του $C[0, 1]$ είναι άπειρη.

Σημείωση 4.2. Ένα σημαντικό αποτέλεσμα της Μαθηματικής Ανάλυσης είναι το Θεώρημα του Weierstrass το οποίο λέει ότι αν $f \in C[0, 1]$, τότε για κάθε $\epsilon > 0$ υπάρχει πολυώνυμο p ώστε

$$\sup_{0 \le x \le 1} |f(x) - p(x)| < \epsilon$$

δηλαδή κάθε συνεχής συνάρτηση μπορεί να προσεγγισθεί με επιθυμητή ακρίβεια με κάποιο πολυώνυμο ομοιόμορφα στο $[0, 1]$. Το αποτέλεσμα αυτό μπορούμε να το γράψουμε στη μορφή

$$f(x) = p(x) + s(x), \qquad \text{με} \qquad |s(x)| < \epsilon \quad \forall x \in [0, 1],$$

με $s(x) = f(x) - p(x)$. Έτσι προκύπτει το αποτέλεσμα ότι κάθε συνεχής συνάρτηση στο $[0, 1]$ είναι ίση με ένα πολυώνυμο συν κάποιο μικρό "σφάλμα". Σε επόμενο κεφάλαιο θα μας απασχολήσουν τα ανάλογα σχετικά προβλήματα:

1ο Πώς προσεγγίζουμε ένα διάνυσμα ενός διανυσματικού χώρου X με διανύσματα από κάποιο υπόχωρο W του X;

2ο Ποια είναι η καλύτερη, ως προς κάποιο καθορισμένο τρόπο μέτρησης, προσέγγιση ενός διανύσματος του X με διανύσματα ενός δοσμένου υπόχωρου W του X, ισοδύναμα ποια προσέγγιση δίνει το ελάχιστο δυνατό σφάλμα;

Στο Παράδειγμα 4.15 υπαινιχθήκαμε ότι εμπλουτίζοντας μια βάση ενός υπόχωρου διανυσματικού χώρου με επιπλέον διανύσματα μπορούμε να πάρουμε μια βάση για ολόκληρο τον χώρο.

Θεώρημα 4.8 (Επέκταση βάσης υποχώρου). *Εάν $S = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_k\}$ είναι ένα σύνολο γραμμικά ανεξαρτήτων διανυσμάτων στο χώρο X πεπερασμένης διάστασης, και $\operatorname{span} S \neq X$, τότε υπάρχουν διανύσματα $\boldsymbol{w}_{k+1}, \boldsymbol{w}_{k+2}, \dots, \boldsymbol{w}_n$ με $n = \dim X$, ώστε το σύνολο $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_k, \boldsymbol{w}_{k+1}, \boldsymbol{w}_{k+2}, \dots, \boldsymbol{w}_n\}$ να είναι μια βάση για τον X.*

Απόδειξη. Εάν $\boldsymbol{w} \notin \operatorname{span} S$, τότε τα διανύσματα $\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_k, \boldsymbol{w}$ είναι γραμμικά ανεξάρτητα. Πράγματι έστω

$$c_1\boldsymbol{w}_1 + c_2\boldsymbol{w}_2 + \cdots + c_k\boldsymbol{w}_k + c\boldsymbol{w} = \boldsymbol{0}. \tag{4.17}$$

Εάν $c \neq 0$, τότε θα ήταν

$$\boldsymbol{w} = -\frac{c_1}{c}\boldsymbol{w}_1 - \frac{c_2}{c}\boldsymbol{w}_2 - \cdots - \frac{c_k}{c}\boldsymbol{w}_k$$

το οποίο είναι άτοπο αφού $\boldsymbol{w} \notin \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_k\}$, επομένως $c = 0$ και από την (4.17) έπεται ότι

$$c_1\boldsymbol{w}_1 + c_2\boldsymbol{w}_2 + \cdots + c_k\boldsymbol{w}_k = \boldsymbol{0}. \tag{4.18}$$

Τότε όμως από τον ορισμό του S έπεται ότι $c_1 = c_2 = \cdots = c_k = 0$, άρα $c_1 = c_2 = \cdots = c_k = c = 0$, κατά συνέπεια τα $\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_k, \boldsymbol{w}$ είναι γραμμικά ανεξάρτητα. Θέτουμε $\boldsymbol{w}_{k+1} = \boldsymbol{w}$. Εάν $k+1 < n$, επαναλαμβάνουμε τη διαδικασία. Έτσι σε $n-k$ βήματα βρίσκουμε ένα σύνολο

$$\mathcal{S}' = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_k, \boldsymbol{w}_{k+1}, \boldsymbol{w}_{k+2}, \dots, \boldsymbol{w}_n\}$$

από γραμμικά ανεξάρτητα διανύσματα στον X. Τότε από το Θεώρημα 4.7 έπεται ότι το $\mathcal{S}'$ είναι μια βάση για το X. □

▶ **Άσκηση 4.5.** Δίνεται το επίπεδο με εξίσωση $x_1 - 2x_2 + x_3 = 0$.

(α΄) Να βρεθεί μια βάση για το επίπεδο.

(β΄) Επεκτείνετε τη βάση στο (α΄) σε μια βάση για το $\mathbb{R}^3$.

Παρατήρηση 4.9 (Το διάνυσμα συντεταγμένων). Εάν $X(\mathbb{F})$ είναι διανυσματικός χώρος διάστασης n και $\mathcal{B} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{n}_n\}$ είναι μια βάση για τον X, τότε σε κάθε διάνυσμα $\boldsymbol{x} \in X$ αντιστοιχούν σταθερές $c_1, c_2, \dots, c_n \in \mathbb{F}$ ώστε

$$\boldsymbol{x} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdot + c_n\boldsymbol{u}_n.$$

Το διάνυσμα

$$(\boldsymbol{x})_{\mathcal{B}} := \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{F}^n$$

θα το λέμε **διάνυσμα συντεταγμένων** του $\boldsymbol{x}$ (coordinate vector) ως προς τη βάση $\mathcal{B}$. Αφήνεται σαν άσκηση να δειχθεί, ότι η απεικόνιση $\Phi : X \to \mathbb{F}^n$ με

$$\Phi(x) = (\boldsymbol{x})_{\mathcal{B}}$$

είναι γραμμική, δηλαδή για όλα τα $\boldsymbol{x}, \boldsymbol{y} \in X$ και $\lambda, \mu \in \mathbb{F}$ ισχύει

$$\Phi(\lambda\boldsymbol{x} + \mu\boldsymbol{y}) = \lambda\Phi(\boldsymbol{x}) + \mu\Phi(\boldsymbol{y}) \Leftrightarrow (\lambda\boldsymbol{x} + \mu\boldsymbol{y})_{\mathcal{B}} = \lambda(\boldsymbol{x})_{\mathcal{B}} + \mu(\boldsymbol{y})_{\mathcal{B}},$$

ένα-προς-ένα και επί, κατά συνέπεια είναι ένας ισομορφισμός ινδεξισομορφισμος!ισομορφισμός, βλέπε Παρατήρηση 4.7. Έτσι κάθε πραγματικός διανυσματικός χώρος διάστασης n είναι ισόμορφος με τον $\mathbb{R}^n$. Κατά συνέπεια κάθε διανυσματικός χώρος πεπερασμένης διάστασης, έστω n, "συμπεριφέρεται" όπως ο $\mathbb{R}^n$ ή ο $\mathbb{C}^n$.

4.5 Ευθύ άθροισμα υποχώρων

Ένα διάνυσμα $\boldsymbol{u}$ στο $\mathbb{R}^3$ γράφεται ως

$$\boldsymbol{u} = c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2 + c_3\boldsymbol{e}_3$$

όπου $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ είναι η κανονική βάση του χώρου. Θέτοντας $\boldsymbol{v} = c_1\boldsymbol{e}_1 + c_2\boldsymbol{e}_2$ και $\boldsymbol{z} = c_3\boldsymbol{e}_3$ βλέπουμε ότι τα $\boldsymbol{v}, \boldsymbol{z}$ είναι, αντίστοιχα, διανύσματα των υποχώρων

$$V = \left\{ \begin{pmatrix} x \\ y \\ 0 \end{pmatrix} : x, y \in \mathbb{R} \right\}, \qquad Z = \left\{ \begin{pmatrix} 0 \\ 0 \\ z \end{pmatrix} : z \in \mathbb{R} \right\}$$

του $\mathbb{R}^3$, και ότι $\boldsymbol{u} = \boldsymbol{v} + \boldsymbol{z}$. Υπό αυτή την έννοια μπορούμε να γράψουμε $\mathbb{R}^3 = V + Z$.

Εάν X είναι ένας διανυσματικός χώρος και V, W είναι υπόχωροι του X ορίζουμε το **άθροισμα** των υποχώρων V και W

$$V + W = \{\boldsymbol{v} + \boldsymbol{w} : \boldsymbol{v} \in V \text{ και } \boldsymbol{w} \in W\}.$$

Παρατηρούμε ότι $V + W \subseteq X$. Επιπλέον, παρατηρούμε ότι το $V + W$ είναι υπόχωρος του X. Πράγματι αν $\boldsymbol{a}, \boldsymbol{b} \in V + W$, τότε $\boldsymbol{a} = \boldsymbol{v} + \boldsymbol{w}$ και $\boldsymbol{b} = \boldsymbol{v}' + \boldsymbol{w}'$ για κάποια $\boldsymbol{v}, \boldsymbol{v}' \in V$ και $\boldsymbol{w}, \boldsymbol{w}' \in W$, και για $\lambda, \mu \in \mathbb{R}$, έχουμε

$$\begin{aligned} \lambda\boldsymbol{a} + \mu\boldsymbol{b} &= \lambda\boldsymbol{v} + \lambda\boldsymbol{w} + \mu\boldsymbol{v}' + \mu\boldsymbol{w}' \\ &= (\lambda\boldsymbol{v} + \mu\boldsymbol{v}') + (\lambda\boldsymbol{w} + \mu\boldsymbol{w}') \in V + W. \end{aligned}$$

Έστω ότι $X = V + W$, όπου V και W είναι υπόχωροι του X. Αν $\boldsymbol{x} \in X$ και $\boldsymbol{x} = \boldsymbol{v} + \boldsymbol{w}$, με $\boldsymbol{v} \in V$ και $\boldsymbol{w} \in W$, τότε για $\boldsymbol{u} \in V \cap W$ ισχύει επίσης ότι $\boldsymbol{x} = \boldsymbol{v} + \boldsymbol{u} + \boldsymbol{w} - \boldsymbol{u}$, με $\boldsymbol{v} + \boldsymbol{u} \in V$ και $\boldsymbol{w} - \boldsymbol{u} \in W$, δηλαδή η αναπαράσταση ενός διανύσματος του X ως άθροισμα ενός διανύσματος του V και ενός του W δεν είναι μοναδική. Αν όμως $V \cap W = \{\boldsymbol{0}\}$, τότε μια τέτοια αναπαράσταση είναι μοναδική. Πράγματι αν $\boldsymbol{x} = \boldsymbol{v} + \boldsymbol{w} = \boldsymbol{v}' + \boldsymbol{w}'$ με $\boldsymbol{v}, \boldsymbol{v}' \in V$ και $\boldsymbol{w}, \boldsymbol{w}' \in W$, τότε

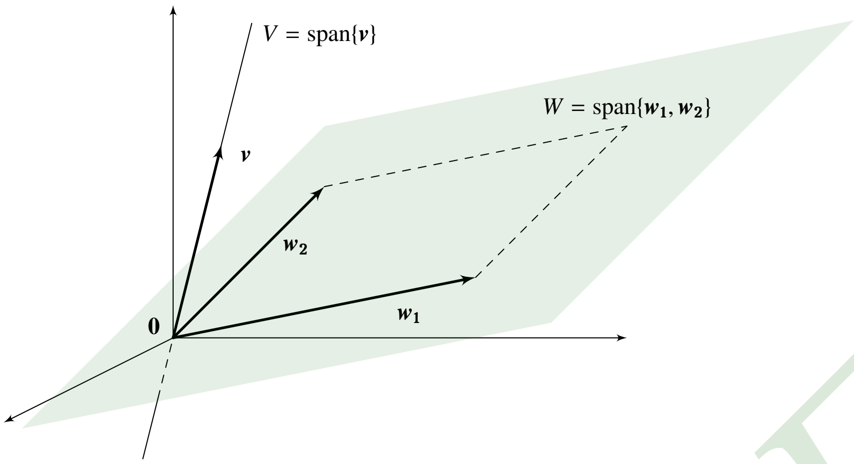
$$\begin{aligned} \boldsymbol{0} &= (\boldsymbol{v} + \boldsymbol{w}) - (\boldsymbol{v}' + \boldsymbol{w}') \\ &= (\boldsymbol{v} - \boldsymbol{v}') + (\boldsymbol{w} - \boldsymbol{w}') \Rightarrow \boldsymbol{v} - \boldsymbol{v}' = \boldsymbol{w}' - \boldsymbol{w} \end{aligned}$$

οπότε $\boldsymbol{v} - \boldsymbol{v}' \in V \cap W$ και $\boldsymbol{w}' - \boldsymbol{w} \in V \cap W$, οπότε από την υπόθεση έπεται ότι $\boldsymbol{v} - \boldsymbol{v}' = \boldsymbol{0}$ και $\boldsymbol{w}' - \boldsymbol{w} = \boldsymbol{0}$, δηλαδή $\boldsymbol{v} = \boldsymbol{v}'$ και $\boldsymbol{w} = \boldsymbol{w}'$.

Ορισμός 4.9. Έστω ότι V και W είναι υπόχωροι του διανυσματικού χώρου X. Θα λέμε ότι ο X είναι **ευθύ άθροισμα** (direct sum) των V και W εάν $X = V + W$ και $V \cap W = \{\boldsymbol{0}\}$. Εάν ο X είναι ευθύ άθροισμα των V και W γράφουμε $X = V \oplus W$.

Παρατήρηση 4.10. Αν $\boldsymbol{w}_1$, και $\boldsymbol{w}_2$ είναι γραμμικά ανεξάρτητα διανύσματα στο $\mathbb{R}^3$ και $\boldsymbol{v} \in \mathbb{R}^3$ δεν περιέχεται στο επίπεδο που ορίζουν τα $\boldsymbol{w}_1$, $\boldsymbol{w}_2$, ισοδύναμα $\boldsymbol{v} \notin \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$, τότε το $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{v}\}$ είναι μια βάση για το $\mathbb{R}^3$, ισοδύναμα (γιατί;) $\mathbb{R}^3 = \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2\} \oplus \operatorname{span}\{\boldsymbol{v}\}$. Βλέπε Παράδειγμα 4.15. Παρατηρείστε ότι $\mathbb{R}^3 = \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2\} + \operatorname{span}\{\boldsymbol{v}, \boldsymbol{w}_1\}$, αλλά το άθροισμα δεν είναι ευθύ αφού η τομή των δύο υποχώρων είναι η ευθεία που ορίζει το διάνυσμα $\boldsymbol{w}_1$.

Το Θεώρημα 4.8 μπορεί να διατυπωθεί και ως



Σχήμα 4.1: $\mathbb{R}^3 = W \oplus V$

Θεώρημα 4.9. *Εάν ο X είναι ένας διανυσματικός χώρος πεπερασμένης διάστασης και V είναι ένας υπόχωρος του X, τότε υπάρχει υπόχωρος W του X ώστε $X = V \oplus W$.*

4.6 Ασκήσεις

1. Θεωρούμε το σύνολο όλων των 2×2 μιγαδικών μητρώων

$$\mathbb{H} = \left\{ \begin{pmatrix} \zeta & \xi \\ -\bar{\xi} & \bar{\zeta} \end{pmatrix} : \zeta, \xi \in \mathbb{C} \right\},$$

όπου $\bar{s}$ είναι ο συζυγής του s. Δείξτε ότι το $\mathbb{H}$ με τις συνήθεις πράξεις της πρόσθεσης και του πολλαπλασιασμού με πραγματική σταθερά είναι πραγματικός διανυσματικός χώρος. Είναι το $\mathbb{H}$ μιγαδικός χώρος;

2. Έστω A ένα 3×3 μητρώο. Δείξτε ότι το σύνολο των μητρώων που μετατίθενται με το A, δηλαδή $W = \{X : AX = XA\}$ με τις συνήθεις πράξεις είναι διανυσματικός χώρος.

3. Εξηγείστε γιατί τα διανύσματα

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \boldsymbol{u}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}$$

δεν αποτελούν βάση για τον $\mathbb{R}^3$. Στη συνέχεια βρείτε το $\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$.

4. Έστω $\mathbb{P}_3^*$ το σύνολο των πολυωνύμων του $\mathbb{P}_3$, του χώρου των πολυωνύμων βαθμού το πολύ τρία, με μηδενικό σταθερό όρο, $\mathbb{P}_3^* = \{p \in \mathbb{P}_3 : p(0) = 0\}$.

   (α΄) Δείξτε ότι το $\mathbb{P}_3^*$ είναι υπόχωρος του $\mathbb{P}_3$.

   (β΄) Να βρεθεί μια βάση για τον $\mathbb{P}_3^*$.

5. Εάν τα διανύσματα $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ κάποιου διανυσματικού χώρου X είναι γραμμικά ανεξάρτητα, εξετάστε ως προς την ανεξαρτησία τα διανύσματα

   (α΄) $\boldsymbol{v}_1 = \boldsymbol{u}_2 + \boldsymbol{u}_3$, $\boldsymbol{v}_2 = \boldsymbol{u}_1 + \boldsymbol{u}_3$, $\boldsymbol{v}_3 = \boldsymbol{u}_1 + \boldsymbol{u}_2$.

(β΄) $\boldsymbol{w}_1 = \boldsymbol{u}_2 - \boldsymbol{u}_3$, $\boldsymbol{w}_2 = \boldsymbol{u}_3 - \boldsymbol{u}_1$, $\boldsymbol{w}_3 = \boldsymbol{u}_1 - \boldsymbol{u}_2$.

6. Εάν τα διανύσματα $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$, $\boldsymbol{v}_4$ αποτελούν βάση κάποιου διανυσματικού χώρου X δείξτε ότι και τα
$$\boldsymbol{v}_1 + \boldsymbol{v}_2, \quad \boldsymbol{v}_2 + \boldsymbol{v}_3, \quad \boldsymbol{v}_3 + \boldsymbol{v}_4, \quad \boldsymbol{v}_4$$
αποτελούν επίσης βάση του X.

7. Εάν V και W είναι υπόχωροι ενός διανυσματικού χώρου X πεπερασμένης διάστασης, δείξτε ότι
$$\dim(V + W) = \dim V + \dim W - \dim(V \cap W).$$

8. Αν $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_r\}$ και $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_s\}$ είναι βάσεις για τους υπόχωρους U και V, αντίστοιχα, του $\mathbb{R}^n$ και
$$W = \operatorname{span}\{\boldsymbol{u}_i + \boldsymbol{v}_j : i = 1, 2, \ldots, r \;\; j = 1, 2, \ldots, s\},$$
δείξτε ότι $\dim W \le \min\{n, r + s\}$.

9. Έστω V ο υπόχωρος του $C(\mathbb{R})$ που παράγεται από τα διανύσματα $\boldsymbol{v}_1 = \cos^2 x$, $\boldsymbol{v}_2 = \sin^2 x$, και $\boldsymbol{v}_3 = \cos 2x$.

   (α΄) Δείξτε ότι το $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ δεν είναι βάση για τον V.

   (β΄) Να βρεθεί μια βάση για τον V.

10. Να βρεθούν βάσεις για κάθε έναν υπόχωρο των διανυσμάτων $\begin{pmatrix} x & y & z \end{pmatrix}^{\mathrm{T}}$ του $\mathbb{R}^3$ για τον οποίο

   (α΄) $x = y = z$.

   (β΄) $x = 4y = 2z$.

   (γ΄) $z = x + y$.

   (δ΄) $ax + by + cz = 0$, όπου a, b, c είναι πραγματικές σταθερές με $a \neq 0$.

11. Δείξτε ότι η διάσταση του χώρου των $n \times n$ συμμετρικών μητρώων είναι
$$n + \frac{n^2 - n}{2} = \frac{n(n+1)}{2}.$$

12. Έστω ότι W είναι υπόχωρος του πεπερασμένης διάστασης διανυσματικού χώρου X, και έστω $\mathcal{B}_W = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ να είναι μια βάση για τον W. Αποδείξτε ότι η $\mathcal{B}_W$ μπορεί να εμπλουτισθεί με στοιχεία $\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k$ του X ώστε το $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ να είναι μια βάση για τον X, βλέπε Παράδειγμα 4.15.

13. Αποδείξτε ή δώστε αντιπαράδειγμα: Έστω $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ να είναι μια βάση για ένα χώρο X. Εάν V είναι ένας υπόχωρος του X και $\boldsymbol{w}_3, \ldots, \boldsymbol{w}_n \notin V$, τότε η $\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$ είναι μια βάση για τον V.

14. Έστω $V \subset \mathbb{R}^4$, όπου
$$V = \left\{ \begin{pmatrix} x \\ x \\ y \\ y \end{pmatrix} : x, y \in \mathbb{R} \right\}.$$

(α΄) Δείξτε ότι το V είναι υπόχωρος του $\mathbb{R}^4$.

(β΄) Βρείτε μια βάση για τον V.

(γ΄) Να βρεθεί υπόχωρος W ώστε $\mathbb{R}^4 = V \oplus W$.

15. Αν
$$V_1 = \left\{\begin{pmatrix} a & b \\ -b & a \end{pmatrix} : a, b \in \mathbb{R}\right\}, \qquad V_2 = \left\{\begin{pmatrix} c & d \\ d & -c \end{pmatrix} : c, d \in \mathbb{R}\right\},$$

(α΄) Δείξτε ότι τα V_1, V_2 είναι υπόχωροι του $\mathbb{M}^{2,2}(\mathbb{R})$.

(β΄) Βρείτε μια βάση για κάθε έναν από τους V_1, V_2.

(γ΄) Δείξτε ότι $\mathbb{M}^{2,2}(\mathbb{R}) = V_1 \oplus V_2$.

16. Αποδείξτε ή δώστε αντιπαράδειγμα: Εάν U, V, W είναι υπόχωροι ενός διανυσματικού χώρου X ώστε
$$X = U \oplus W \qquad \text{και} \qquad X = V \oplus W$$
τότε $U = V$.

17. Αν $\mathbb{R}^7 = U \oplus V$, δείξτε ότι ισχύει ένα από τα δύο, ή $\dim U < \dim V$, ή $\dim V < \dim U$. Ισχύει το ίδιο συμπέρασμα αν το $\mathbb{R}^7$ αντικατασταθεί με το $\mathbb{R}^6$;

18. Αν $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ και $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ είναι βάσεις στον $\mathbb{R}^3$, δείξτε ότι υπάρχει 3×3 μητρώο M ώστε
$$\begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3 \end{pmatrix} M.$$
Είναι το M αντιστρέψιμο;
Υπόδειξη: Εκφράστε κάθε $\boldsymbol{u}_j$ ως γραμμικό συνδυασμό των $\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3$.

4.7 Παράρτημα V: Άλγεβρες

Ορισμός 4.10. Μια **άλγεβρα** $\mathcal{A}$ είναι ένας διανυσματικός χώρος $\mathcal{A}(\mathbb{F})$ στον οποίο ορίζεται επιπλέον πολλαπλασιασμός μεταξύ των διανυσμάτων έτσι ώστε αν $x, y \in \mathcal{A}$, τότε $xy \in \mathcal{A}$, και ισχύουν οι νόμοι

(A1) $x(yz) = (xy)z$, για κάθε $x, y, z \in \mathcal{A}$.

(A2) $x(y + z) = xy + xz$, και $(x + y)z = xz + yz$ για κάθε $x, y, z \in \mathcal{A}$.

(A3) $a(xy) = (ax)y = x(ay)$, για κάθε $a \in \mathbb{F}$ και για όλα τα $x, y \in \mathcal{A}$.

Ορισμός 4.11. Μια άλγεβρα $\mathcal{A}$ λέγεται **άλγεβρα με μονάδα** εάν

(AI) Υπάρχει $I \in \mathcal{A}$ ώστε $Ix = xI = x$, για κάθε $x \in \mathcal{A}$,

υπάρχει δηλαδή μοναδιαίο διάνυσμα.

Ορισμός 4.12. Μια άλγεβρα $\mathcal{A}$ λέγεται **μεταθετική** εάν

(AC) $xy = yx$, για κάθε $x, y \in \mathcal{A}$,

δηλαδή ο πολλαπλασιασμός διανυσμάτων είναι μεταθετικός.

Παράδειγμα 4.16. Ο διανυσματικός χώρος $C[0,1]$ των συνεχών πραγματικών συναρτήσεων στο $[0,1]$, βλέπε Παράδειγμα 4.1, με τον πολλαπλασιασμό συναρτήσεων

$$(fg)(x) = f(x)g(x), \qquad 0 \leq x \leq 1$$

είναι μεταθετική άλγεβρα με μονάδα τη συνάρτηση I με $I(x) = 1$ για κάθε $x \in [0,1]$.

Παράδειγμα 4.17. Ο χώρος των πολυωνύμων με πραγματικούς συντελεστές $\mathbb{R}[x]$ με τις γνωστές πράξεις πρόσθεση, πολλαπλασιασμό και πολλαπλασιασμό με σταθερά είναι μεταθετική άλγεβρα με μονάδα.

► **Άσκηση 4.6.** Δείξτε ότι το σύνολο $\mathbb{C}$ των μιγαδικών αριθμών με πράξεις την πρόσθεση και τον πολλαπλασιασμό μιγαδικών αριθμών και την εξωτερική πράξη του πολλαπλασιασμού με πραγματική σταθερά είναι μεταθετική άλγεβρα με μονάδα, την οποία ας συμβολίσουμε με $\mathbb{C}(\mathbb{R})$.

► **Άσκηση 4.7.** Δείξτε ότι η δομή $(\mathbb{M}^{n,n}(\mathbb{R}), +, \cdot, *)$, όπου $*$ είναι ο πολλαπλασιασμός μητρώων είναι άλγεβρα με μονάδα. Είναι η άλγεβρα μεταθετική;

Κεφάλαιο 5

Εσωτερικό γινόμενο και ορθογωνιότητα

5.1 Εσωτερικό γινόμενο

Εάν $\boldsymbol{a}$ και $\boldsymbol{b}$ είναι διανύσματα στο $\mathbb{R}^n$ ορίσαμε το σύνηθες εσωτερικό γινόμενο των $\boldsymbol{a}$ και $\boldsymbol{b}$ με τη σχέση

$$\boldsymbol{a}\cdot\boldsymbol{b} = \boldsymbol{a}^{\mathrm{T}}\boldsymbol{b} = \begin{pmatrix} a_1 & a_2 & \cdots & a_n \end{pmatrix}\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_3 \end{pmatrix} = a_1b_1 + a_2b_2 + \cdots + a_nb_n,$$

όπου $\boldsymbol{a}^{\mathrm{T}}\boldsymbol{b}$ ορίζεται, σαν γινόμενο μητρώων. Είδαμε ότι το εσωτερικό γινόμενο ορίζει μια σχέση καθετότητας στο $\mathbb{R}^n$, επιπλέον στο Θεώρημα 1.2 δείξαμε ότι το εσωτερικό γινόμενο έχει τέσσερεις χαρακτηριστικές ιδιότητες. Μέσω αυτών των ιδιοτήτων μπορούμε να ορίσουμε εσωτερικό γινόμενο σε γενικούς διανυσματικούς χώρους.

Ορισμός 5.1. Εάν X είναι ένας πραγματικός διανυσματικός χώρος μια συνάρτηση $\langle\cdot,\cdot\rangle :$ $X\times X\to\mathbb{R}$ θα λέγεται **εσωτερικό γινόμενο** (inner product) εάν ισχύουν τα παρακάτω:

(1) $\langle \boldsymbol{u},\boldsymbol{u}\rangle \geq 0$ και $\langle \boldsymbol{u},\boldsymbol{u}\rangle = 0$ αν και μόνο αν $\boldsymbol{u} = \boldsymbol{0}$.

(2) $\langle \boldsymbol{u}+\boldsymbol{v},\boldsymbol{w}\rangle = \langle \boldsymbol{u},\boldsymbol{w}\rangle + \langle \boldsymbol{v},\boldsymbol{w}\rangle$, για κάθε $\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}\in X$.

(3) $\langle \lambda\boldsymbol{u},\boldsymbol{v}\rangle = \lambda\langle \boldsymbol{u},\boldsymbol{v}\rangle$, για κάθε $\boldsymbol{u},\boldsymbol{v}\in X$ και $\lambda\in\mathbb{R}$.

(4) $\langle \boldsymbol{v},\boldsymbol{u}\rangle = \langle \boldsymbol{u},\boldsymbol{v}\rangle$, για κάθε $\boldsymbol{u},\boldsymbol{v}\in X$.

Ένα εσωτερικό γινόμενο $\langle \boldsymbol{u},\boldsymbol{v}\rangle$ συμβολίζεται συνήθως και ως $\boldsymbol{u}\cdot\boldsymbol{v}$ και για τον λόγο αυτό λέγεται και **στικτό γινόμενο** (dot product). Το εσωτερικό γινόμενο λέγεται επίσης και **βαθμωτό γινόμενο** (scalar product). Ένας διανυσματικός χώρος στον οποίο έχει οριστεί ένα εσωτερικό γινόμενο λέγεται **χώρος με εσωτερικό γινόμενο** .

Παράδειγμα 5.1. Ο χώρος $\mathbb{R}^n$ με $\langle \boldsymbol{u},\boldsymbol{v}\rangle = \boldsymbol{u}^{\mathrm{T}}\boldsymbol{v}$ είναι χώρος με εσωτερικό γινόμενο.

Παρατήρηση 5.1. Εάν το A είναι ένα $n\times m$ μητρώο και το B ένα $m\times k$ μητρώο, με το γνωστό συμβολισμό για τις γραμμές και τις στήλες, από τον ορισμό του πολλαπλασιασμού μητρώων, βλέπε και (2.21), έπεται ότι

$$AB = \left(A_{i*}^{\mathrm{T}} B_{*j}\right) = \begin{pmatrix} \langle A_{1*}, B_{*1}\rangle & \langle A_{1*}, B_{*2}\rangle & \cdots & \langle A_{1*}, B_{*k}\rangle \\ \langle A_{2*}, B_{*1}\rangle & \langle A_{2*}, B_{*2}\rangle & \cdots & \langle A_{2*}, B_{*k}\rangle \\ \vdots & \vdots & & \vdots \\ \langle A_{n*}, B_{*1}\rangle & \langle A_{n*}, B_{*2}\rangle & \cdots & \langle A_{n*}, B_{*k}\rangle \end{pmatrix},$$

έτσι $AB = \left(A_{i*} \cdot B_{*j}\right)$, με $i = 1, 2, \ldots, n$ και $j = 1, 2, \ldots, k$, όπου έχουμε γράψει

$$\langle A_{i*}, B_{*j}\rangle = A_{i*} \cdot B_{*j} = A_{i*}^{\mathrm{T}} B_{*j}.$$

► **Άσκηση 5.1.** Στο $\mathbb{R}^3$ ορίζουμε τη σχέση

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = u_1 v_1 + 2u_2 v_2 + 3u_3 v_3. \tag{5.1}$$

(α΄) Δείξτε ότι η σχέση (5.1) ορίζει ένα εσωτερικό γινόμενο στο R^3.

(β΄) Δείξτε ότι

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = \boldsymbol{u}^{\mathrm{T}} \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{pmatrix} \boldsymbol{v}.$$

► **Άσκηση 5.2.** Εάν $\boldsymbol{u} \cdot \boldsymbol{v}$ είναι το σύνηθες εσωτερικό γινόμενο στο $\mathbb{R}^3$, δηλαδή $\boldsymbol{u} \cdot \boldsymbol{v} = u_1 v_1 + u_2 v_2 + u_3 v_3$, και A είναι ένα αντιστρέψιμο 3×3 μητρώο δείξτε ότι η σχέση

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = A\boldsymbol{u} \cdot A\boldsymbol{v} \tag{5.2}$$

ορίζει ένα εσωτρικό γινόμενο στο $\mathbb{R}^3$. Σημειώνουμε ότι

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = A\boldsymbol{u} \cdot A\boldsymbol{v} = (A\boldsymbol{u})^{\mathrm{T}} A\boldsymbol{v} = \boldsymbol{u}^{\mathrm{T}} A^{\mathrm{T}} A\boldsymbol{v}.$$

Παρατήρηση 5.2. Υπό το πρίσμα του αποτελέσματος της Άσκησης 5.2 βλέπουμε ότι το σύνηθες εσωτερικό γινόμενο στο $\mathbb{R}^3$ υλοποιείται και ως

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle = I\boldsymbol{u} \cdot I\boldsymbol{v}$$

όπου I είναι το 3×3 ταυτοτικό μητρώο. Επιπλέον το εσωτερικό γινόμενο στην Άσκηση 5.1 δίνεται από τη σχέση (5.2) με

$$A = \begin{pmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{3} \end{pmatrix}.$$

Τα αποτελέσματα των Ασκήσεων 5.1 και 5.2 γενικεύονται και ισχύουν στον $\mathbb{R}^n$.

► **Άσκηση 5.3.** Για τα πολυώνυμα $p(x) = a_0 + a_1 x + a_2 x^2$ και $q(x) = b_0 + b_1 x + b_2 x^2$ ορίζουμε

$$\langle p, q\rangle = a_0 b_0 + a_1 b_1 + a_2 b_2.$$

Δείξτε ότι η σχέση αυτή είναι ένα εσωτερικό γινόμενο στο $\mathbb{P}_2$, τον χώρο των πολυωνύμων βαθμού το πολύ δύο. Βλέπε Παρατήρηση 4.7.

Παράδειγμα 5.2. Στο χώρο $C[a,b]$ των πραγματικών συνεχών συναρτήσεων στο διάστημα $[a,b]$ η σχέση

$$\langle f,g\rangle = \int_a^b f(x)g(x)\,\mathrm{d}x \tag{5.3}$$

ορίζει ένα εσωτερικό γινόμενο.
Πράγματι αν $f,g,h \in C[a,b]$ και $\lambda \in \mathbb{R}$, τότε
(1)

$$\langle f,f\rangle = \int_a^b f^2(x)\,\mathrm{d}x \geq 0$$

από τις ιδιότητες του ολοκληρώματος αφού $f^2(x) \geq 0$ για κάθε $x \in [a,b]$. Επιπλέον από τις ιδιότητες των συνεχών συναρτήσεων και του ολοκληρώματος έπεται ότι[1]

$$\langle f,f\rangle = \int_a^b f^2(x)\,\mathrm{d}x = 0 \Leftrightarrow f(x) = 0 \quad \forall x \in [a,b] \Leftrightarrow f = 0.$$

(2) Από τη γραμμικότητα του ολοκληρώματος

$$\begin{aligned}\langle f+g,h\rangle &= \int_a^b (f(x)+g(x))h(x)\,\mathrm{d}x = \int_a^b (f(x)h(x)+g(x)h(x))\,\mathrm{d}x\\ &= \int_a^b f(x)h(x)\,\mathrm{d}x + \int_a^b g(x)h(x)\,\mathrm{d}x\\ &= \langle f,h\rangle + \langle g,h\rangle.\end{aligned}$$

(3) Όμοια

$$\langle \lambda f,g\rangle = \int_a^b \lambda f(x)g(x)\,\mathrm{d}x = \lambda\int_a^b f(x)g(x)\,\mathrm{d}x = \lambda\langle f,g\rangle.$$

(4) Τέλος

$$\langle f,g\rangle = \int_a^b f(x)g(x)\,\mathrm{d}x = \int_a^b g(x)f(x)\,\mathrm{d}x = \langle g,f\rangle.$$

Κατά συνέπεια η (5.3) ορίζει ένα εσωτερικό γινόμενο στο χώρο $C[a,b]$.

▲ **Πρόταση 5.1.** *Σε κάθε διανυσματικό χώρο X με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$ ισχύουν*

(1) *Εάν για κάποιο $\boldsymbol{u} \in X$ είναι $\langle \boldsymbol{x},\boldsymbol{u}\rangle = 0$ για κάθε $\boldsymbol{x} \in X$, τότε $\boldsymbol{u} = \mathbf{0}$.*

(2) *$\langle \mathbf{0},\boldsymbol{u}\rangle = 0$ για κάθε $\boldsymbol{u} \in X$.*

---

[1] **Λήμμα.** *Έστω ότι η h είναι συνεχής στο $[a,b]$, $h(x) \geq 0$ και $\int_a^b h(x)\,\mathrm{d}x = 0$, τότε $h(x) = 0$ στο $[a,b]$.*

Απόδειξη. Πράγματι υποθέτοντας ότι υπάρχει $c \in [a,b]$ με $h(c) = \epsilon > 0$, μπορούμε να υποθέσουμε ότι $c \in (a,b)$ (γιατί;) και από τη συνέχεια της h επεται ότι υπάρχει $\delta > 0$ ώστε $(c-\delta,c+\delta) \subset [a,b]$ και $h(x) \geq \epsilon/2$ στο $(c-\delta,c+\delta)$, οπότε αφού $h(x) \geq 0$ θα είναι

$$\int_a^b h(x)\,\mathrm{d}x \geq \int_{c-\delta}^{c+\delta} h(x)\,\mathrm{d}x \geq \int_{c-\delta}^{c+\delta} \frac{\epsilon}{2}\,\mathrm{d}x = \epsilon\delta > 0$$

το οποίο είναι άτοπο αφού $\int_a^b h(x)\,\mathrm{d}x = 0$. Καταλήξαμε σε άτοπο γιατί υποθέσαμε ότι υπάρχει c με $h(c) \neq 0$, κατά συνέπεια $h(x) = 0$ για κάθε x στο $[a,b]$. □

(3) $\langle \boldsymbol{u}, \lambda \boldsymbol{v} + \mu \boldsymbol{w}\rangle = \lambda\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \mu\langle \boldsymbol{u}, \boldsymbol{w}\rangle$ *για* $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in X$ *και* $\lambda, \mu \in \mathbb{R}$.

Απόδειξη. (1) Πράγματι για $\boldsymbol{x} = \boldsymbol{u}$ έχουμε

$$\langle \boldsymbol{u}, \boldsymbol{u}\rangle = 0 \Rightarrow \boldsymbol{u} = \boldsymbol{0}$$

από τον ορισμό του εσωτερικού γινομένου.
(2) Αν $\boldsymbol{u} \in X$, τότε για $\boldsymbol{v} \in X$, έχουμε

$$\begin{aligned}\langle \boldsymbol{0}, \boldsymbol{u}\rangle &= \langle \boldsymbol{v} - \boldsymbol{v}, \boldsymbol{u}\rangle \\ &= \langle \boldsymbol{v}, \boldsymbol{u}\rangle - \langle \boldsymbol{v}, \boldsymbol{u}\rangle && \text{(από τις (2) και (3))} \\ &= 0.\end{aligned}$$

(3) Αν $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w}$ είναι διανύσματα στον X και λ, μ σταθερές, υπολογίζουμε

$$\begin{aligned}\langle \boldsymbol{u}, \lambda \boldsymbol{v} + \mu \boldsymbol{w}\rangle &= \langle \lambda \boldsymbol{v} + \mu \boldsymbol{w}, \boldsymbol{u}\rangle && \text{(από την (4))} \\ &= \lambda\langle \boldsymbol{v}, \boldsymbol{u}\rangle + \mu\langle \boldsymbol{w}, \boldsymbol{u}\rangle && \text{(από τις (2) και (3))} \\ &= \lambda\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \mu\langle \boldsymbol{u}, \boldsymbol{w}\rangle && \text{(από την (4))}\end{aligned}$$

το οποίο είναι το ζητούμενο. □

Στην περίπτωση μιγαδικού διανυσματικού χώρου, δηλαδή ενός διανυσματικού χώρου πάνω στο σώμα των μιγαδικών αριθμών, έχουμε

Ορισμός 5.2. Εάν X είναι ένας μιγαδικός διανυσματικός χώρος μια συνάρτηση $\langle\cdot,\cdot\rangle : X \times X \to \mathbb{C}$ θα λέγεται **εσωτερικό γινόμενο** (inner product) εάν ισχύουν τα παρακάτω:

(1) $\langle \boldsymbol{u}, \boldsymbol{u}\rangle \geq 0$ και $\langle \boldsymbol{u}, \boldsymbol{u}\rangle = 0$ αν και μόνο αν $\boldsymbol{u} = \boldsymbol{0}$.

(2) $\langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{w}\rangle = \langle \boldsymbol{u}, \boldsymbol{w}\rangle + \langle \boldsymbol{v}, \boldsymbol{w}\rangle$, για κάθε $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in X$.

(3) $\langle \lambda \boldsymbol{u}, \boldsymbol{v}\rangle = \lambda\langle \boldsymbol{u}, \boldsymbol{v}\rangle$, για κάθε $\boldsymbol{u}, \boldsymbol{v} \in X$ και $\lambda \in \mathbb{C}$.

(4) $\langle \boldsymbol{v}, \boldsymbol{u}\rangle = \overline{\langle \boldsymbol{u}, \boldsymbol{v}\rangle}$, για κάθε $\boldsymbol{u}, \boldsymbol{v} \in X$.

Όπου στην (4) με $\overline{z}$ συμβολίζουμε τον συζυγή του μιγαδικού αριθμού z.

Παρατηρούμε ότι εάν $\boldsymbol{u}, \boldsymbol{v}, \boldsymbol{w} \in X$ και $\zeta, \xi \in \mathbb{C}$ έχουμε

$$\begin{aligned}\langle \boldsymbol{u}, \zeta \boldsymbol{v} + \xi \boldsymbol{w}\rangle &= \overline{\langle \zeta \boldsymbol{v} + \xi \boldsymbol{w}, \boldsymbol{u}\rangle} && \text{(από την (4))} \\ &= \overline{\zeta\langle \boldsymbol{v}, \boldsymbol{u}\rangle} + \overline{\xi\langle \boldsymbol{w}, \boldsymbol{u}\rangle} && \text{(από τις (2) και (3))} \\ &= \overline{\zeta}\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \overline{\xi}\langle \boldsymbol{u}, \boldsymbol{w}\rangle && \text{(από την (4)).}\end{aligned}$$

► **Άσκηση 5.4.** Αν $\boldsymbol{a} = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}^{\mathrm{T}}$ και $\boldsymbol{b} = \begin{pmatrix} b_1 & b_2 & \dots & b_n \end{pmatrix}^{\mathrm{T}}$ είναι διανύσματα στο $\mathbb{C}^n$ δείξτε ότι η σχέση

$$\langle \boldsymbol{a}, \boldsymbol{b}\rangle := a_1\overline{b_1} + a_2\overline{b_2} + \cdots + a_n\overline{b_n} \tag{5.4}$$

είναι ένα εσωτρικό γινόμενο.

Ορισμός 5.3. Έστω ότι X είναι ένας διανυσματικός χώρος με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$. Θα λέμε ότι τα διανύσματα $\boldsymbol{u}$ και $\boldsymbol{v}$ του X είναι **ορθογώνια** (orthogonal) ή **κάθετα** (perpendicular) μεταξύ τους αν $\langle\boldsymbol{u},\boldsymbol{v}\rangle = 0$. Αν τα $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι ορθογώνια γράφουμε $\boldsymbol{u}\perp\boldsymbol{v}$.

Για παράδειγμα αν $\boldsymbol{u} = \begin{pmatrix}1 & 2 & 3\end{pmatrix}^{\mathrm{T}}$ και $\boldsymbol{v} = \begin{pmatrix}-2 & 1 & 0\end{pmatrix}^{\mathrm{T}}$, τότε

$$\boldsymbol{u}\cdot\boldsymbol{v} = (1)(-2) + (2)(1) + (3)(0) = -2 + 2 = 0,$$

κατά συνέπεια τα διανύσματα $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι ορθογώνια μεταξύ τους ως προς το σύνηθες εσωτερικό γινόμενο στο $\mathbb{R}^3$.

Παράδειγμα 5.3. Αν $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ είναι η συνήθης βάση του $\mathbb{R}^n$ και $\langle\cdot,\cdot\rangle$ είναι το σύνηθες εσωτερικό γινόμενο, τότε

$$\langle\boldsymbol{e}_i, \boldsymbol{e}_j\rangle = 0, \qquad \text{για } i \neq j,$$

κατά συνέπεια τα διανύσματα $\boldsymbol{e}_j$, ως προς το σύνηθες εσωτωρικό γινόμενο, είναι ορθογώνια μεταξύ τους.

Παρατήρηση 5.3. Εάν X είναι ένας διανυσματικός χώρος με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$ και τα μη μηδενικά διανύσματα $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\}$ είναι ανά δύο ορθογώνια, δηλαδή $\langle\boldsymbol{u}_i, \boldsymbol{u}_j\rangle = 0$ για $i \neq j$, τότε τα διανύσματα είναι γραμμικά ανεξάρτητα. Πράγματι αν

$$c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_n\boldsymbol{u}_n = \boldsymbol{0}$$

δείχνουμε ότι $c_k = 0$ για κάθε $k \in \{1, 2, \dots, n\}$. Για τυχαίο k υπολογίζουμε

$$\begin{aligned} 0 = \langle\boldsymbol{0}, \boldsymbol{u}_k\rangle &= \langle c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + \cdots + c_n\boldsymbol{u}_n, \boldsymbol{u}_k\rangle \\ &= c_1\langle\boldsymbol{u}_1, \boldsymbol{u}_k\rangle + c_2\langle\boldsymbol{u}_2, \boldsymbol{u}_k\rangle + \cdots + c_n\langle\boldsymbol{u}_n, \boldsymbol{u}_k\rangle \\ &= c_k\langle\boldsymbol{u}_k, \boldsymbol{u}_k\rangle \\ &= c_k\|\boldsymbol{u}_k\|^2 \end{aligned}$$

επομένως $c_k = 0$, αφού $\boldsymbol{u}_k \neq \boldsymbol{0}$ και ο ισχυρισμός αποδείχθηκε.

Παράδειγμα 5.4. Ας συμβολίσουμε με $\mathbb{P}_n[-1,1]$ τον χώρο των πολυωνύμων βαθμού έως n περιορισμένων στο διάστημα $[-1,1]$. Μια βάση, η συνήθης, για τον χώρο αυτό είναι η $\mathcal{B} = \{1, x, x^2, \dots, x^n\}$. Παρατηρούμε ότι

$$\begin{aligned} \langle x^k, x^l\rangle &= \int_{-1}^{1} x^k x^l \,\mathrm{d}x \\ &= \frac{1}{k+l+1} x^{k+l+1}\Big]_{-1}^{1} \\ &= \frac{1-(-1)^{k+l+1}}{k+l+1}. \end{aligned}$$

όπου $k, l = 0, 1, \dots, n$ ($x^0 = 1$.) Συμπεραίνουμε λοιπόν ότι αν $k+l$ είναι περιττός αριθμός, και μόνο τότε τα x^k και x^l είναι ορθογώνια μεταξύ τους στο διάστημα $[-1,1]$.

5.2 Ορθογώνιο συμπλήρωμα

Παρατήρηση 5.4. Ας θεωρήσουμε το επίπεδο-υπόχωρο του $\mathbb{R}^3$ με εξίσωση $ax + by + cz = 0$. Αν (x_1, y_1, z_1) είναι σημείο του επιπέδου, τότε το διάνυσμα $\boldsymbol{u} = \begin{pmatrix} a & b & c \end{pmatrix}^{\mathrm{T}}$ είναι ορθογώνιο στο διάνυσμα $\boldsymbol{\xi} = \begin{pmatrix} x_1 & y_1 & z_1 \end{pmatrix}^{\mathrm{T}}$ του επιπέδου αφού

$$\langle \boldsymbol{u}, \boldsymbol{\xi} \rangle = ax_1 + by_1 + cz_1 = 0,$$

δηλαδή το $\boldsymbol{u}$ είναι ορθογώνιο σε κάθε διάνυσμα του επιπέδου, και η εξίσωση του επιπέδου εκφράζει ακριβώς αυτή την ορθογωνιότητα.

Ορισμός 5.4. Έστω X ένας διανυσματικός χώρος με εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$ και έστω S ένα υποσύνολο του X.

(1) Ένα διάνυσμα $\boldsymbol{u}$ του X θα λέγεται ορθογώνιο στο S εάν είναι ορθογώνιο σε κάθε διάνυσμα του S.

(2) Το σύνολο όλων των διανυσμάτων του X τα οποία είναι ορθογώνια στο S λέγεται **ορθογώνιο συμπλήρωμα** (orthogonal complement) του S και συμβολίζεται με $S^{\perp}$, έτσι

$$S^{\perp} = \{\boldsymbol{x} \in X : \langle \boldsymbol{x}, \boldsymbol{s} \rangle = 0 \ \text{ για κάθε } \boldsymbol{s} \in S\}.$$

Παράδειγμα 5.5. Αν X είναι ένας διανυσματικός χώρος με εσωτερικό γινόμενο παρατηρούμε ότι $\{\mathbf{0}\}^{\perp} = X$, αφού για κάθε $\boldsymbol{x} \in X$ είναι $\langle \mathbf{0}, \boldsymbol{x} \rangle = 0$, δηλαδή το ορθογώνιο συμπλήρωμα του $\{\mathbf{0}\}$ είναι ολόκληρος ο χώρος X. Όμοια $X^{\perp} = \{\mathbf{0}\}$ (γιατί;).

Παράδειγμα 5.6. Το υποσύνολο του $\mathbb{R}^2$

$$W = \left\{ \begin{pmatrix} x \\ x \end{pmatrix} : x \in \mathbb{R} \right\}$$

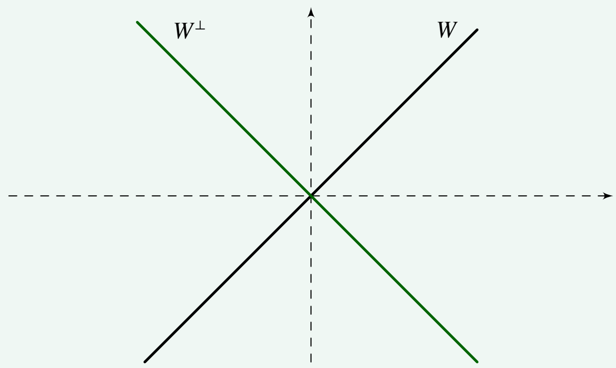
είναι υπόχωρος του $\mathbb{R}^2$. Είναι η ευθεία με εξίσωση $y = x$. Να βρεθεί το ορθογώνιο συμπλήρωμα του $W^{\perp}$.
Αν $\begin{pmatrix} y \\ z \end{pmatrix} \in W^{\perp}$, τότε θα πρέπει να ισχύει

$$\begin{pmatrix} x \\ x \end{pmatrix} \cdot \begin{pmatrix} y \\ z \end{pmatrix} = xy + xz = x(y + z) = 0$$

για κάθε $x \in \mathbb{R}$. Θα πρέπει, επομένως να είναι $z = -y$, κατά συνέπεια, ως προς το σύνηθες εσωτερικό γινόμενο, το ορθογώνιο συμπλήρωμα του W είναι το

$$W^{\perp} = \left\{ \begin{pmatrix} x \\ -x \end{pmatrix} : x \in \mathbb{R} \right\}$$

δηλαδή η ευθεία με εξίσωση $y = -x$.



► **Άσκηση 5.5.** Να βρεθεί το ορθογώνιο συμπλήρωμα καθενός υποσυνόλου του $\mathbb{R}^2$

(α΄) $S_1 = \left\{\begin{pmatrix} x \\ y \end{pmatrix} : y - 2x = 0\right\}$.

(β΄) $S_2 = \left\{\begin{pmatrix} x \\ y \end{pmatrix} : y - x = 1\right\}$.

Απάντηση. $S_2^{\perp} = \{\mathbf{0}\}$.

Παράδειγμα 5.7. Αν

$$W = \left\{\begin{pmatrix} x \\ y \\ z \end{pmatrix} : ax + by + cz = 0\right\}$$

από την Παρατήρηση 5.4 προκύπτει ότι

$$W^{\perp} = \operatorname{span}\left\{\begin{pmatrix} a \\ b \\ c \end{pmatrix}\right\}.$$

Πράγματι αν, δίχως βλάβη της γενικότητας, $c \neq 0$, τότε

$$W = \left\{\begin{pmatrix} x \\ y \\ \alpha x + \beta y \end{pmatrix} : x, y \in \mathbb{R}\right\},$$

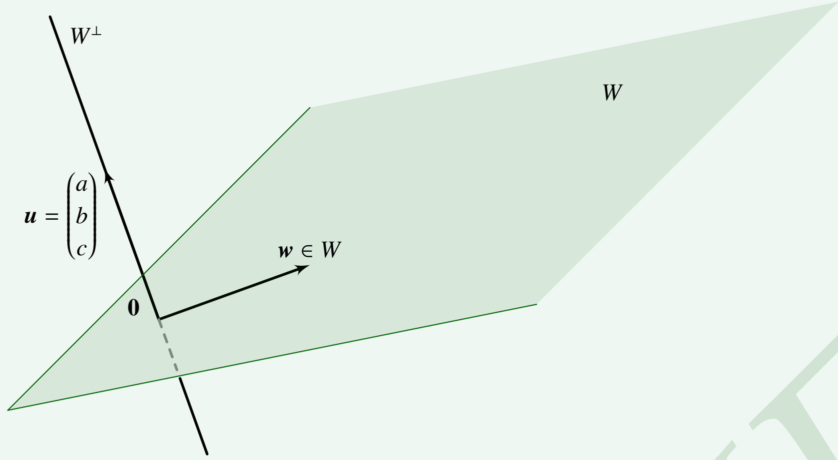
όπου $\alpha = -a/c$ και $\beta = -b/c$. Έτσι αν $\begin{pmatrix} r & s & t \end{pmatrix}^{\mathrm{T}} \in W^{\perp}$, τότε θα πρέπει

$$rx + sy + t(\alpha x + \beta y) = 0 \Leftrightarrow (r + t\alpha)x + (s + t\beta)y = 0$$

για όλα τα x και για όλα τα y, κατά συνέπεια $r = -t\alpha$ και $s = -t\beta$, επομένως

$$\begin{pmatrix} r \\ s \\ t \end{pmatrix} = \begin{pmatrix} -t\alpha \\ -t\beta \\ t \end{pmatrix} = \frac{t}{c}\begin{pmatrix} a \\ b \\ c \end{pmatrix} = \lambda\begin{pmatrix} a \\ b \\ c \end{pmatrix}$$

με $\lambda \in \mathbb{R}$.



Σχήμα 5.1: Ορθογώνιο συμπλήρωμα επιπέδου

Παράδειγμα 5.8 (Διαφορετική προσέγγιση στο Παράδειγμα 5.7). Ας θεωρήσουμε το επίπεδο W με εξίσωση $ax + by + cz = 0$, με $c \neq 0$. Το W είναι υπόχωρος του $\mathbb{R}^3$. Δείξτε ότι

$$W^{\perp} = \operatorname{span}\left\{\begin{pmatrix} a \\ b \\ c \end{pmatrix}\right\}.$$

Βρίσκουμε μια βάση για το W. Για τις λύσεις της $ax + by + cz = 0$ βρίσκουμε για $x = 1$ και $y = 0$, $z = -a/c$, ενώ για $x = 0$ και $y = 1$, $z = -b/c$, κατά συνέπεια τα διανύσματα $\begin{pmatrix} 1 & 0 & -a/c \end{pmatrix}^{\mathrm{T}}$ και $\begin{pmatrix} 0 & 1 & -b/c \end{pmatrix}^{\mathrm{T}}$, ισοδύναμα τα

$$\boldsymbol{w}_1 = \begin{pmatrix} c \\ 0 \\ -a \end{pmatrix} \quad \text{και} \quad \boldsymbol{w}_2 = \begin{pmatrix} 0 \\ c \\ -b \end{pmatrix}$$

αποτελούν μια βάση για τον W. Αν θέσουμε $\boldsymbol{u} = \begin{pmatrix} a & b & c \end{pmatrix}^{\mathrm{T}}$, τότε $\langle \boldsymbol{u}, \boldsymbol{w}_1 \rangle = \langle \boldsymbol{u}, \boldsymbol{w}_2 \rangle = 0$, κατά συνέπεια $\langle \boldsymbol{u}, \lambda \boldsymbol{w}_1 + \mu \boldsymbol{w}_2 \rangle = 0$ για κάθε λ και για κάθε μ, που είναι το ζητούμενο.

Έστω ότι το S είναι υποσύνολο ενός διανυσματικού χώρου X με εσωτερικό γινόμενο, τότε για κάθε διάνυσμα z του S είναι $\langle \mathbf{0}, \boldsymbol{z} \rangle = 0$, κατά συνέπεια $\mathbf{0} \in S^{\perp}$, ειδικά το $S^{\perp}$ είναι μη κενό σύνολο. Αν τώρα $\boldsymbol{x}, \boldsymbol{y} \in S^{\perp}$, $\boldsymbol{z} \in S$ και $\lambda, \mu \in \mathbb{R}$, τότε

$$\begin{aligned} \langle \lambda \boldsymbol{x} + \mu \boldsymbol{y}, \boldsymbol{z} \rangle &= \langle \lambda \boldsymbol{x}, \boldsymbol{z} \rangle + \langle \mu \boldsymbol{y}, \boldsymbol{z} \rangle \\ &= \lambda \langle \boldsymbol{x}, \boldsymbol{z} \rangle + \mu \langle \boldsymbol{y}, \boldsymbol{z} \rangle \\ &= 0 \end{aligned}$$

από τον ορισμό του $S^{\perp}$, κατά συνέπεια $\lambda \boldsymbol{x} + \mu \boldsymbol{y} \in S^{\perp}$. Δείξαμε λοιπόν ότι

Θεώρημα 5.1. *Αν το S είναι υποσύνολο του διανυσματικού χώρου X με εσωτερικό γινόμενο, το ορθογώνιο συμπλήρωμα $S^{\perp}$ του S είναι υπόχωρος του X.*

Παρατήρηση 5.5. Σε σχέση με το Παράδειγμα 5.6 παρατηρούμε ότι οι υπόχωροι W και $W^{\perp}$

αποτελούν ένα ορθογώνιο σύστημα αξόνων και

$$\begin{pmatrix} x \\ y \end{pmatrix} = \frac{1}{2}\begin{pmatrix} x+y \\ x+y \end{pmatrix} + \frac{1}{2}\begin{pmatrix} x-y \\ y-x \end{pmatrix},$$

δηλαδή το τυχαίο διάνυσμα $\boldsymbol{x} \in \mathbb{R}^2$ γράφεται ως το άθροισμα $\boldsymbol{x} = \boldsymbol{w} + \boldsymbol{w}'$, όπου $\boldsymbol{w} \in W$ και $\boldsymbol{w}' \in W^{\perp}$. Το αποτέλεσμα αυτό γενικεύεται.

Θεώρημα 5.2. *Εάν X είναι ένας διανυσματικός χώρος με εσωτερικό γινόμενο και W είναι ένας πεπερασμένης διάστασης υπόχωρος του X, τότε $X = W \oplus W^{\perp}$, δηλαδή ο X είναι το ευθύ άθροισμα των W και $W^{\perp}$.*

Απόδειξη. Για να αποδείξουμε ότι $X = W \oplus W^{\perp}$ πρέπει να δείξουμε ότι αφενός $X = W + W^{\perp}$ και αφετέρου $W \cap W^{\perp} = \{\mathbf{0}\}$. Το Θεώρημα 5.8 σε συνδυασμό με το Πόρισμα 5.2 εξασφαλίζει ότι $X = W + W^{\perp}$. Αρκεί λοιπόν να δείξουμε ότι $W \cap W^{\perp} = \{\mathbf{0}\}$. Πράγματι αν $\boldsymbol{x} \in W \cap W^{\perp}$ τότε

$$\langle \boldsymbol{x}, \boldsymbol{x} \rangle = \mathbf{0}$$

(γιατί;) απ' όπου, από τον ορισμό του εσωτερικού γινομένου, έπεται ότι $\boldsymbol{x} = \mathbf{0}$ το οποίο είναι και το ζητούμενο. □

▼ **Πόρισμα 5.1.** *Αν το W είναι υπόχωρος του διανυσματικού χώρου X διάστασης n, με εσωτερικό γινόμενο, τότε* $\dim W^{\perp} = n - \dim W$.

▲ **Πρόταση 5.2.** *Αν το W είναι υπόχωρος του διανυσματικού χώρου X διάστασης n, με εσωτερικό γινόμενο, τότε $(W^{\perp})^{\perp} = W$.*

Αναρωτιόμαστε αν θα μπορούσε να είναι διαφορετικό το συμπέρασμα αφού

$$X = W \oplus W^{\perp} \qquad \text{και} \qquad X = W^{\perp} \oplus (W^{\perp})^{\perp},$$

σαν συνέπεια του Θεωρήματος 5.2.

Απόδειξη. Δείχνουμε πρώτα ότι $(W^{\perp})^{\perp} \subseteq W$. Αν $\boldsymbol{x} \in (W^{\perp})^{\perp}$, τότε ως στοιχείο του χώρου X γράφεται ως $\boldsymbol{x} = \boldsymbol{w} + \boldsymbol{z}$, όπου $\boldsymbol{w} \in W$ και $\boldsymbol{z} \in W^{\perp}$, βλέπε Θεώρημα 5.2. Έτσι υπολογίζουμε

$$\mathbf{0} = \langle \boldsymbol{x}, \boldsymbol{z} \rangle = \langle \boldsymbol{w} + \boldsymbol{z}, \boldsymbol{z} \rangle = \langle \boldsymbol{w}, \boldsymbol{z} \rangle + \langle \boldsymbol{z}, \boldsymbol{z} \rangle = \langle \boldsymbol{z}, \boldsymbol{z} \rangle$$

αφού $\boldsymbol{w} \perp \boldsymbol{z}$, συνεπώς, από τον ορισμό του εσωτερικού γινομένου, παίρνουμε $\boldsymbol{z} = \mathbf{0}$, ισοδύναμα $\boldsymbol{x} = \boldsymbol{w} \in W$, δηλαδή $(W^{\perp})^{\perp} \subseteq W$. Απομένει να δείξουμε ότι $W \subseteq (W^{\perp})^{\perp}$. Πράγματι αν $\boldsymbol{x} \in W$, τότε $\langle \boldsymbol{x}, \boldsymbol{y} \rangle = 0$ για κάθε $\boldsymbol{y} \in W^{\perp}$, επομένως $\boldsymbol{x} \in (W^{\perp})^{\perp}$, που είναι το ζητούμενο. □

Παρατήρηση 5.6. Κλείνουμε την παράγραφο αυτή με μερικές παρατηρήσεις.

(1) Εάν A είναι ένα $n \times m$ μητρώο, $\boldsymbol{x} \in \mathbb{R}^m$ και $\boldsymbol{y} \in \mathbb{R}^n$, τότε

$$\langle A\boldsymbol{x}, \boldsymbol{y} \rangle = (A\boldsymbol{x})^{\mathrm{T}} \boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}} A^{\mathrm{T}} \boldsymbol{y} = \langle \boldsymbol{x}, A^{\mathrm{T}} \boldsymbol{y} \rangle,$$

όπου το $\langle \cdot, \cdot \rangle$ στα αριστερά είναι το σύνηθες εσωτερικό γινόμενο στο $\mathbb{R}^n$ και στα δεξιά το σύνηθες εσωτερικό γινόμενο στο $\mathbb{R}^m$.

(2) Εάν V και W είναι διανυσματικοί υπόχωροι ενός χώρου X με εσωτερικό γινόμενο και

$$X = V \oplus W = V \oplus V^{\perp}$$

είναι αλήθεια ότι $W = V^{\perp}$; Για την απάντηση βλέπε το Παράδειγμα που ακολουθεί.

Παράδειγμα 5.9 (ή αντιπαράδειγμα). Στο $\mathbb{R}^2$ θέτουμε

$$L_1 = \operatorname{span}\left\{\begin{pmatrix}1\\0\end{pmatrix}\right\}, \qquad L_2 = \operatorname{span}\left\{\begin{pmatrix}0\\1\end{pmatrix}\right\}, \qquad L_3 = \operatorname{span}\left\{\begin{pmatrix}1\\1\end{pmatrix}\right\}.$$

Τότε $L_2 = L_1^{\perp}$ (γιατί;), $\mathbb{R}^2 = L_1 \oplus L_2$ και $\mathbb{R}^2 = L_1 \oplus L_3$, αφού

$$\begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x\\0\end{pmatrix} + \begin{pmatrix}0\\y\end{pmatrix} \qquad \text{και} \qquad \begin{pmatrix}x\\y\end{pmatrix} = \begin{pmatrix}x-y\\0\end{pmatrix} + \begin{pmatrix}y\\y\end{pmatrix},$$

αλλά κάθε άλλο παρά $L_2 = L_3$.

5.3 Νόρμα διανύσματος

Ορισμός 5.5. Εάν X είναι ένας διανυσματικός χώρος μια συνάρτηση $\|\cdot\| : X \to \mathbb{R}$ λέγεται **νόρμα** (norm) εάν για κάθε $\boldsymbol{u}$ και $\boldsymbol{v}$ στο X ικανοποιούνται οι σχέσεις

(1) $\|\boldsymbol{u}\| \geq 0$ και $\|\boldsymbol{u}\| = 0$ αν και μόνο αν $\boldsymbol{u} = \mathbf{0}$.

(2) $\|\lambda \boldsymbol{u}\| = |\lambda| \|\boldsymbol{u}\|$, για κάθε $\lambda \in \mathbb{R}$ (ή $\mathbb{C}$).

(3) $\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$.

Για $\boldsymbol{u} \in X$ ο (μη αρνητικός) πραγματικός αριθμός $\|\boldsymbol{u}\|$ λέγεται νόρμα του $\boldsymbol{u}$.

Η ιδιότητα (3) είναι η **τριγωνική ανισότητα** .

Παρατήρηση 5.7. Απόρροια της τριγωνικής ανισότητας είναι η παρακάτω χρήσιμη ανισότητα:

$$\big|\|\boldsymbol{u}\| - \|\boldsymbol{v}\|\big| \leq \|\boldsymbol{u} - \boldsymbol{v}\|. \tag{5.5}$$

Για την απόδειξη της (5.5) παρατηρούμε ότι

$$\|\boldsymbol{u}\| = \|\boldsymbol{u} - \boldsymbol{v} + \boldsymbol{v}\| \leq \|\boldsymbol{u} - \boldsymbol{v}\| + \|\boldsymbol{v}\| \Rightarrow \|\boldsymbol{u}\| - \|\boldsymbol{v}\| \leq \|\boldsymbol{u} - \boldsymbol{v}\|.$$

Όμοια $\|\boldsymbol{v}\| - \|\boldsymbol{u}\| \leq \|\boldsymbol{u} - \boldsymbol{u}\|$, απ΄ όπου έπεται τελικά η (5.5).

Παρατήρηση 5.8. Στο $\mathbb{R}^n$ ο μέτρο του διανύσματος $\boldsymbol{u} = \begin{pmatrix}u_1 & u_2 & \dots & u_n\end{pmatrix}^{\mathrm{T}}$

$$\|\boldsymbol{u}\| = \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2},$$

συνδέεται με το σύνηθες εσωτερικό γινόμενο μέσω της

$$\|\boldsymbol{u}\| = \sqrt{\boldsymbol{u}^{\mathrm{T}}\boldsymbol{u}} = \sqrt{\langle \boldsymbol{u}, \boldsymbol{u} \rangle},$$

απ΄ όπου προκύπτει, βλέπε Παρατήρηση 1.9, η τριγωνική ανισότητα

$$\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|,$$

κατά συνέπεια το μέτρο του $\boldsymbol{u} \in \mathbb{R}^n$ είναι νόρμα. Έτσι αν $\boldsymbol{a} = \begin{pmatrix}a_1 & a_2 & \dots & a_n\end{pmatrix}^{\mathrm{T}} \in \mathbb{R}^n$ η σχέση

$$\|\boldsymbol{a}\| = \left(\sum_{k=1}^{n} |a_k|^2\right)^{1/2}, \tag{5.6}$$

είναι μια νόρμα στο $\mathbb{R}^n$. Τη νόρμα αυτή λέμε **ℓ_2-νόρμα** στο $\mathbb{R}^n$ και τη συμβολίζουμε με $\|\cdot\|_2$.

Παράδειγμα 5.10. Αν $\boldsymbol{a} = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}^{\mathrm{T}} \in \mathbb{R}^n$ ορίζουμε

$$\|\boldsymbol{a}\| = \sum_{k=1}^{n} |a_k|.$$

Η $\|\cdot\|$ είναι νόρμα. Πράγματι $\|\boldsymbol{a}\| = \sum_{k=1}^{n} |a_k| \geq 0$, σαν άθροισμα μη αρνητικών ποσοτήτων, ενώ $\|\boldsymbol{a}\| = 0$ αν και μόνο αν $|a_k| = 0$, για όλα τα k, άρα ο νόμος (1) στον ορισμό της νόρμας ικανοποιείται. Έστω $\lambda \in \mathbb{R}$, τότε

$$\|\lambda \boldsymbol{a}\| = \sum_{k=1}^{n} |\lambda a_k| = |\lambda| \sum_{k=1}^{n} |a_k|$$

από την οποία προκύπτει η (2). Η (3), η τριγωνική ανισότητα, είναι άμεση συνέπεια της τριγωνικής ανισότητας της απόλυτης τιμής

$$|a_k + b_k| \leq |a_k| + |b_k|, \qquad k = 1, 2, \dots, n.$$

Τη νόρμα αυτή λέμε **ℓ_1-νόρμα** στο $\mathbb{R}^n$ και τη συμβολίζουμε με $\|\cdot\|_1$.

Παράδειγμα 5.11. Αν $\boldsymbol{a} = \begin{pmatrix} a_1 & a_2 & \dots & a_n \end{pmatrix}^{\mathrm{T}} \in \mathbb{R}^n$ ορίζουμε

$$\|\boldsymbol{a}\| = \max_{1 \leq k \leq n} |a_k|.$$

Οι συνθήκες (1) και (2) του ορισμού της νόρμας ικανοποιούνται η δε (3), η τριγωνική ανισότητα, έπεται από το γεγονός ότι

$$|a_k + b_k| \leq |a_k| + |b_k| \leq \max_{1 \leq k \leq n} |a_k| + \max_{1 \leq k \leq n} |b_k|,$$

για κάθε $k = 1, 2, \dots, n$, επομένως

$$\max_{1 \leq k \leq n} |a_k + b_k| \leq \max_{1 \leq k \leq n} |a_k| + \max_{1 \leq k \leq n} |b_k|$$

ισοδύναμα $\|\boldsymbol{a} + \boldsymbol{b}\| \leq \|\boldsymbol{a}\| + \|\boldsymbol{b}\|$. Συνεπώς η $\|\cdot\|$ είναι νόρμα. Συνήθως τη συμβολίζουμε με $\|\cdot\|_\infty$ και τη λέμε **maximum-νόρμα** ή **ℓ_∞-νόρμα** στο $\mathbb{R}^n$.

Παρατήρηση 5.9. Αν $\boldsymbol{u} = \begin{pmatrix} u_1 & u_2 & \dots & u_n \end{pmatrix}^{\mathrm{T}}$ είναι ένα διάνυσμα στον $\mathbb{R}^n$ το μέτρο ή η ℓ_2-νόρμα του $\boldsymbol{u}$ είναι η **ευκλείδεια απόσταση** του σημείου $(u_1, u_2, \dots, u_n)$ από την αρχή των αξόνων, και σαν τέτοια η ℓ_2-νόρμα είναι γενίκευση της απόλυτης τιμής πραγματικού αριθμού στις περισσότερες διαστάσεις. Όμοια, αν $\boldsymbol{v} = \begin{pmatrix} v_1 & v_2 & \dots & v_n \end{pmatrix}^{\mathrm{T}} \in \mathbb{R}^n$, η νόρμα

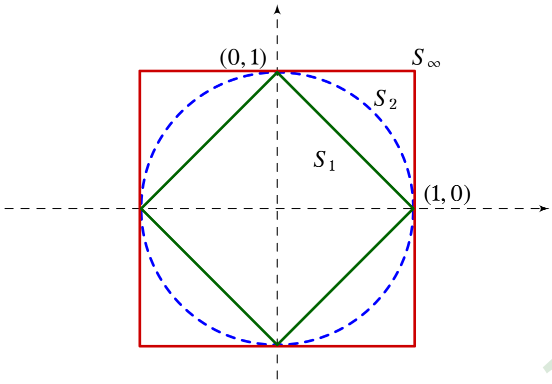
$$\|\boldsymbol{u} - \boldsymbol{v}\| = \sqrt{(u_1 - v_1)^2 + (u_2 - v_2)^2 + \cdots + (u_n - v_n)^2}$$

είναι η απόσταση μεταξύ των σημείων $(u_1, u_2, \dots, u_n)$ και $(v_1, v_2, \dots, v_n)$. Ορίζουμε λοιπόν την απόσταση (distance) των διανυσμάτων $\boldsymbol{u}$ και $\boldsymbol{v}$ του $\mathbb{R}^n$ με τη σχέση

$$d(\boldsymbol{u}, \boldsymbol{v}) = \|\boldsymbol{u} - \boldsymbol{v}\|.$$

Αν $\|\cdot\|$ είναι οποιαδήποτε νόρμα στον $\mathbb{R}^n$, ή και σε τυχαίο διανυσματικό χώρο, η σχέση $\|\boldsymbol{u} - \boldsymbol{v}\|$ είναι και πάλι απόσταση αλλά σε διαφορετική μετρική. Στον $\mathbb{R}^n$ ορίζουμε τη μοναδιαία σφαίρα

S με τη σχέση $S = \{\boldsymbol{x} : \|\boldsymbol{x}\| = 1\}$. Στο $\mathbb{R}^2$ η μοναδιαία σφαίρα είναι ο μοναδιαίος κύκλος. Στο Σχήμα 5.2 εικονίζονται οι μοναδιαίες σφαίρες S_1, S_2, και S_∞ στο $\mathbb{R}^2$, ως προς τις νόρμες ℓ_1, ℓ_2, και ℓ_∞ αντίστοιχα.



Σχήμα 5.2: Οι μοναδιαίες σφαίρες S_1, S_2, S_∞ στο $\mathbb{R}^2$

► **Άσκηση 5.6.** Έστω ότι το $\boldsymbol{a}$ είναι ένα τυχαίο αλλά σταθερό διάνυσμα στο $\mathbb{R}^n$.

(α΄) Να δειχθεί ότι $\|\boldsymbol{a}\|_2 \le \|\boldsymbol{a}\|_1$.

(β΄) Να δειχθεί ότι $\|\boldsymbol{a}\|_\infty \le \|\boldsymbol{a}\|_p \le n^{1/p}\|\boldsymbol{a}\|_\infty$ για $p = 1, 2$.

Παρατήρηση 5.10. Για το σύνηθες εσωτερικό γινόμενο στο $\mathbb{R}^n$ παρατηρούμε ότι

$$|\langle \boldsymbol{u}, \boldsymbol{v}\rangle| = \|\boldsymbol{u}\|\|\boldsymbol{v}\||\cos\theta| \le \|\boldsymbol{u}\|\|\boldsymbol{v}\| \tag{5.7}$$

ισοδύναμα

$$|u_1v_1 + u_2v_2 + \cdots + u_nv_n| \le \sqrt{u_1^2 + u_2^2 + \cdots + u_n^2}\sqrt{v_1^2 + v_2^2 + \cdots + v_n^2} \tag{5.8}$$

(προσπαθήστε να την αποδείξετε). Στην πραγματικότητα η (5.7) είναι ειδική περίπτωση της ανισότητας Cauchy-Bunyakovsky-Schwarz.

Θεώρημα 5.3 (Η ανισότητα Cauchy-Bunyakovsky-Schwarz). *Εάν X είναι ένας πραγματικός διανυσματικός χώρος με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$, τότε*

$$\langle \boldsymbol{u}, \boldsymbol{v}\rangle^2 \le \langle \boldsymbol{u}, \boldsymbol{u}\rangle\langle \boldsymbol{v}, \boldsymbol{v}\rangle$$

για όλα τα διανύσματα $\boldsymbol{u}$ και $\boldsymbol{v}$ του X. Ισότητα ισχύει στην ανισότητα αν και μόνο αν τα $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι γραμμικά εξαρτημένα.

Απόδειξη. Για $t \in \mathbb{R}$ από τις ιδιότητες του εσωτερικού γινομένου έχουμε

$$\begin{aligned} 0 &\le \langle \boldsymbol{u} - t\boldsymbol{v}, \boldsymbol{u} - t\boldsymbol{v}\rangle \\ &= \langle \boldsymbol{u}, \boldsymbol{u}\rangle - 2t\langle \boldsymbol{u}, \boldsymbol{v}\rangle + t^2\langle \boldsymbol{v}, \boldsymbol{v}\rangle. \end{aligned}$$

Αν $\boldsymbol{v} = \boldsymbol{0}$ η προς απόδειξη ανισότητα ισχύει σαν ισότητα με τα δύο μέλη ίσα με μηδέν. Για $\boldsymbol{v} \ne \boldsymbol{0}$ και $t = \langle \boldsymbol{u}, \boldsymbol{v}\rangle/\langle \boldsymbol{v}, \boldsymbol{v}\rangle$ από την τλευταία ανισότητα παίρνουμε

$$0 \le \langle \boldsymbol{u}, \boldsymbol{u}\rangle - \frac{\langle \boldsymbol{u}, \boldsymbol{v}\rangle^2}{\langle \boldsymbol{v}, \boldsymbol{v}\rangle} \Rightarrow \langle \boldsymbol{u}, \boldsymbol{v}\rangle^2 \le \langle \boldsymbol{u}, \boldsymbol{u}\rangle\langle \boldsymbol{v}, \boldsymbol{v}\rangle$$

που είναι το ζητούμενο. Ισχύει η ισότητα στην ανισότητα αν και μόνο αν $\langle \boldsymbol{u} - t\boldsymbol{v}, \boldsymbol{u} - t\boldsymbol{v}\rangle = 0$, ισοδύναμα, από τον ορισμό του εσωτερικού γινομένου $\boldsymbol{u} - t\boldsymbol{v} = \boldsymbol{0}$, ή $\boldsymbol{u} = t\boldsymbol{v}$, δηλαδή τα $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι γραμμικά εξαρτημένα. □

Θεώρημα 5.4. *Σε ένα διανυσματικό χώρο X με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$ η σχέση $\|\boldsymbol{u}\| = \langle \boldsymbol{u}, \boldsymbol{u}\rangle^{1/2}$, $\boldsymbol{u} \in X$ ορίζει νόρμα. Τη νόρμα αυτή λέμε* **επαγόμενη νόρμα***.*

Απόδειξη. Αν $\boldsymbol{u} \in X$, τότε $\langle \boldsymbol{u}, \boldsymbol{u}\rangle^{1/2} \geq 0$ και η συνθήκη (1) του ορισμού της νόρμας έπεται από την αυτήν του εσωτερικού γινομένου. Αν τώρα $\lambda \in \mathbb{R}$ υπολογίζουμε

$$\|\lambda\boldsymbol{u}\|^2 = \langle \lambda\boldsymbol{u}, \lambda\boldsymbol{u}\rangle = \lambda\langle \boldsymbol{u}, \lambda\boldsymbol{u}\rangle = \lambda^2\langle \boldsymbol{u}, \boldsymbol{u}\rangle = \lambda^2\|\boldsymbol{u}\|^2,$$

οπότε $\|\lambda\boldsymbol{u}\| = |\lambda|\|\boldsymbol{u}\|$. Η μόνη μη τετριμμένη συνθήκη είναι η τριγωνική ανισότητα. Αν $\boldsymbol{u}, \boldsymbol{v} \in X$ τότε από τον ορισμό της $\|\cdot\|$ και όπως στην απόδειξη της ανισότητας Cauchy-Bunyakovsky-Schwarz για $t = -1$ έχουμε

$$\begin{aligned}
\|\boldsymbol{u} + \boldsymbol{v}\|^2 &= \langle \boldsymbol{u} + \boldsymbol{v}, \boldsymbol{u} + \boldsymbol{v}\rangle \\
&= \langle \boldsymbol{u}, \boldsymbol{u}\rangle + 2\langle \boldsymbol{u}, \boldsymbol{v}\rangle + \langle \boldsymbol{v}, \boldsymbol{v}\rangle \\
&\leq \|\boldsymbol{u}\|^2 + 2\langle \boldsymbol{u}, \boldsymbol{u}\rangle^{1/2}\langle \boldsymbol{v}, \boldsymbol{v}\rangle^{1/2} + \|\boldsymbol{v}\|^2 \qquad \text{(από Cauchy-Bunyakovsky-Schwarz)} \\
&= \|\boldsymbol{u}\|^2 + 2\|\boldsymbol{u}\|\|\boldsymbol{v}\| + \|\boldsymbol{v}\|^2 \\
&= (\|\boldsymbol{u}\| + \|\boldsymbol{v}\|)^2,
\end{aligned}$$

απ΄ όπου έπεται ότι $\|\boldsymbol{u} + \boldsymbol{v}\| \leq \|\boldsymbol{u}\| + \|\boldsymbol{v}\|$. Η απόδειξη ότι η $\|\cdot\|$ είναι νόρμα είναι πλήρης. □

Υπό το πρίσμα του τελευταίου αποτελέσματος επαναδιατυπώνουμε

Θεώρημα 5.5 (Η ανισότητα Cauchy-Bunyakovsky-Schwarz). *Εάν X είναι ένας πραγματικός διανυσματικός χώρος με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$ και νόρμα που ορίζεται με τη σχέση $\|\cdot\| = \langle\cdot,\cdot\rangle^{1/2}$, τότε*

$$|\langle \boldsymbol{u}, \boldsymbol{v}\rangle| \leq \|\boldsymbol{u}\|\|\boldsymbol{v}\|$$

για όλα τα διανύσματα $\boldsymbol{u}$ και $\boldsymbol{v}$ του X. Ισότητα ισχύει στην ανισότητα αν και μόνο αν τα $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι γραμμικά εξαρτημένα.

► **Άσκηση 5.7.** Κατά την απόδειξη της ανισότητας Cauchy-Bunyakovsky-Schwarz δείξαμε ότι

$$\|\boldsymbol{u}\|^2 - 2t\langle \boldsymbol{u}, \boldsymbol{v}\rangle + t^2\|\boldsymbol{v}\|^2 \geq 0$$

για κάθε $t \in \mathbb{R}$. Από το συμπέρασμα που εξάγεται για την διακρίνουσα αυτού του τριωνύμου (ποιό;) προκύπτει μια άλλη απόδειξη της ανισότητας (πώς;).

Σημείωση 5.1. Στη περίπτωση μιγαδικού διανυσματικού χώρου με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$ ισχύει και πάλι η ανισότητα Cauchy-Bunyakovsky-Schwarz

$$|\langle \boldsymbol{u}, \boldsymbol{v}\rangle|^2 \leq \langle \boldsymbol{u}, \boldsymbol{u}\rangle\langle \boldsymbol{v}, \boldsymbol{v}\rangle.$$

Η απόδειξη είναι παρόμοια με αυτήν της πραγματικής περίπτωσης. Έτσι ξεκινώντας από την ίδια με την πραγματική περίπτωση ανισότητα $0 \leq \langle \boldsymbol{u} - t\boldsymbol{v}\rangle$ με $t \in \mathbb{C}$ παίρνουμε

$$\begin{aligned}
0 &\leq \langle \boldsymbol{u} - t\boldsymbol{v}, \boldsymbol{u} - t\boldsymbol{v}\rangle \\
&= \langle \boldsymbol{u}, \boldsymbol{u}\rangle - \langle \boldsymbol{u}, t\boldsymbol{v}\rangle - \langle t\boldsymbol{v}, \boldsymbol{u}\rangle + \langle t\boldsymbol{v}, t\boldsymbol{v}\rangle \\
&= \langle \boldsymbol{u}, \boldsymbol{u}\rangle - \bar{t}\langle \boldsymbol{u}, \boldsymbol{v}\rangle - t\overline{\langle \boldsymbol{u}, \boldsymbol{v}\rangle} + |t|^2\langle \boldsymbol{v}, \boldsymbol{v}\rangle,
\end{aligned}$$

όπου και πάλι για $\boldsymbol{v} \neq \boldsymbol{0}$ επιλέγοντας $t = \langle \boldsymbol{u}, \boldsymbol{v} \rangle / \langle \boldsymbol{v}, \boldsymbol{v} \rangle$ προκύπτει το ζητούμενο.

► **Άσκηση 5.8.** Αν με $C([a,b],\mathbb{C})$ συμβολίσουμε τον χώρο των συνεχών συναρτήσεων στο διάστημα $[a,b]$ οι οποίες παίρνουν μιγαδικές τιμές, δείξτε ότι η σχέση

$$\langle f, g \rangle = \int_a^b f(x)\overline{g(x)}\,\mathrm{d}x \tag{5.9}$$

ορίζει ένα εσωτερικό γινόμενο. Συμπεράνατε ότι η σχέση

$$\|f\| := \sqrt{\langle f, f \rangle} = \left(\int_a^b f(x)\overline{f(x)}\,\mathrm{d}x \right)^{1/2} = \left(\int_a^b |f(x)|^2\,\mathrm{d}x \right)^{1/2} \tag{5.10}$$

είναι μια νόρμα στο $C([a,b],\mathbb{C})$.

Θεώρημα 5.6 (Το πυθαγόρειο Θεώρημα). *Έστω X ένας διανυσματικός χώρος με εσωτερικό γινόμενο $\langle \cdot,\cdot \rangle$ και επαγόμενη νόρμα $\|\cdot\|$. Εάν $\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n$ είναι διανύσματα του X ανά δύο κάθετα μεταξύ τους, δηλαδή $\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = 0$, αν $i \neq j$, τότε*

$$\|\boldsymbol{u}_1 + \boldsymbol{u}_2 + \cdots + \boldsymbol{u}_n\|^2 = \|\boldsymbol{u}_1\|^2 + \|\boldsymbol{u}_2\|^2 + \cdots \|\boldsymbol{u}_n\|^2.$$

Απόδειξη. Από τον ορισμό της επαγόμενης νόρμας είναι

$$\begin{aligned}
\|\boldsymbol{u}_1 + \boldsymbol{u}_2 + \cdots + \boldsymbol{u}_n\|^2 &= \langle \boldsymbol{u}_1 + \boldsymbol{u}_2 + \cdots + \boldsymbol{u}_n, \boldsymbol{u}_1 + \boldsymbol{u}_2 + \cdots + \boldsymbol{u}_n \rangle \\
&= \langle \boldsymbol{u}_1, \sum_{j=1}^n \boldsymbol{u}_j \rangle + \langle \boldsymbol{u}_2, \sum_{j=1}^n \boldsymbol{u}_j \rangle + \cdots + \langle \boldsymbol{u}_n, \sum_{j=1}^n \boldsymbol{u}_j \rangle \\
&= \sum_{j=1}^n \langle \boldsymbol{u}_1, \boldsymbol{u}_j \rangle + \sum_{j=1}^n \langle \boldsymbol{u}_2, \boldsymbol{u}_j \rangle + \cdots + \sum_{j=1}^n \langle \boldsymbol{u}_n, \boldsymbol{u}_j \rangle \\
&= \langle \boldsymbol{u}_1, \boldsymbol{u}_1 \rangle + \langle \boldsymbol{u}_2, \boldsymbol{u}_2 \rangle + \cdots + \langle \boldsymbol{u}_n, \boldsymbol{u}_n \rangle \\
&= \|\boldsymbol{u}_1\|^2 + \|\boldsymbol{u}_2\|^2 + \cdots \|\boldsymbol{u}_n\|^2.
\end{aligned}$$

Η περίπτωση $n = 2$ δικαιολογεί το όνομα για το θεώρημα. □

► **Άσκηση 5.9 (Ο νόμος του παραλληλογράμμου).** Έστω X ένας διανυσματικός χώρος με εσωτερικό γινόμενο $\langle \cdot,\cdot \rangle$ και επαγόμενη νόρμα $\|\cdot\|$. Δείξτε ότι

$$\|\boldsymbol{x} + \boldsymbol{y}\|^2 + \|\boldsymbol{x} - \boldsymbol{y}\|^2 = 2(\|\boldsymbol{x}\|^2 + \|\boldsymbol{y}\|^2) \tag{5.11}$$

για όλα τα $\boldsymbol{x}$ και $\boldsymbol{y}$ στο X. Σχεδιάστε δύο διανύσματα στο επίπεδο και διαπιστώστε ότι η ταυτότητα εκφράζει το γεγονός ότι το άθροισμα των τετραγώνων των διαγωνίων παραλληλογράμμου ισούται με το διπλάσιο του αθροίσματος των τετραγώνων των πλευρών του.

5.4 Ορθοκανονικά σύνολα

Ορισμός 5.6. Σε ένα διανυσματικό χώρο με εσωτερικό γινόμενο $\langle \cdot,\cdot \rangle$ και επαγόμενη νόρμα $\|\cdot\|$ ένα σύνολο διανυσμάτων $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \cdots, \boldsymbol{u}_n\}$ λέγεται **ορθοκανονικό** (orthonormal) αν

$$\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = 0, \qquad \text{αν } i \neq j, \qquad \text{και} \qquad \|\boldsymbol{u}_i\| = 1, \qquad i = 1, 2, \ldots, n,$$

ισοδύναμα $\langle \boldsymbol{u}_i, \boldsymbol{u}_j \rangle = \delta_{ij}$, όπου δ_{ij} είναι το δέλτα του Kronecker. Μια βάση η οποία είναι ορθοκανονικό σύνολο θα λέγεται **ορθοκανονική βάση**.

Για παράδειγμα η συνήθης βάση $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ του $\mathbb{R}^3$, με

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{e}_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

όπως εύκολα διαπιστώνεται, είναι μια ορθοκανονική βάση ως προς το σύνηθες εσωτερικό γινόμενο. Αντίθετα η βάση $\mathcal{B}_1 = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$ του $\mathbb{R}^3$, με

$$\boldsymbol{b}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{b}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{b}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

δεν είναι ορθοκανονική ως προς το σύνηθες εσωτερικό γινόμενο αφού

$$\boldsymbol{b}_1 \cdot \boldsymbol{b}_2 = (1)(1) + (0)(1) + (0)(0) = 1 \neq 0.$$

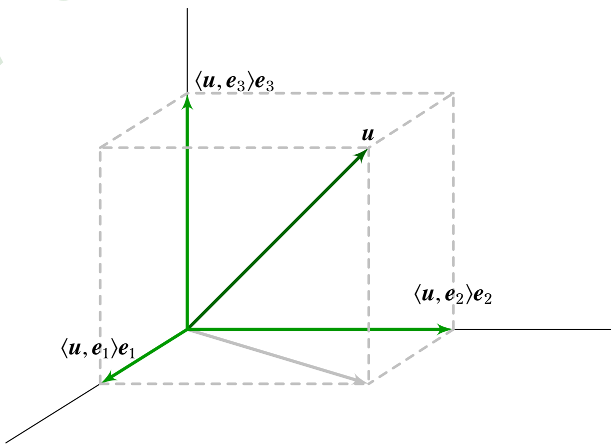
Αν $\boldsymbol{u} = \begin{pmatrix} a & b & c \end{pmatrix}^{\mathrm{T}}$ είναι ένα διάνυσμα στον $\mathbb{R}^3$ και $\langle \cdot, \cdot \rangle$ είναι το σύνηθες εσωτερικό γινόμενο παρατηρούμε ότι

$$\langle \boldsymbol{u}, \boldsymbol{e}_1 \rangle = a, \qquad \langle \boldsymbol{u}, \boldsymbol{e}_2 \rangle = b, \qquad \langle \boldsymbol{u}, \boldsymbol{e}_3 \rangle = c$$

κατά συνέπεια η συνήθης αναπαράσταση $\boldsymbol{u} = a\boldsymbol{e_1} + b\boldsymbol{e_2} + c\boldsymbol{e_3}$ γράφεται ως

$$\boldsymbol{u} = \langle \boldsymbol{u}, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 + \langle \boldsymbol{u}, \boldsymbol{e}_2 \rangle \boldsymbol{e}_2 + \langle \boldsymbol{u}, \boldsymbol{e}_3 \rangle \boldsymbol{e}_3.$$

Οι συνιστώσες δηλαδή του $\boldsymbol{u}$ είναι στην πραγματικότητα τα εσωτερικά γινόμενα του $\boldsymbol{u}$ με τα αντίστοιχα στοιχεία της ορθοκανονικής βάσης $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Γεωμετρικά οι ποσότητες $\langle \boldsymbol{u}, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1$, $\langle \boldsymbol{u}, \boldsymbol{e}_2 \rangle \boldsymbol{e}_2$, και $\langle \boldsymbol{u}, \boldsymbol{e}_3 \rangle \boldsymbol{e}_3$ είναι οι **προβολές** του $\boldsymbol{u}$ στους άξονες που περιέχουν τα αντίστοιχα διανύσματα της βάσης.



Σχήμα 5.3: Συνιστώσες και προβολές στο $\mathbb{R}^3$

Θεώρημα 5.7. *Έστω ότι X είναι ένας διανυσματικός χώρος με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$ και έστω ότι $\mathcal{B} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_n\}$ είναι μία ορθοκανονική βάση για το X, τότε κάθε διάνυσμα $\boldsymbol{x} \in X$ εκφράζεται ως*

$$\boldsymbol{x} = \langle\boldsymbol{x}, \boldsymbol{w}_1\rangle\boldsymbol{w}_1 + \langle\boldsymbol{x}, \boldsymbol{w}_2\rangle\boldsymbol{w}_2 + \cdots + \langle\boldsymbol{x}, \boldsymbol{w}_n\rangle\boldsymbol{w}_n$$

Απόδειξη. Έστω $\boldsymbol{x} \in X$ και έστω $\boldsymbol{x} = x_1\boldsymbol{w}_1 + x_2\boldsymbol{w}_2 + \cdots + x_n\boldsymbol{w}_n$, τότε αν $k \in \{1, 2, \ldots, n\}$

$$\begin{aligned}
\langle\boldsymbol{x}, \boldsymbol{w}_k\rangle &= \langle x_1\boldsymbol{w}_1 + x_2\boldsymbol{w}_2 + \cdots + x_n\boldsymbol{w}_n, \boldsymbol{w}_k\rangle \\
&= x_1\langle\boldsymbol{w}_1, \boldsymbol{w}_k\rangle + x_2\langle\boldsymbol{w}_2, \boldsymbol{w}_k\rangle + \cdots + x_n\langle\boldsymbol{w}_n, \boldsymbol{w}_k\rangle \\
&= x_k\langle\boldsymbol{w}_k, \boldsymbol{w}_k\rangle \\
&= x_k,
\end{aligned}$$

συνέπεια της ορθοκανονικότητας, δηλαδή ο συντελεστής του $\boldsymbol{w}_k$ στο ανάπτυγμα του $\boldsymbol{x}$ ως προς τη βάση $\mathcal{B}$ είναι ίσος με $\langle\boldsymbol{x}, \boldsymbol{w}_k\rangle$. Επειδή το k είναι τυχαίο έπεται το ζητούμενο. □

Ορισμός 5.7. Εάν $\mathcal{B} = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_n\}$ είναι μία ορθοκανονική βάση για τον διανυσματικό χώρο X με εσωτερικό γινόμενο, η έκφραση

$$\boldsymbol{x} = \langle\boldsymbol{x}, \boldsymbol{w}_1\rangle\boldsymbol{w}_1 + \langle\boldsymbol{x}, \boldsymbol{w}_2\rangle\boldsymbol{w}_2 + \cdots + \langle\boldsymbol{x}, \boldsymbol{w}_n\rangle\boldsymbol{w}_n$$

λέγεται **ανάπτυγμα Fourier** του $\boldsymbol{x}$ ως προς τη βάση $\mathcal{B}$. Οι συντελεστές $\langle\boldsymbol{x}, \boldsymbol{w}_k\rangle$ των $\boldsymbol{w}_k$ λέγονται **συντελεστές Fourier** του $\boldsymbol{x}$ ως προς τη βάση $\mathcal{B}$.

Παράδειγμα 5.12. Τα διανύσματα

$$\boldsymbol{b}_1 = \begin{pmatrix}1\\1\end{pmatrix}, \qquad \boldsymbol{b}_2 = \begin{pmatrix}1\\-1\end{pmatrix},$$

είναι γραμμικά ανεξάρτητα, αφού $\boldsymbol{b}_1 \neq \lambda\boldsymbol{b}_2$ για κάθε $\lambda \in \mathbb{R}$ (γιατί;), κατά συνέπεια αποτελούν μια βάση για το $\mathbb{R}^2$. Παρατηρούμε ότι

$$\boldsymbol{b}_1 \cdot \boldsymbol{b}_2 = (1)(1) + (1)(-1) = 0$$

δηλαδή $\boldsymbol{b}_1 \perp \boldsymbol{b}_2$. Επειδή $\|\boldsymbol{b}_1\| = \|\boldsymbol{b}_2\| = \sqrt{2}$, τα διανύσματα

$$\boldsymbol{b}_1' = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\1\end{pmatrix}, \qquad \boldsymbol{b}_2' = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\end{pmatrix}$$

αποτελούν μια ορθοκανονική βάση για το $\mathbb{R}^2$. Αν $\boldsymbol{x} = \begin{pmatrix}2 & -3\end{pmatrix}^{\mathrm{T}}$, τότε

$$\begin{aligned}
\boldsymbol{x} \cdot \boldsymbol{b}_1' &= (2)(1/\sqrt{2}) + (-3)(1/\sqrt{2}) = -1/\sqrt{2} \\
\boldsymbol{x} \cdot \boldsymbol{b}_2' &= (2)(1/\sqrt{2}) + (-3)(-1/\sqrt{2}) = 5/\sqrt{2}.
\end{aligned}$$

Έτσι το ανάπτυγμα Fourier του $\boldsymbol{x}$ ως πρός τη βάση $\{\boldsymbol{b}_1', \boldsymbol{b}_2'\}$ είναι

$$\begin{pmatrix}2\\-3\end{pmatrix} = -\frac{1}{\sqrt{2}}\boldsymbol{b}_1' + \frac{5}{\sqrt{2}}\boldsymbol{b}_2'.$$

Παρατήρηση 5.11. Έστω X ένας διανυσματικός χώρος με εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$, επαγόμενη νόρμα $\| \cdot \|$, και έστω $\boldsymbol{w}$ ένα μοναδιαίο διάνυσμα του X. Αν $\boldsymbol{x}$ είναι διάνυσμα του X, τότε

$$\boldsymbol{x} = \langle \boldsymbol{x}, \boldsymbol{w} \rangle \boldsymbol{w} + \boldsymbol{y} \tag{5.12}$$

για κάποιο $\boldsymbol{y} \in X$ με $\boldsymbol{y} \perp \boldsymbol{w}$. Πράγματι παρατηρούμε ότι το μοναδικό $\boldsymbol{y}$ για το οποίο ισχύει η (5.12) είναι το $\boldsymbol{y} = \boldsymbol{x} - \langle \boldsymbol{x}, \boldsymbol{w} \rangle \boldsymbol{w}$, επομένως

$$\begin{aligned} \langle \boldsymbol{y}, \boldsymbol{w} \rangle &= \langle \boldsymbol{x} - \langle \boldsymbol{x}, \boldsymbol{w} \rangle \boldsymbol{w}, \boldsymbol{w} \rangle \\ &= \langle \boldsymbol{x}, \boldsymbol{w} \rangle - \langle \boldsymbol{x}, \boldsymbol{w} \rangle \langle \boldsymbol{w}, \boldsymbol{w} \rangle \\ &= 0 \end{aligned}$$

αφού $\langle \boldsymbol{w}, \boldsymbol{w} \rangle = \|\boldsymbol{w}\|^2 = 1$. Το αποτέλεσμα γενικεύεται.

Θεώρημα 5.8. *Εάν X είναι ένας διανυσματικός χώρος με εσωτερικό γινόμενο και W είναι ένας μη τετριμμένος, πεπερασμένης διάστασης υπόχωρος του X με ορθοκανονική βάση $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \cdots, \boldsymbol{w}_n\}$, τότε κάθε διάνυσμα $\boldsymbol{u}$ του X μπορεί να εκφραστεί στη μορφή*

$$\boldsymbol{u} = \widehat{\boldsymbol{u}} + \boldsymbol{u}',$$

όπου $\widehat{\boldsymbol{u}} \in W$ και $\boldsymbol{u}'$ είναι ορθογώνιο στον W.

Απόδειξη. Εάν $\boldsymbol{u} \in X$, θέτοντας

$$\widehat{\boldsymbol{u}} = \langle \boldsymbol{u}, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1 + \langle \boldsymbol{u}, \boldsymbol{w}_2 \rangle \boldsymbol{w}_2 + \cdots + \langle \boldsymbol{u}, \boldsymbol{w}_n \rangle \boldsymbol{w}_n$$

έχουμε ότι $\widehat{\boldsymbol{u}} \in W$. Στη συνέχεια ορίζουμε

$$\boldsymbol{u}' = \boldsymbol{u} - \widehat{\boldsymbol{u}}$$

και δείχνουμε ότι αν $\boldsymbol{v} \in W$, τότε $\langle \boldsymbol{v}, \boldsymbol{u}' \rangle = 0$. Επειδή

$$\langle \boldsymbol{v}, \boldsymbol{u}' \rangle = \langle \boldsymbol{v}, \boldsymbol{u} - \widehat{\boldsymbol{u}} \rangle = \langle \boldsymbol{v}, \boldsymbol{u} \rangle - \langle \boldsymbol{v}, \widehat{\boldsymbol{u}} \rangle, \tag{5.13}$$

και

$$\boldsymbol{v} = \langle \boldsymbol{v}, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1 + \langle \boldsymbol{v}, \boldsymbol{w}_2 \rangle \boldsymbol{w}_2 + \cdots + \langle \boldsymbol{v}, \boldsymbol{w}_n \rangle \boldsymbol{w}_n,$$

υπολογίζοντας

$$\langle \boldsymbol{v}, \boldsymbol{u} \rangle = \langle \boldsymbol{v}, \boldsymbol{w}_1 \rangle \langle \boldsymbol{u}, \boldsymbol{w}_1 \rangle + \langle \boldsymbol{v}, \boldsymbol{w}_2 \rangle \langle \boldsymbol{u}, \boldsymbol{w}_2 \rangle + \cdots + \langle \boldsymbol{v}, \boldsymbol{w}_n \rangle \langle \boldsymbol{u}, \boldsymbol{w}_n \rangle \tag{5.14}$$

$$\langle \boldsymbol{v}, \widehat{\boldsymbol{u}} \rangle = \langle \boldsymbol{u}, \boldsymbol{w}_1 \rangle \langle \boldsymbol{v}, \boldsymbol{w}_1 \rangle + \langle \boldsymbol{u}, \boldsymbol{w}_2 \rangle \langle \boldsymbol{v}, \boldsymbol{w}_2 \rangle + \cdots + \langle \boldsymbol{u}, \boldsymbol{w}_n \rangle \langle \boldsymbol{v}, \boldsymbol{w}_n \rangle \tag{5.15}$$

από τις (5.13), (5.14), και (5.15) έπεται ότι $\langle \boldsymbol{v}, \boldsymbol{w}^{\perp} \rangle = 0$. □

5.5 Ορθοκανονικοποίηση Gram-Schmidt

Μια ορθοκανονική βάση σε ένα χώρο με εσωτερικό γινόμενο παράγει ένα ορθογώνιο σύστημα "συντεταγμένων", δηλαδή γενικεύει την έννοια του ορθογωνίου συστήματος αξόνων του $\mathbb{R}^n$. Έτσι οι συντελεστές Fourier του τυχαίου διανύσματος του χώρου είναι οι συντεταγμένες ως προς το ορθογώνιο αυτό σύστημα και είναι εύκολο να υπολογισθούν. Κατά συνέπεια είναι σημαντικό και πρακτικά χρήσιμο σε χώρους με εσωτρικό γινόμενο να γνωρίζουμε μια ορθοκανονική βάση. Ενώ στο $\mathbb{R}^n$ η εκ των προτέρων ύπαρξη μιας τέτοιας βάσης είναι γνωστή δεν συμβαίνει το ίδιο για άλλους χώρους. Για παράδειγμα στο χώρο των πολυωνύμων $\mathbb{P}_3[-1,1]$, βλέπε Παράδειγμα 5.4, με εσωτερικό γινόμενο

$$\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)\,\mathrm{d}x$$

η γνωστή βάση $\{1, x, x^2, x^3\}$, όπως είδαμε, δεν είναι ορθοκανονική. Στη συνέχεια περιγράφουμε μια διαδικασία ορθοκανονικοποίησης, δηλαδή μια διαδικασία όπου ξεκινώντας από ένα γραμμικά ανεξάρτητο σύνολο διανυσμάτων $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ παράγεται με συστηματικό τρόπο ένα ορθοκανονικό σύνολο $S' = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ έτσι ώστε $\operatorname{span} S = \operatorname{span} S'$. Η διαδικασία αυτή λέγεται **ορθοκανονικοποίηση Gram-Schmidt**.

Έστω λοιπόν ένας χώρος X με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$ και επαγόμενη νόρμα $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$, και έστω $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ ένα σύνολο γραμμικά ανεξαρτήτων διανυσμάτων του X.

• Θέτουμε

$$\boldsymbol{w}_1 = \boldsymbol{u}_1/\|\boldsymbol{u}_1\|.$$

Είναι προφανές ότι $\operatorname{span}\{\boldsymbol{u}_1\} = \operatorname{span}\{\boldsymbol{w}_1\}$.

• Θέτουμε

$$\boldsymbol{v}_2 = \boldsymbol{u}_2 - \langle \boldsymbol{u}_2, \boldsymbol{w}_1\rangle \boldsymbol{w}_1$$

αφαιρούμε δηλαδή από το $\boldsymbol{u}_2$ την $\boldsymbol{w}_1$-Fourier συνιστώσα του. Ισχυριζόμαστε ότι $\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\} = \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{v}_2\}$. Αρκεί να δείξουμε ότι αν $\alpha, \beta \in \mathbb{R}$, τότε υπάρχουν $a, b \in \mathbb{R}$ ώστε

$$\alpha\boldsymbol{u}_1 + \beta\boldsymbol{u}_2 = a\boldsymbol{w}_1 + b\boldsymbol{v}_2.$$

Από τον ορισμό των $\boldsymbol{w}_1$ και $\boldsymbol{v}_2$ υπολογίζουμε

$$\begin{aligned}\alpha\boldsymbol{u}_1 + \beta\boldsymbol{u}_2 &= a\boldsymbol{w}_1 + b\boldsymbol{v}_2\\ &= \frac{a\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|} + b\Big(\boldsymbol{u}_2 - \Big\langle \boldsymbol{u}_2, \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|}\Big\rangle \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|}\Big)\end{aligned}$$

οπότε

$$\Big(\alpha - \frac{a}{\|\boldsymbol{u}_1\|} + \frac{b\langle \boldsymbol{u}_2, \boldsymbol{u}_1\rangle}{\|\boldsymbol{u}_1\|^2}\Big)\boldsymbol{u}_1 + (\beta - b)\boldsymbol{u}_2 = \boldsymbol{0}.$$

Επειδή τα $\boldsymbol{u}_1$, $\boldsymbol{u}_2$ είναι γραμμικά ανεξάρτητα βρίσκουμε

$$a = \alpha\|\boldsymbol{u}_1\| + \beta\langle \boldsymbol{u}_2, \boldsymbol{u}_1\rangle/\|\boldsymbol{u}_1\| \qquad \text{και} \qquad b = \beta.$$

Επίσης

$$\begin{aligned}\langle \boldsymbol{w}_1, \boldsymbol{v}_2\rangle &= \langle \boldsymbol{w}_1, \boldsymbol{u}_2 - \langle \boldsymbol{u}_2, \boldsymbol{w}_1\rangle \boldsymbol{w}_1\rangle\\ &= \langle \boldsymbol{w}_1, \boldsymbol{u}_2\rangle - \langle \boldsymbol{u}_2, \boldsymbol{w}_1\rangle\langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle\\ &= 0\end{aligned}$$

αφού $\langle \boldsymbol{w}_1, \boldsymbol{w}_1\rangle = 1$. Θέτοντας

$$\boldsymbol{w}_2 = \boldsymbol{v}_2/\|\boldsymbol{v}_2\|$$

το $\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$ είναι ορθοκανονικό και $\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2\} = \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$.

• Έστω ότι για $k < n$ έχει κατασκευαστεί ένα ορθοκανονικό σύνολο $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}$ τέτοιο ώστε $\operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\} = \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}$. Θέτουμε

$$\boldsymbol{v}_{k+1} = \boldsymbol{u}_{k+1} - \langle \boldsymbol{u}_{k+1}, \boldsymbol{w}_1\rangle \boldsymbol{w}_1 - \langle \boldsymbol{u}_{k+1}, \boldsymbol{w}_2\rangle \boldsymbol{w}_2 - \cdots - \langle \boldsymbol{u}_{k+1}, \boldsymbol{w}_k\rangle \boldsymbol{w}_k$$

αφαιρούμε δηλαδή από το $\boldsymbol{u}_{k+1}$ τις $\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k$-Fourier συνιστώσες του. Για $j < k+1$ υπολογίζουμε

$$\begin{aligned}
\langle \boldsymbol{w}_j, \boldsymbol{v}_{k+1} \rangle &= \langle \boldsymbol{w}_j, \boldsymbol{u}_{k+1} \rangle - \langle \boldsymbol{u}_{k+1}, \boldsymbol{w}_1 \rangle \langle \boldsymbol{w}_j, \boldsymbol{w}_1 \rangle - \langle \boldsymbol{u}_{k+1}, \boldsymbol{w}_2 \rangle \langle \boldsymbol{w}_j, \boldsymbol{w}_2 \rangle \\
&\qquad - \cdots - \langle \boldsymbol{u}_{k+1}, \boldsymbol{w}_k \rangle \langle \boldsymbol{w}_j, \boldsymbol{w}_k \rangle \\
&= \langle \boldsymbol{w}_j, \boldsymbol{u}_{k+1} \rangle - \langle \boldsymbol{u}_{k+1}, \boldsymbol{w}_j \rangle \langle \boldsymbol{w}_j, \boldsymbol{w}_j \rangle \\
&= 0
\end{aligned}$$

αφού $\langle \boldsymbol{w}_j, \boldsymbol{w}_j \rangle = 1$. Θέτοντας $\boldsymbol{w}_{k+1} = \boldsymbol{v}_{k+1}/\|\boldsymbol{v}_{k+1}\|$ το $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k, \boldsymbol{w}_{k+1}\}$ παράγει ότι και το $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k, \boldsymbol{u}_{k+1}\}$, αφού

$$\begin{aligned}
\boldsymbol{u}_{k+1} &= \langle \boldsymbol{u}_{k+1}, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1 + \langle \boldsymbol{u}_{k+1}, \boldsymbol{w}_2 \rangle \boldsymbol{w}_2 + \cdots + \langle \boldsymbol{u}_{k+1}, \boldsymbol{w}_k \rangle \boldsymbol{w}_k + \|\boldsymbol{v}_{k+1}\| \boldsymbol{w}_{k+1} \\
&= c_1 \boldsymbol{w}_1 + c_2 \boldsymbol{w}_2 + \cdots + c_k \boldsymbol{w}_k + c_{k+1} \boldsymbol{w}_{k+1},
\end{aligned}$$

κατά συνέπεια, από την υπόθεση της επαγωγής, έπεται ότι

$$\operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_{k+1}\} = \operatorname{span}\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_{k+1}\}.$$

Αποδείξαμε λοιπόν, με επαγωγή, το

Θεώρημα 5.9 (Ορθοκανονικοποίηση Gram-Schmidt). *Αν σε χώρο X με εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$ και επαγόμενη νόρμα $\| \cdot \| = \sqrt{\langle \cdot, \cdot \rangle}$ το $S = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ είναι ένα σύνολο γραμμικά ανεξαρτήτων διανυσμάτων, τότε η ακολουθία των διανυσμάτων που ορίζονται με τη σχέση*

$$\boldsymbol{w}_1 = \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|}, \qquad \boldsymbol{w}_k = \frac{\boldsymbol{u}_k - \langle \boldsymbol{u}_k, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1 - \cdots - \langle \boldsymbol{u}_k, \boldsymbol{w}_{k-1} \rangle \boldsymbol{w}_{k-1}}{\|\boldsymbol{u}_k - \langle \boldsymbol{u}_k, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1 - \cdots - \langle \boldsymbol{u}_k, \boldsymbol{w}_{k-1} \rangle \boldsymbol{w}_{k-1}\|}, \tag{5.16}$$

$k = 2, \ldots, n$ είναι ένα ορθοκανονικό σύνολο στο X το οποίο παράγει τον ίδιο υπόχωρο με το S. Ειδικά αν το S είναι μια βάση του X, το $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ είναι μια ορθοκανονική βάση για τον X.

Παράδειγμα 5.13. Τα διανύσματα

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{u}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

αποτελούν μια βάση για τον $\mathbb{R}^3$. Ως προς το σύνηθες εσωτερικό γινόμενο τα διανύσματα δεν είναι ανά δύο ορθογώνια. Ακολουθώντας τη διαδικασία Gram-Schmidt ορίζουμε

$$\boldsymbol{w}_1 = \frac{\boldsymbol{u}_1}{\|\boldsymbol{u}_1\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

και

$$\boldsymbol{v}_2 = \boldsymbol{u}_2 - \langle \boldsymbol{u}_2, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{3} 2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \qquad \|\boldsymbol{v}_2\| = \frac{\sqrt{2}}{\sqrt{3}},$$

τότε

$$\boldsymbol{w}_2 = \frac{\boldsymbol{v}_2}{\|\boldsymbol{v}_2\|} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

Όμοια

$$\boldsymbol{v}_3 = \boldsymbol{u}_3 - \langle \boldsymbol{u}_3, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1 - \langle \boldsymbol{u}_3, \boldsymbol{w}_2 \rangle \boldsymbol{w}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{6}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}, \qquad \|\boldsymbol{v}_3\| = \frac{1}{\sqrt{2}},$$

οπότε

$$\boldsymbol{w}_3 = \frac{\boldsymbol{v}_3}{\|\boldsymbol{v}_3\|} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Έτσι αρχίζοντας με τα $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{u}_3$ η διαδικασία Gram-Schmidt παράγει την ορθοκανονική βάση

$$\boldsymbol{w}_1 = \frac{1}{\sqrt{3}}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \boldsymbol{w}_2 = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \qquad \boldsymbol{w}_3 = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Παρατήρηση 5.12. Παρατηρούμε ότι η ορθοκανονική βάση $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \boldsymbol{w}_3\}$ του Παραδείγματος 5.13 που προέκυψε από την διαδικασία ορθοκανονικοποίησης της $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ δεν είναι τόσο εύχρηστη όσο η κανονική βάση $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$. Για παράδειγμα ποιά είναι η αναπαράσταση του $\begin{pmatrix} 2 & -1 & 3 \end{pmatrix}^{\mathrm{T}}$ σε αυτή τη βάση; Αλλάζοντας την αρίθμηση στα $\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3$ βλέπουμε ότι προκύπτουν $3!$ ορθοκανονικές βάσεις για το $\mathbb{R}^3$.

► **Άσκηση 5.10.** Αν $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{u}_3$ είναι όπως στο Παράδειγμα 5.13, θέτουμε $\boldsymbol{b}_1 = \boldsymbol{u}_3$, $\boldsymbol{b}_2 = \boldsymbol{u}_2$, $\boldsymbol{b}_3 = \boldsymbol{u}_1$. Να βρεθεί η ορθοκανονική βάση του $\mathbb{R}^3$ που παράγει η διαδικασία Gram-Schmidt από την βάση $\{\boldsymbol{b}_1, \boldsymbol{b}_2, \boldsymbol{b}_3\}$.

▼ **Πόρισμα 5.2.** *Κάθε μη τετριμμένος, διάφορος δηλαδή του $\{\boldsymbol{0}\}$, διανυσματικός χώρος πεπερασμένης διάστασης έχει μια ορθοκανονική βάση.*

Απόδειξη. Εφαρμόζοντας την διαδικασία Gram-Schmidt σε μια βάση $\{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ του χώρου προκύπτει μια ορθοκανονική βάση. □

Παρατήρηση 5.13. Άμεση συνέπεια του Πορίσματος 5.2 είναι ότι το Θεώρημα 5.8 ισχύει πάντα και είναι ανεξάρτητο από την ορθοκανονική βάση του W.

5.6 Ορθογώνια μητρώα

Τα 2×2 μητρώα

$$\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix}, \quad \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}, \quad \begin{pmatrix} 0 & -1 \\ -1 & 0 \end{pmatrix},$$

έχουν ορθοκανονικές στήλες. Την ίδια δομή έχουν τα δύο 2×2 μητρώα μετάθεσης

$$\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \qquad \text{και} \qquad \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}.$$

καθώς και τα 3×3 μητρώα μετάθεσης της παραγράφου 2.3.1. Γενικότερα αν $\boldsymbol{q}_1, \boldsymbol{q}_2, \cdots, \boldsymbol{q}_n$ είναι ορθοκανονικά διανύσματα στον $\mathbb{R}^n$, τότε $\boldsymbol{q}_i \cdot \boldsymbol{q}_j = \delta_{ij}$ (το δέλτα του Kronecker), κατά συνέπεια αν ορίσουμε το μητρώο $Q = \begin{pmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots & \boldsymbol{q}_n \end{pmatrix}$, τότε

$$Q^{\mathrm{T}}Q = \begin{pmatrix} \boldsymbol{q}_1^{\mathrm{T}} \\ \boldsymbol{q}_2^{\mathrm{T}} \\ \vdots \\ \boldsymbol{q}_n^{\mathrm{T}} \end{pmatrix} \begin{pmatrix} \boldsymbol{q}_1 & \boldsymbol{q}_2 & \cdots & \boldsymbol{q}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{q}_i^{\mathrm{T}}\boldsymbol{q}_j \end{pmatrix} = I.$$

Επειδή το μητρώο Q είναι τετραγωνικό έπεται ότι $QQ^{\mathrm{T}} = I$ και $Q^{\mathrm{T}} = Q^{-1}$. Δείξαμε λοιπόν ότι για ένα τετραγωνικό μητρώο Q οι ισχυρισμοί που ακολουθούν είναι ισοδύναμοι:

(1) Οι στήλες του Q είναι ορθοκανονικά διανύσματα.

(2) Οι γραμμές του Q είναι ορθοκανονικά διανύσματα.

(3) $Q^{-1} = Q^{\mathrm{T}}$.

Ορισμός 5.8. Ένα τετραγωνικό μητρώο Q λέγεται **ορθογώνιο** (orthogonal) αν οι στήλες του είναι ορθοκανονικά διανύσματα.

Στη συνέχεια δίνουμε κάποια παραδείγματα χαρακτηριστικών ορθογωνίων 2×2 μητρώων. Σημειώνουμε ότι κάθε ορθογώνιο μητρώο στο επίπεδο είναι της μορφής

$$Q_\omega^+ = \begin{pmatrix} \cos\omega & -\sin\omega \\ \sin\omega & \cos\omega \end{pmatrix}, \quad \text{ή} \quad Q_\omega^- = \begin{pmatrix} \cos\omega & \sin\omega \\ \sin\omega & -\cos\omega \end{pmatrix}, \tag{5.17}$$

για κατάλληλο $\omega \in \mathbb{R}$.

Παράδειγμα 5.14 (Περιστροφή στο επίπεδο). Το μητρώο

$$Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix} \tag{5.18}$$

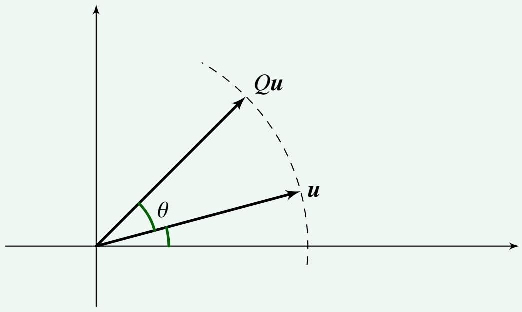
είναι ορθογώνιο μητρώο. Ας δούμε το αποτέλεσμα της δράσης του Q επί ενός μοναδιαίου διανύσματος του $\mathbb{R}^2$, $\begin{pmatrix} \cos\phi & \sin\phi \end{pmatrix}^{\mathrm{T}}$. Υπολογίζοντας

$$\begin{aligned} Q\begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix} &= \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}\begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix} \\ &= \begin{pmatrix} \cos\theta\cos\phi - \sin\theta\sin\phi \\ \sin\theta\cos\phi + \cos\theta\sin\phi \end{pmatrix} \\ &= \begin{pmatrix} \cos(\phi+\theta) \\ \sin(\phi+\theta) \end{pmatrix} \end{aligned}$$

βλέπουμε ότι ο πολλαπλασιασμός επί Q έχει ως αποτέλεσμα την περιστροφή του μοναδιαίου διανύσματος, άρα και κάθε διανύσματος, κατά γωνία θ, αφού αν $\boldsymbol{u}$ είναι ένα διάνυσμα στο $\mathbb{R}^2$ με $r = \|\boldsymbol{u}\|$, τότε

$$Q\boldsymbol{u} = Q\begin{pmatrix} r\cos\phi \\ r\sin\phi \end{pmatrix} = rQ\begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix}.$$

Το αποτέλεσμα αποτυπώνεται στο Σχήμα 5.4.



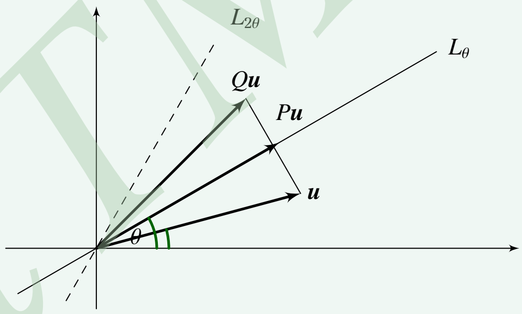
Σχήμα 5.4: Περιστροφή του διανύσματος $\boldsymbol{u}$ κατά γωνία θ.

Αν γράψουμε $Q(\theta)$ για το μητρώο στην (5.18) παρατηρούμε ότι

$$Q(-\theta) = \begin{pmatrix} \cos(-\theta) & -\sin(-\theta) \\ \sin(-\theta) & \cos(-\theta) \end{pmatrix} = \begin{pmatrix} \cos\theta & \sin\theta \\ -\sin\theta & \cos\theta \end{pmatrix} = Q(\theta)^{\mathrm{T}},$$

κατά συνέπεια, αφού το μητρώο είναι ορθογώνιο, $Q(\theta)^{-1} = Q(-\theta)$.

Παράδειγμα 5.15 (Ανάκλαση στο επίπεδο). Αν $\boldsymbol{q} = \begin{pmatrix}\cos\theta & \sin\theta\end{pmatrix}^{\mathrm{T}}$ είναι ένα μοναδιαίο διάνυσμα και L_θ η ευθεία που το περιέχει θέλουμε να βρούμε το μητρώο Q το οποίο υλοποιεί την ανάκλαση οποιουδήποτε διανύσματος ως προς την ευθεία.



Σχήμα 5.5: Ανάκλαση διανύσματος ως προς ευθεία.

Αν $\boldsymbol{u}$ είναι τυχαίο διάνυσμα στο επίπεδο, $Q\boldsymbol{u}$ είναι η ανάκλασή του ως προς την ευθεία L_θ, και $P\boldsymbol{u}$ η προβολή του επί της ευθείας L_θ, τότε από από το παρραλληλόγραμμο με προσκείμενες πλευρές τα διανύσματα $\boldsymbol{u}$ και $Q\boldsymbol{u}$ στο Σχήμα 5.5, παίρνουμε

$$Q\boldsymbol{u} + \boldsymbol{u} = 2P\boldsymbol{u} \Rightarrow Q\boldsymbol{u} = 2P\boldsymbol{u} - \boldsymbol{u}$$

επομένως $Q\boldsymbol{u} = (2P - I)\boldsymbol{u}$, ισοδύναμα $Q = 2P - I$. Επειδή

$$P\boldsymbol{u} = \langle \boldsymbol{u}, \boldsymbol{q}\rangle \boldsymbol{q} = \boldsymbol{q}\langle \boldsymbol{q}, \boldsymbol{u}\rangle = \boldsymbol{q}(\boldsymbol{q}^{\mathrm{T}}\boldsymbol{u}) = (\boldsymbol{q}\boldsymbol{q}^{\mathrm{T}})\boldsymbol{u}$$

τελικά βρίσκουμε

$$Q = 2\boldsymbol{q}\boldsymbol{q}^{\mathrm{T}} - I. \tag{5.19}$$

Αναλυτικά το μητρώο Q είναι

$$\begin{aligned} Q &= 2\begin{pmatrix} \cos^2\theta & \sin\theta\cos\theta \\ \sin\theta\cos\theta & \sin^2\theta \end{pmatrix} - I \\ &= \begin{pmatrix} 2\cos^2\theta - 1 & 2\sin\theta\cos\theta \\ 2\sin\theta\cos\theta & 2\sin^2\theta - 1 \end{pmatrix} \\ &= \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix} \end{aligned}$$

του οποίου οι στήλες είναι ορθοκανονικά διανύσματα στο $\mathbb{R}^2$, επομένως είναι ένα ορθογώνιο μητρώο. Η δράση του Q επί ενός μοναδιαίου διανύσματος $\begin{pmatrix}\cos\phi & \sin\phi\end{pmatrix}^{\mathrm{T}}$ έχει σαν αποτέλεσμα το διάνυσμα

$$\begin{aligned} Q\begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix} &= \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}\begin{pmatrix} \cos\phi \\ \sin\phi \end{pmatrix} \\ &= \begin{pmatrix} \cos 2\theta\cos\phi + \sin 2\theta\sin\phi \\ \sin 2\theta\cos\phi - \cos 2\theta\sin\phi \end{pmatrix} \\ &= \begin{pmatrix} \cos(2\theta-\phi) \\ \sin(2\theta-\phi) \end{pmatrix} \end{aligned}$$

το οποίο συμφωνεί με την εικόνα του σχήματος ότι το αρχικό διάνυσμα, ας πούμε, $\boldsymbol{u} = \begin{pmatrix}\cos\phi & \sin\phi\end{pmatrix}^{\mathrm{T}}$ περιστρέφεται κατά 2θ αρχικά και στη συνέχεια κατά $-\phi$, δηλαδή κατά ϕ με αντίθετη φορά, προκειμένου να προκύψει τελικά το $Q\boldsymbol{u}$.

Παρατήρηση 5.14. Στο Παράδειγμα 5.15 δείξαμε ότι η προβολή ενός διανύσματος επί ενός μοναδιαίου διανύσματος $\boldsymbol{q}$ υλοποιείται μέσω πολλαπλασιασμού του διανύσματος με το μητρώο

$$P = \boldsymbol{q}\boldsymbol{q}^{\mathrm{T}}.$$

Τις προβολές θα μελετήσουμε αναλυτικά σε επόμενο κεφάλαιο.

Θεώρημα 5.10. *Αν Q είναι ένα $n\times n$ πραγματικό μητρώο, $\langle\cdot,\cdot\rangle$ είναι το σύνηθες εσωτερικό γινόμενο και $\|\cdot\|$ η επαγόμενη νόρμα, οι παρακάτω προτάσεις είναι ισοδύναμες*

(1) *Το Q είναι ορθογώνιο.*

(2) *Το Q διατηρεί το μήκος, δηλαδή για κάθε $\boldsymbol{x} \in \mathbb{R}^n$ ισχύει $\|Q\boldsymbol{x}\| = \|\boldsymbol{x}\|$.*

(3) *Το Q διατηρεί το εσωτερικό γινόμενο, δηλαδή για $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ ισχύει $\langle Q\boldsymbol{x}, Q\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle$.*

Απόδειξη. Αποδεικνύουμε τις ισοδυναμίες δείχνοντας ότι (1) $\Rightarrow$ (2) $\Rightarrow$ (3) $\Rightarrow$ (1). Θυμίζουμε ότι αν A, B είναι $n \times n$ μητρώα, τότε

$$\langle A\boldsymbol{x}, B\boldsymbol{y}\rangle = (A\boldsymbol{x})^{\mathrm{T}}B\boldsymbol{y} = \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}B\boldsymbol{y} = \langle x, A^{\mathrm{T}}B\boldsymbol{y}\rangle.$$

(1) $\Rightarrow$ (2) Από τον ορισμό του εσωτερικού γινομένου και της νόρμας στο $\mathbb{R}^n$ για κάθε $\boldsymbol{x} \in \mathbb{R}^n$ είναι

$$\|Q\boldsymbol{x}\|^2 = \langle Q\boldsymbol{x}, Q\boldsymbol{x}\rangle = \langle \boldsymbol{x}, Q^{\mathrm{T}}Q\boldsymbol{x}\rangle = \langle \boldsymbol{x}, \boldsymbol{x}\rangle = \|\boldsymbol{x}\|^2$$

απ΄ όπου έπεται το ζητούμενο.

(2) ⇒ (3) Για κάθε $\boldsymbol{x} \in \mathbb{R}^n$ και για κάθε $\boldsymbol{y} \in \mathbb{R}^n$, από την υπόθεση έχουμε

$$\begin{aligned}\|Q(\boldsymbol{x}+\boldsymbol{y})\|^2 = \|\boldsymbol{x}+\boldsymbol{y}\|^2 &\Leftrightarrow \langle Q(\boldsymbol{x}+\boldsymbol{y}), Q(\boldsymbol{x}+\boldsymbol{y})\rangle = \langle \boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}\rangle \\ &\Leftrightarrow \langle Q\boldsymbol{x}+Q\boldsymbol{y}, Q\boldsymbol{x}+Q\boldsymbol{y}\rangle = \langle \boldsymbol{x}+\boldsymbol{y}, \boldsymbol{x}+\boldsymbol{y}\rangle\end{aligned}$$

απ΄ όπου κάνοντας πράξεις παίρνουμε

$$\begin{aligned}\langle Q\boldsymbol{x}, Q\boldsymbol{x}\rangle + 2\langle Q\boldsymbol{x}, Q\boldsymbol{y}\rangle + \langle Q\boldsymbol{y}, Q\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{x}\rangle + 2\langle \boldsymbol{x}, \boldsymbol{y}\rangle + \langle \boldsymbol{y}, \boldsymbol{y}\rangle \qquad &\Leftrightarrow \\ \langle Q\boldsymbol{x}, Q\boldsymbol{y}\rangle = \langle \boldsymbol{x}, \boldsymbol{y}\rangle\end{aligned}$$

που είναι το ζητούμενο.

(3) ⇒ (1) Για κάθε $\boldsymbol{x} \in \mathbb{R}^n$ και για κάθε $\boldsymbol{y} \in \mathbb{R}^n$, από την υπόθεση έχουμε

$$\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \langle Q\boldsymbol{x}, Q\boldsymbol{y}\rangle = \langle \boldsymbol{x}, Q^{\mathrm{T}}Q\boldsymbol{y}\rangle \Rightarrow \langle \boldsymbol{x}, Q^{\mathrm{T}}Q\boldsymbol{y}\rangle - \langle \boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, (Q^{\mathrm{T}}Q - I)\boldsymbol{y}\rangle = 0$$

όπου I είναι το ταυτοτικό μητρώο. Από την τελευταία σχέση για κάθε $\boldsymbol{y}$ και $\boldsymbol{x} = (Q^{\mathrm{T}}Q - I)\boldsymbol{y}$ έχουμε

$$\langle (Q^{\mathrm{T}}Q - I)\boldsymbol{y}, (Q^{\mathrm{T}}Q - I)\boldsymbol{y}\rangle = 0$$

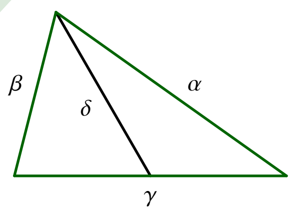
για κάθε $\boldsymbol{y} \in \mathbb{R}^n$, οπότε από τον ορισμό του εσωτερικού γινομένου έπεται ότι το $Q^{\mathrm{T}}Q - I$ είναι το μηδενικό μητρώο, ισοδύναμα $Q^{\mathrm{T}}Q = I$ δηλαδή το Q είναι ορθογώνιο. □

5.7 Ασκήσεις

1. Έστω X ένας διανυσματικός χώρος με εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$. Εάν $\boldsymbol{u} \in X$ και $\langle \boldsymbol{u}, \boldsymbol{v}\rangle$ για κάθε διάνυσμα $\boldsymbol{v} \neq \boldsymbol{u}$, δείξτε ότι $\boldsymbol{u} = \boldsymbol{0}$.

2. Βρείτε δύο ορθοκανονικά διανύσματα του επιπέδου $x - 2y + z = 0$. Είναι τα διανύσματα που βρήκατε μια βάση για το επίπεδο; Εξηγείστε.

3. Για το τρίγωνο του σχήματος



χρησιμοποιώντας το εσωτερικό γινόμενο αποδείξτε το Θεώρημα του Απολλωνίου:

$$\alpha^2 + \beta^2 = \frac{\gamma^2}{2} + 2\delta^2,$$

όπου δ είναι το μήκος της διαμέσου.

4. Σε πραγματικό χώρο X με εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$ και επαγόμενη νόρμα $\|\cdot\|$ δείξτε ότι για $\boldsymbol{u}, \boldsymbol{v} \in X$ ισχύει

   (α΄) $\|a\boldsymbol{u} + b\boldsymbol{v}\| = \|b\boldsymbol{u} + a\boldsymbol{v}\|$ για όλα τα $a, b \in \mathbb{R}$ αν και μόνο αν $\|\boldsymbol{u}\| = \|\boldsymbol{v}\|$.

   (β΄) $\|\boldsymbol{u}\| \leq \|\boldsymbol{u} + a\boldsymbol{v}\|$ για όλα τα $a \in \mathbb{R}$ αν και μόνο αν $\langle \boldsymbol{u}, \boldsymbol{v}\rangle = 0$.

5. Έστω X ένας πραγματικός χώρος με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$ και επαγόμενη νόρμα $\|\cdot\|$. Αν $\boldsymbol{u},\boldsymbol{v}\in X$ και $\|\boldsymbol{u}\|=\|\boldsymbol{v}\|$, δείξτε ότι τα διανύσματα $\boldsymbol{u}+\boldsymbol{v}$ και $\boldsymbol{u}-\boldsymbol{v}$ είναι ορθογώνια. Ισχύει το συμπέρασμα αν ο χώρος είναι μιγαδικός;

6. Έστω X ένας πραγματικός χώρος με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$ και επαγόμενη νόρμα $\|\cdot\|$. Αν τα διανύσματα $\boldsymbol{u},\boldsymbol{v}\in X$ είναι ορθογώνια τότε $\|\boldsymbol{u}+\boldsymbol{v}\|^2=\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2$.

   (α΄) Δείξτε ότι ισχύει και το αντίστροφο, δηλαδή αν $\|\boldsymbol{u}+\boldsymbol{v}\|^2=\|\boldsymbol{u}\|^2+\|\boldsymbol{v}\|^2$, τότε $\boldsymbol{u}\perp\boldsymbol{v}$.

   (β΄) Εξετάστε αν το αντίστροφο ισχύει αν ο χώρος είναι μιγαδικός.

   Υπόδειξη: Θεωρήστε το $\mathbb{C}^2$ με εσωτερικό γινόμενο όπως στην (5.4) και

$$\boldsymbol{u}=\begin{pmatrix} i+1\\ 1-i\end{pmatrix},\qquad \boldsymbol{v}=\begin{pmatrix} i-1\\ 1+i\end{pmatrix}.$$

7. Εάν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα στον $\mathbb{R}^n$ με $\|\boldsymbol{u}\|=3$ και $\|\boldsymbol{v}\|=2$ να βρεθεί η μέγιστη και η ελάχιστη τιμή της $\|\boldsymbol{u}-\boldsymbol{v}\|$.

8. **Η νόρμα Frobenius.** Εάν A είναι ένα $n\times n$ μητρώο, ορίζουμε το **ίχνος** (trace) του A να είναι το άθροισμα των στοιχείων της κύριας διαγωνίου του A. Το συμβολίζουμε με $\operatorname{tr}A$, έτσι

$$\operatorname{tr}A=a_{11}+a_{22}+\cdots+a_{nn}.$$

   (α΄) Δείξτε ότι στον χώρο $\mathbb{M}^{n,m}(\mathbb{R})$ η σχέση

$$\langle A,B\rangle:=\operatorname{tr}(A^{\mathrm{T}}B)$$

   είναι ένα εσωτερικό γινόμενο. Την επαγόμενη νόρμα ονομάζουμε **νόρμα Frobenius**.

   (β΄) Βρείτε μια αναλυτική έκφραση για την νόρμα Frobenius $\|\cdot\|_F$ του μητρώου $A=\left(a_{ij}\right)\in\mathbb{M}^{2,3}$.

9. **Διάζευξη του Fredholm**: Ακριβώς ένα από τα δύο προβλήματα έχει λύση

   Π1: $A\boldsymbol{x}=\boldsymbol{b}$

   Π2: $A^{\mathrm{T}}\boldsymbol{y}=\boldsymbol{0}$ με $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b}\neq 0$.

   Ισοδύναμη διατύπωση: Το σύστημα $A\boldsymbol{x}=\boldsymbol{b}$ έχει λύση ($\boldsymbol{b}\in\operatorname{range}A$) αν και μόνο αν για κάθε $\boldsymbol{y}$ με $A^{\mathrm{T}}\boldsymbol{y}=\boldsymbol{0}$ ισχύει $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b}=0$ ($\boldsymbol{b}\in(\operatorname{null}A^{\mathrm{T}})^{\perp}$).

10. Εάν

$$S=\left\{\begin{pmatrix} x\\ y\\ x\\ y\end{pmatrix}: x,y\in\mathbb{R}\right\},$$

   να βρεθεί το ορθογώνιο συμπλήρωμα $S^{\perp}$ καθώς και μια βάση του $S^{\perp}$.

11. Αν $\boldsymbol{q}=\begin{pmatrix}\cos\theta & \sin\theta\end{pmatrix}^{\mathrm{T}}$ είναι ένα μοναδιαίο διάνυσμα δείξτε ότι το μητρώο $Q=I-2\boldsymbol{q}\boldsymbol{q}^{\mathrm{T}}$

   (α΄) Είναι ορθογώνιο.

   (β΄) Ικανοποιεί τη σχέση $Q^2=I$.

(γ′) Ανακλά κάθε διάνυσμα ως προς την ευθεία η οποία είναι ορθογώνια στο $\boldsymbol{q}$, δηλαδή κατά μήκος του $\boldsymbol{q}^{\perp} = \begin{pmatrix} -\cos\theta & \sin\theta \end{pmatrix}^{\mathrm{T}}$.

12. Δείξτε ότι το μητρώο

$$Q = \begin{pmatrix} \cos\theta & -\sin\theta & 0 \\ \sin\theta & \cos\theta & 0 \\ 0 & 0 & 1 \end{pmatrix}$$

(α′) Είναι ορθογώνιο.

(β′) Περιστρέφει κάθε διάνυσμα στο $\mathbb{R}^3$ γύρω από τον z-άξονα κατά γωνία θ.

13. Να βρεθούν, αν υπάρχουν τέτοιες, οι τιμές των a, b για τις οποίες το μητρώο

$$Q = \frac{1}{3}\begin{pmatrix} a & a & 1 \\ b & 1 & b \\ 1 & a & a \end{pmatrix}$$

είναι ορθογώνιο.

14. Εάν Q_1 και Q_2 είναι $n \times n$ ορθογώνια μητρώα δείξτε ότι το γινόμενο $Q_1 Q_2$ είναι επίσης ορθογώνιο μητρώο.

5.8 Παράρτημα VI: Ορθογώνια πολυώνυμα

Στο χώρο των πολυωνύμων $\mathbb{P}[-1,1]$ με πραγματικούς συντελεστές, η σχέση

$$\langle p, q \rangle = \int_{-1}^{1} p(x) q(x)\,\mathrm{d}x$$

ορίζει ένα εσωτερικό γινόμενο. Ενώ, όπως έχουμε δει, βλέπε Παράδειγμα 5.4, τα πολυώνυμα της φυσιολογικής βάσης $\mathcal{B} = \{1, x, x^2, x^3, \dots\}$ δεν είναι ορθογώνια μεταξύ τους η διαδικασία οθοκανονικοποίησης Gram-Schmidt παράγει μια ακολουθία από ορθογώνια πολυώνυμα. Στη πράξη και σε πολλές εφαρμογές στα Μαθηματικά και στη Φυσική εμφανίζονται διάφορες ακολουθίες ορθογωνίων πολυωνύμων, τα οποία προκύπτουν από την $\mathcal{B}$ με τη διαδικασία **ορθογωνοποίησης** και όχι ορθοκανονικοποίησης, ως προς κατάλληλο εσωτερικό γινόμενο

$$\langle p, q \rangle = \int_{a}^{b} p(x) q(x) w(x)\,\mathrm{d}x$$

όπου $-\infty \leq a < b \leq +\infty$ και $w(x) > 0$ και ολοκληρώσιμη στο σχετικό διάστημα. Στη συνέχεια περιγράφουμε τη διαδικασία ορθογωνοποίησης ως προς κάποιο εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$. Θέτουμε $\mathcal{B} = \{p_0, p_1, p_2, p_3, \dots\}$ και ορίζουμε $q_0 = p_0 = 1$. Έχοντας κατασκευάσει τα ορθογώνια μεταξύ τους $q_0, q_1, \dots, q_k$ στο επόμενο βήμα γράφουμε, κατασκευάζοντας έμμεσα το q_{k+1}, το ανάπτυγμα του p_{k+1} ως

$$p_{k+1} = \mu_1 q_1 + \mu_2 q_2 + \cdots + \mu_k q_k + q_{k+1}. \tag{5.20}$$

Για τον υπολογισμό των συντελεστών υπολογίζουμε για $j = 1, 2, \dots, k$

$$\begin{aligned} \langle p_{k+1}, q_j \rangle &= \mu_1 \langle q_1, q_j \rangle + \mu_2 \langle q_2, q_j \rangle + \cdots + \langle q_{k+1}, q_j \rangle \\ &= \mu_j \langle q_j, q_j \rangle \end{aligned}$$

από ορθογωνιότητα, οπότε $\mu_j = \langle p_{k+1}, q_j\rangle / \langle q_j, q_j\rangle$. Έτσι έχουμε

$$q_{k+1} = p_{k+1} - \frac{\langle p_{k+1}, q_1\rangle}{\langle q_1, q_1\rangle} q_1 - \frac{\langle p_{k+1}, q_2\rangle}{\langle q_2, q_2\rangle} q_2 - \cdots - \frac{\langle p_{k+1}, q_k\rangle}{\langle q_k, q_k\rangle} q_k. \tag{5.21}$$

Αν τα $q_0, q_1, \ldots, q_k$ ήταν κανονικοποιημένα οι παρονομαστές στην (5.21) θα ήταν ίσοι με ένα και η έκφραση του q_{k+1} θα ήταν η ίδια με αυτή που βρήκαμε με τη διαδικασία όπως στην διαδικασία Gram-Schmidt πριν την κανονικοποίηση.

Στη συνέχεια δίνουμε μερικά χακτηριστικά παραδείγματα βάσεων-συστημάτων ορθογωνίων πολυωνύμων.

Πολυώνυμα Legendre

Παράγονται από την $\mathcal{B}$ με τη διαδικασία ορθογωνοποίησης ως προς το εσωτερικό γινόμενο

$$\langle p, q\rangle = \int_{-1}^{1} p(x) q(x)\, \mathrm{d}x.$$

Συμβολίζονται με $P_n(x)$, $n = 0, 1, 2, \ldots$. Τα πρώτα πέντε τέτοια είναι

1. $P_0(x) = 1$
2. $P_1(x) = x$
3. $P_2(x) = \frac{1}{2}(3x^2 - 1)$
4. $P_3(x) = \frac{1}{2}(5x^3 - 3x)$
5. $P_4(x) = \frac{1}{8}(35x^4 - 30x^2 + 3)$

Αποδεικνύεται ότι τα πολυώνυμα Legendre δίνονται από την αναδρομική σχέση

$$P_0(x) = 1, \qquad P_1(x) = x$$
$$(n+1) P_{n+1}(x) = (2n+1) x P_n(x) - n P_{n-1}(x), \quad n = 1, 2, 3, \ldots$$

Κάποιες χαρακτηριστικές ιδιότητες

$$\int_{-1}^{1} P_n(x) P_m(x)\, \mathrm{d}x = \frac{2}{2n+1} \delta_{nm}$$
$$P_n(-x) = (-1)^n P_n(x), \qquad \int_{-1}^{1} P_n(x)\, \mathrm{d}x = 0, \quad n \geq 1.$$

Πολυώνυμα Chebyshev πρώτου είδους

Παράγονται από την $\mathcal{B}$ με τη διαδικασία ορθογωνοποίησης ως προς το εσωτερικό γινόμενο

$$\langle p, q\rangle = \int_{-1}^{1} p(x) q(x) \frac{1}{\sqrt{1 - x^2}}\, \mathrm{d}x.$$

Συμβολίζονται με $T_n(x)$, $n = 0, 1, 2, \ldots$. Τα πρώτα πέντε πολυώνυμα Chebyshev είναι

1. $T_0(x) = 1$
2. $T_1(x) = x$
3. $T_2(x) = 2x^2 - 1$
4. $T_3(x) = 4x^3 - 3x$
5. $T_4(x) = 8x^4 - 8x^2 + 1$

Αποδεικνύεται ότι τα πολυώνυμα Chebyshev πρώτου είδους δίνονται από την αναδρομική σχέση

$$T_0(x) = 1, \qquad T_1(x) = x$$
$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n = 1, 2, 3, \dots$$

Κάποιες χαρακτηριστικές ιδιότητες

$$\int_{-1}^{1} T_n(x)T_m(x)\frac{1}{\sqrt{1-x^2}}\,\mathrm{d}x = \begin{cases} 0 & n \neq m \\ \pi/2 & n = m \neq 0 \\ \pi & n = m = 0 \end{cases}$$
$$T_n(-x) = (-1)^n T_n(x), \qquad \int_{-1}^{1} T_n(x)\,\mathrm{d}x = \frac{(-1)^n + 1}{1 - n^2}, \quad n \neq 1.$$

Πολυώνυμα Chebyshev δεύτερου είδους

Παράγονται από την $\mathcal{B}$ με τη διαδικασία ορθογωνοποίησης ως προς το εσωτερικό γινόμενο

$$\langle p, q\rangle = \int_{-1}^{1} p(x)q(x)\sqrt{1-x^2}\,\mathrm{d}x.$$

Συμβολίζονται με $U_n(x)$, $n = 0, 1, 2, \dots$. Τα πρώτα πέντε πολυώνυμα Chebyshev είναι

1. $U_0(x) = 1$
2. $U_1(x) = 2x$
3. $U_2(x) = 4x^2 - 1$
4. $U_3(x) = 8x^3 - 4x$
5. $U_4(x) = 16x^4 - 12x^2 + 1$

Αποδεικνύεται ότι τα πολυώνυμα Chebyshev δεύτερου είδους δίνονται από την αναδρομική σχέση

$$U_0(x) = 1, \qquad U_1(x) = 2x$$
$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), \quad n = 1, 2, 3, \dots$$

Κάποιες χαρακτηριστικές ιδιότητες

$$\int_{-1}^{1} U_n(x)U_m(x)\sqrt{1-x^2}\,\mathrm{d}x = \frac{\pi}{2}\delta_{nm}$$
$$U_n(-x) = (-1)^n U_n(x), \qquad \int_{-1}^{1} U_n(x)\,\mathrm{d}x = \dots, \quad n > 1.$$

Πολυώνυμα Laguerre

Παράγονται από την $\mathcal{B}$ με τη διαδικασία ορθογωνοποίησης ως προς το εσωτερικό γινόμενο

$$\langle p, q\rangle = \int_{0}^{+\infty} p(x)q(x)e^{-x}\,\mathrm{d}x.$$

Συμβολίζονται με $L_n(x)$, $n = 0, 1, 2, \dots$. Τα πρώτα πέντε πολυώνυμα Laguerre είναι

1. $L_0(x) = 1$

2. $L_1(x) = -x + 1$
3. $L_2(x) = \frac{1}{2!}(x^2 - 4x + 2)$
4. $L_3(x) = \frac{1}{3!}(-x^3 + 9x^2 - 18x + 6)$
5. $L_4(x) = \frac{1}{4!}(x^4 - 16x^3 + 72x^2 - 96x + 24)$

Αποδεικνύεται ότι τα πολυώνυμα Laguerre δίνονται από την αναδρομική σχέση

$$L_0(x) = 1, \qquad L_1(x) = -x + 1$$
$$(n+1)L_{n+1}(x) = (2n + 1 - x)L_n(x) - nL_{n-1}(x), \quad n = 1, 2, 3, \ldots$$

Κάποιες χαρακτηριστικές ιδιότητες

$$\int_0^{+\infty} L_n(x)L_m(x)e^{-x}\,\mathrm{d}x = \delta_{nm}$$
$$L_n(x) = \sum_{k=0}^{n} \binom{n}{k}\frac{(-1)^k}{k!}x^k, \qquad \text{όπου} \qquad \binom{n}{k} = \frac{n!}{k!(n-k)!}$$

είναι ο διωνυμικός συντελεστής.

Πολυώνυμα Hermite

Παράγονται από την $\mathcal{B}$ με τη διαδικασία ορθογωνοποίησης ως προς το εσωτερικό γινόμενο

$$\langle p, q\rangle = \int_{-\infty}^{+\infty} p(x)q(x)e^{-x^2}\,\mathrm{d}x.$$

Συμβολίζονται με $H_n(x)$, $n = 0, 1, 2, \ldots$. Τα πρώτα πέντε πολυώνυμα Hermite είναι

1. $H_0(x) = 1$
2. $H_1(x) = 2x$
3. $H_2(x) = 4x^2 - 2$
4. $H_3(x) = 8x^3 - 12x + 6)$
5. $H_4(x) = 16x^4 - 48x^2 + 12$

Αποδεικνύεται ότι τα πολυώνυμα Hermite δίνονται από την αναδρομική σχέση

$$H_0(x) = 1, \qquad H_1(x) = 2x$$
$$H_{n+1}(x) = 2xH_n(x) - 2nH_{n-1}(x), \quad n = 1, 2, 3, \ldots$$

και η σχέση ορθογωνιότητας είναι

$$\int_{-\infty}^{+\infty} H_n(x)H_m(x)e^{-x^2}\,\mathrm{d}x = n!2^n\sqrt{\pi}\delta_{nm}.$$

Κεφάλαιο 6

Υπόχωροι παραγόμενοι από μητρώο

6.1 Τάξη μητρώου

Αν το A είναι ένα $n \times m$ μητρώο τότε οι οδηγοί του A καθορίζουν τις γραμμικά ανεξάρτητες στήλες και τις γραμμικά ανεξάρτητες γραμμές του A. Συγκεκριμένα οι γραμμικά ανεξάρτητες στήλες του A και οι γραμμικά ανεξάρτητες γραμμές του A είναι αντίστοιχα οι στήλες που περιέχουν τους οδηγούς και οι γραμμές που περιέχουν τους οδηγούς, κατά συνέπεια το πλήθος των γραμμικά ανεξαρτήτων στηλών του A είναι ίσο με το πλήθος των γραμμικά ανεξαρτήτων γραμμών του A.

Ορισμός 6.1. Εάν A είναι ένα $n \times m$ μητρώο ορίζουμε την **τάξη** (rank) του A να είναι το πλήθος των οδηγών του A. Συμβολίζουμε την τάξη του A με $\operatorname{rank} A$.

Άμεση συνέπεια του ορισμού είναι η σχέση $\operatorname{rank} A \le \min\{n, m\}$ αφού το μητρώο δεν μπορεί να περιέχει περισσότερους οδηγούς απ΄ ότι το πλήθος των γραμμών του ή το πλήθος των στηλών του. Επιπλέον η τάξη τετραγωνικού μητρώου σχετίζεται με την ύπαρξη του αντίστροφου, όπως δείχνει το

Θεώρημα 6.1. *Εάν A είναι ένα $n \times n$ μητρώο, τα παρακάτω είναι ισοδύναμα*

(1) *Το A είναι αντιστρέψιμο.*

(2) *Οι στήλες του A είναι γραμμικά ανεξάρτητες.*

(3) $\operatorname{rank} A = n$.

Απόδειξη. Αποδεικνύουμε τις συνεπαγωγές (1) ⇒ (2) ⇒ (3) ⇒ (1).

(1) ⇒ (2) Αν το A^{-1} υπάρχει, τότε για κάθε $\boldsymbol{b} \in \mathbb{K}^n$ το σύστημα $A\boldsymbol{x} = \boldsymbol{b}$ έχει λύση, $\boldsymbol{x} = A^{-1}\boldsymbol{b}$, ισοδύναμα κάθε $\boldsymbol{b} \in \mathbb{K}^n$ εκφράζεται ως γραμμικός συνδυασμός των στηλών του A, κατά συνέπεια οι n το πλήθος στήλες του A αποτελούν μια βάση για τον $\mathbb{K}^n$, επομένως είναι γραμμικά ανεξάρτητες.

(2) ⇒ (3) Αν οι στήλες του A είναι γραμμικά ανεξάρτητες, τότε το πλήθος των οδηγών του A είναι n, οπότε από τον ορισμό της τάξης έχουμε $\operatorname{rank} A = n$.

(3) ⇒ (1) Αν $\operatorname{rank} A = n$, τότε το πλήθος των οδηγών του A είναι n, επομένως η ανηγμένη κλιμακωτή μορφή του A είναι το ταυτοτικό μητρώο κατά συνέπεια από το Θεώρημα 3.4 έπεται ότι το A είναι αντιστρέψιμο. □

Παρατήρηση 6.1. Το Θεώρημα 6.1 επιτρέπει μια διαφορετική απόδειξη του Θεωρήματος 3.5. Ας υποθέσουμε ότι τα $n\times n$ μητρώα A και B ικανοποιούν την σχέση $AB = I$ και ας υποθέσουμε ότι το A δεν έχει αντίστροφο. Τότε θα ήταν $\operatorname{rank} A = m < n$. Επειδή οι στήλες του γινομένου είναι γραμμικοί συνδυασμοί των στηλών του A, και επειδή

$$AB = I = \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \dots & \boldsymbol{e}_n \end{pmatrix},$$

όπου $\boldsymbol{e}_j$, $j = 1, 2, \dots, n$ είναι τα διανύσματα της κανονικής βάσης του $\mathbb{R}^n$, θα είχαμε ότι m γραμμικά ανεξάρτητα διανύσματα με $m < n$ παράγουν το $\mathbb{R}^n$, αφού κάθε στοιχείο της βάσης του είναι συνδυασμός αυτών των m διανυσμάτων. Αυτό όμως είναι άτοπο, και καταλήξαμε σε άτοπο επειδή υποθέσαμε ότι το A δεν αντιστρέφεται. Επομένως το A έχει αντίστροφο και από την $AB = I$ πολλαπλασιάζοντας από αριστερά με A^{-1} βρίσκουμε $B = A^{-1}$ γεγονός που έχει ως συνέπεια ότι και το B αντιστρέφεται. Όμοια υποθέτοντας ότι το B δεν έχει αντίστροφο το ανάλογο επιχείρημα για τις γραμμές του γινομένου οδηγεί πάλι σε άτοπο.

Ορισμός 6.2. Ένα $n\times m$ μητρώο A λέγεται **πλήρους τάξης** , αν $\operatorname{rank} A = \min\{n, m\}$.

Μητρώα τάξης ένα

Αν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα, ας πούμε στο $\mathbb{R}^3$, τότε

$$\begin{aligned}
\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} \begin{pmatrix} v_1 & v_2 & v_3 \end{pmatrix} &= \begin{pmatrix} u_1v_1 & u_1v_2 & u_1v_3 \\ u_2v_1 & u_2v_2 & u_2v_3 \\ u_3v_1 & u_3v_2 & u_3v_3 \end{pmatrix} \\
&= \begin{pmatrix} v_1\boldsymbol{u} & v_2\boldsymbol{u} & v_3\boldsymbol{u} \end{pmatrix} \\
&= \begin{pmatrix} u_1\boldsymbol{v}^{\mathrm{T}} \\ u_2\boldsymbol{v}^{\mathrm{T}} \\ u_3\boldsymbol{v}^{\mathrm{T}} \end{pmatrix},
\end{aligned}$$

δηλαδή $\operatorname{rank} \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}} = 1$. Όμοια αν $\boldsymbol{u} \in \mathbb{R}^n$ και $\boldsymbol{v} \in \mathbb{R}^m$, τότε το μητρώο $\boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$ έχει τάξη ένα. Το αυτό ισχύει και για το μητρώο $\boldsymbol{v}\boldsymbol{u}^{\mathrm{T}}$. Ισχύει και το αντίστροφο, δηλαδή αν το $n\times m$ μητρώο A είναι τάξης ένα, τότε υπάρχουν διανύσματα $\boldsymbol{u} \in \mathbb{R}^n$ και $\boldsymbol{v} \in \mathbb{R}^m$ ώστε $A = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$. Πράγματι αν $\operatorname{rank} A = 1$, τότε το A έχει μία μόνο γραμμικά ανεξάρτητη στήλη, έτσι αν $\boldsymbol{a}_k$, $k = 1, 2, \dots, m$ είναι οι στήλες του A, τότε $\boldsymbol{a}_k = \lambda_k\boldsymbol{u}$, για κάποιο διάνυσμα $\boldsymbol{u} \in \mathbb{R}^n$ και σταθερά λ_k, έτσι

$$A = \begin{pmatrix} \lambda_1\boldsymbol{u} & \lambda_2\boldsymbol{u} & \cdots & \lambda_m\boldsymbol{u} \end{pmatrix} = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}},$$

όπου $\boldsymbol{v}^{\mathrm{T}} = \begin{pmatrix} \lambda_1 & \lambda_2 & \cdots & \lambda_m \end{pmatrix}$.

Παρατήρηση 6.2. Από τον ορισμό του γινομένου μητρώων έχουμε, για παράδειγμα,

$$\begin{aligned}
\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} w & x \\ y & z \end{pmatrix} &= \begin{pmatrix} aw + by & ax + bz \\ cw + dy & cx + dz \end{pmatrix} \\
&= \begin{pmatrix} aw & ax \\ cw & cx \end{pmatrix} + \begin{pmatrix} by & bz \\ dy & dz \end{pmatrix},
\end{aligned}$$

έτσι

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} w & x \\ y & z \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix} \begin{pmatrix} w & x \end{pmatrix} + \begin{pmatrix} b \\ d \end{pmatrix} \begin{pmatrix} y & z \end{pmatrix}. \tag{6.1}$$

Κάθε μητρώο στο δεξί μέλος της (6.1) είναι τάξης ένα. Τα αποτελέσματα αυτά ισχύουν προφανώς για γινόμενα μητρώων γενικότερων συμβατών διαστάσεων. Έτσι αν το $A = \left(a_{ij}\right)$ είναι ένα $n \times m$ μητρώο, το $B = \left(b_{ij}\right)$ είναι ένα $m \times k$ μητρώο και για το τυπικό μητρώο $M \in \mathbb{M}^{n,m}$ με M_{i*} συμβολίσουμε, ως συνήθως, την i-γραμμή του ως διάνυσμα-στήλη, και με M_{*j} την j-στήλη του, τότε

$$AB = A\begin{pmatrix} B_{*1} & B_{*2} & \dots & B_{*k} \end{pmatrix} = \begin{pmatrix} AB_{*1} & AB_{*2} & \dots & AB_{*k} \end{pmatrix},$$

κατά συνέπεια $(AB)_{*j} = AB_{*j}$, ισοδύναμα

$$(AB)_{*j} = A_{*1}b_{1j} + A_{*2}b_{2j} + \cdots + A_{*m}b_{mj}, \tag{6.2}$$

με $j = 1, 2, \dots, k$. Για τον υπολογισμό του $(AB)_{i*}$, παρατηρούμε ότι

$$(AB)_{i*} = ((AB)^{\mathrm{T}})_{*i} = (B^{\mathrm{T}}A^{\mathrm{T}})_{*i} = B^{\mathrm{T}}(A^{\mathrm{T}})_{*i} = B^{\mathrm{T}}A_{i*}$$

κατά συνέπεια

$$(AB)_{i*} = a_{i1}B_{1*} + a_{i2}B_{2*} + \cdots + a_{im}B_{m*}, \tag{6.3}$$

με $i = 1, 2, \dots, n$. Η γενίκευση της (6.1) είναι

$$AB = A_{*1}B_{1*}^{\mathrm{T}} + A_{*2}B_{2*}^{\mathrm{T}} + \cdots + A_{*m}B_{m*}^{\mathrm{T}}$$

από την οποία έπεται ότι το γινόμενο δύο μητρώων διασπάται σε ένα άθροισμα μητρώων τάξης ένα.

6.2 Η δομή της λύσης συστήματος

Ας θεωρήσουμε το σύστημα $Ax = \boldsymbol{b}$, όπου A είναι ένα $n \times m$ μητρώο και $\boldsymbol{b}$ είναι ένα διάνυσμα στο $\mathbb{R}^n$. Στόχος μας είναι να μελετήσουμε τη δομή της λύσης, ή των λύσεων του συστήματος, εφόσον υπάρχουν τέτοιες. Ας μελετήσουμε λοιπόν ένα τυπικό παράδειγμα.

Παράδειγμα 6.1. Θεωρούμε το σύστημα $A\boldsymbol{x} = \boldsymbol{b}$ όπου

$$A = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 1 & 0 & -1 & 1 \\ 2 & 1 & 1 & 3 \end{pmatrix} \qquad \text{και} \qquad \boldsymbol{b} = \begin{pmatrix} 3 \\ 4 \\ b \end{pmatrix},$$

όπου b είναι μια παράμετρος. Ας δούμε αρχικά για ποια ή ποιες τιμές της παραμέτρου το σύστημα έχει λύση.

Διαμορφώνοντας το επαυξημένο μητρώο $\begin{pmatrix} A & \boldsymbol{b} \end{pmatrix}$ του συστήματος, με απαλοιφή βρίσκουμε

$$\begin{pmatrix} 1 & 1 & 2 & 2 & 3 \\ 1 & 0 & -1 & 1 & 4 \\ 2 & 1 & 1 & 3 & b \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 2 & 2 & 3 \\ 0 & -1 & -3 & -1 & 1 \\ 0 & -1 & -3 & -1 & b-6 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 & 1 & 4 \\ 0 & 1 & 3 & 1 & -1 \\ 0 & 0 & 0 & 0 & b-7 \end{pmatrix} \tag{6.4}$$

κατά συνέπεια το σύστημα έχει λύση αν και μόνον αν $b = 7$. Παρατηρούμε ότι η τάξη του μητρώου A είναι δύο, δηλαδή $\operatorname{rank} A = 2$, κατά συνέπεια δύο από τους αγνώστους προσδιορίζονται μέσω των υπολοίπων $4 - \operatorname{rank} A = 2$. Πράγματι για την τιμή $b = 2$ της παραμέτρου επιλύοντας το αντίστοιχο σύστημα βρίσκουμε

$$\left.\begin{aligned} x_1 - x_3 + x_4 &= 4 \\ x_2 + 3x_3 + x_4 &= -1 \end{aligned}\right\} \Leftrightarrow \begin{cases} x_1 = x_3 - x_4 + 4 \\ x_2 = -3x_3 - x_4 - 1 \end{cases}$$

Έτσι θέτοντας $x_3 = s$ και $x_4 = t$ ή λύση του συστήματος είναι

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} s - t + 4 \\ -3s - t - 1 \\ s \\ t \end{pmatrix} = s \begin{pmatrix} 1 \\ -3 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 4 \\ -1 \\ 0 \\ 0 \end{pmatrix}. \tag{6.5}$$

Όλες λοιπόν οι λύσεις του $A\boldsymbol{x} = \boldsymbol{b}$ δίνονται από την (6.5), με s και t να παίρνουν πραγματικές τιμές. Ειδικά για $s = t = 0$ η

$$\boldsymbol{x}_p = \begin{pmatrix} 4 \\ -1 \\ 0 \\ 0 \end{pmatrix}$$

είναι λύση του συστήματος. Τότε όμως θέτοντας

$$\boldsymbol{x}_0 = s \begin{pmatrix} 1 \\ -3 \\ 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix}, \tag{6.6}$$

θα έχουμε, ως αποτέλεσμα της γραμμικότητας,

$$A(\boldsymbol{x}_0 + \boldsymbol{x}_p) = \boldsymbol{b} \Leftrightarrow A\boldsymbol{x}_0 + A\boldsymbol{x}_p = \boldsymbol{b} \Leftrightarrow A\boldsymbol{x}_0 + \boldsymbol{b} = \boldsymbol{b} \Leftrightarrow A\boldsymbol{x}_0 = \mathbf{0},$$

ισοδύναμα η $\boldsymbol{x}_0$ είναι λύση του ομοιογενούς συστήματος. Πράγματι από την (6.4) για $\boldsymbol{b} = \mathbf{0}$ βρίσκουμε ότι η λύση του ομοιογενούς συστήματος $A\boldsymbol{x} = \mathbf{0}$ είναι η $\boldsymbol{x}_0$, ισοδύναμα για όλες τις τιμές των s και t το αντίστοιχο διάνυσμα (6.6) είναι λύση του ομοιογενούς συστήματος, ισοδύναμα οι άπειρες το πλήθος λύσεις του ομοιογενούς συστήματος είναι το

$$\operatorname{span}\left\{ \begin{pmatrix} 1 \\ -3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}.$$

Παρατήρηση 6.3. Ας υποθέσουμε ότι οι $\boldsymbol{x}_1$ και $\boldsymbol{x}_2$ είναι λύσεις του $A\boldsymbol{x} = \mathbf{0}$ και $\lambda, \mu \in \mathbb{R}$, τότε

$$A(\lambda \boldsymbol{x}_1 + \mu \boldsymbol{x}_2) = \lambda A\boldsymbol{x}_1 + \mu A\boldsymbol{x}_2 = \lambda \mathbf{0} + \mu \mathbf{0} = \mathbf{0},$$

κατά συνέπεια οι λύσεις του $A\boldsymbol{x} = \mathbf{0}$ αποτελούν διανυσματικό υπόχωρο του $\mathbb{R}^m$, όπως δείχνει εξάλλου και το Παράδειγμα 6.1. Το χώρο λύσεων του ομοιογενούς συστήματος $A\boldsymbol{x} = \mathbf{0}$ θα τον ορίσουμε παρακάτω ως **μηδενόχωρο** του μητρώου A.

Εάν $\{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_k\}$ είναι μια βάση για τον χώρο λύσεων του ομοιογενούς συστήματος $A\boldsymbol{x} = \mathbf{0}$, τότε κάθε λύση του $A\boldsymbol{x} = \mathbf{0}$ είναι της μορφής

$$c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 + \cdots + c_k\boldsymbol{x}_k. \tag{6.7}$$

Ονομάζουμε την (6.7) **γενική λύση του ομοιογενούς συστήματος**. Ο χαρακτηρισμός γενική λύση δικαιολογείται από το γεγονός ότι αν $\boldsymbol{x}'$ είναι μια λύση του $A\boldsymbol{x} = \mathbf{0}$, τότε η $\boldsymbol{x}'$ προκύπτει από την (6.7) με κατάλληλη επιλογή των c_k. Ας υποθέσουμε τώρα ότι η $\boldsymbol{x}_p$ είναι μια λύση του $A\boldsymbol{x} = \boldsymbol{b}$, μια όπως λέμε **ειδική λύση** (particular solution). Αν $\boldsymbol{x}$ είναι μια επίσης λύση του $A\boldsymbol{x} = \boldsymbol{b}$, τότε η $\boldsymbol{x} - \boldsymbol{x}_p$ είναι λύση του ομοιογενούς προβλήματος. Πράγματι

$$A(\boldsymbol{x} - \boldsymbol{x}_p) = A\boldsymbol{x} - A\boldsymbol{x}_p = \boldsymbol{b} - \boldsymbol{b} = \mathbf{0},$$

κατά συνέπεια υπάρχουν σταθερές $c_1, c_2, \ldots, c_k$ ώστε

$$\boldsymbol{x} - \boldsymbol{x}_p = c_1\boldsymbol{x}_1 + c_2\boldsymbol{x}_2 \cdots + c_k\boldsymbol{x}_k = \boldsymbol{x}_0 \Rightarrow \boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{x}_p.$$

Την έκφραση $\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{x}_p$ ορίζουμε ως **γενική λύση** του συστήματος, σε αναλογία της γενικής λύσης του ομοιογενούς, αφού κάθε λύση του $A\boldsymbol{x} = \boldsymbol{b}$ προκύπτει από αυτή την έκφραση με κατάλληλη επιλογή των c_k. Αποδείξαμε λοιπόν το

Θεώρημα 6.2. *Εάν $\boldsymbol{x}_p$ είναι μια ειδική λύση του συστήματος $A\boldsymbol{x} = \boldsymbol{b}$, τότε η γενική λύση του συστήματος είναι της μορφής*

$$\boldsymbol{x} = \boldsymbol{x}_0 + \boldsymbol{x}_p,$$

όπου $\boldsymbol{x}_0$ είναι η γενική λύση του ομοιογενούς συστήματος.

Παρατήρηση 6.4. Σε σχέση με το Παράδειγμα 6.1 παρατηρούμε ότι

$$\begin{pmatrix} 1 & 1 & 2 & 2 \\ 1 & 0 & -1 & 1 \\ 2 & 1 & 1 & 3 \end{pmatrix} \begin{pmatrix} 4 \\ -5 \\ 1 \\ 1 \end{pmatrix} = \begin{pmatrix} 3 \\ 4 \\ 7 \end{pmatrix},$$

κατά συνέπεια το $\boldsymbol{x}_p = \begin{pmatrix} 4 & -5 & 1 & 1 \end{pmatrix}^{\mathrm{T}}$ είναι μια ειδική λύση για το σύστημα $A\boldsymbol{x} = \boldsymbol{b}$, με $\boldsymbol{b} = \begin{pmatrix} 3 & 4 & 7 \end{pmatrix}^{\mathrm{T}}$. Είδαμε ότι μια βάση για τον χώρο λύσεων του $A\boldsymbol{x} = \mathbf{0}$ είναι η

$$\left\{ \begin{pmatrix} 1 \\ -3 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix} \right\}$$

επομένως, σύμφωνα με το Θεώρημα 6.2, η γενική λύση του συστήματος $A\boldsymbol{x} = \boldsymbol{b}$ είναι

$$\boldsymbol{x} = c_1 \begin{pmatrix} 1 \\ -3 \\ 1 \\ 0 \end{pmatrix} + c_2 \begin{pmatrix} -1 \\ -1 \\ 0 \\ 1 \end{pmatrix} + \begin{pmatrix} 4 \\ -5 \\ 1 \\ 1 \end{pmatrix}, \tag{6.8}$$

όπου c_1, c_2 είναι πραγματικές παράμετροι. Διατυπώνεται το ερώτημα κατά πόσον οι (6.5) και (6.8) είναι ισοδύναμες. Ένας απλός υπολογισμός δείχνει ότι για $s = \lambda + 1$ και $t = \mu + 1$ έχουμε

$$(\lambda+1)\begin{pmatrix}1\\-3\\1\\0\end{pmatrix} + (\mu+1)\begin{pmatrix}-1\\-1\\0\\1\end{pmatrix} + \begin{pmatrix}4\\-1\\0\\0\end{pmatrix} = \lambda\begin{pmatrix}1\\-3\\1\\0\end{pmatrix} + \mu\begin{pmatrix}-1\\-1\\0\\1\end{pmatrix} + \begin{pmatrix}4\\-5\\1\\1\end{pmatrix},$$

κατά συνέπεια οι δύο λύσεις είναι ισοδύναμες. Αυτό ακριβώς το αποτέλεσμα δικαιολογεί τον χαρακτηρισμό μιας τέτοιας λύσης ως γενική. Ειδικά η $\boldsymbol{x}_p = \begin{pmatrix}4 & -5 & 1 & 1\end{pmatrix}^{\mathrm{T}}$ προκύπτει από την (6.5) για $s = t = 1$.

6.3 Οι τέσσερεις υπόχωροι παραγόμενοι από μητρώο

Έστω ότι A είναι ένα $n \times m$ μητρώο και ας θεωρήσουμε το σύστημα $A\boldsymbol{x} = \boldsymbol{b}$, όπου $\boldsymbol{b} \in \mathbb{R}^n$ είναι ένα σταθερό διάνυσμα. Το ερώτημα κατά πόσον το σύστημα έχει λύση είναι ισοδύναμο με το αν υπάρχει $\boldsymbol{x} \in \mathbb{R}^m$ ώστε $A\boldsymbol{x} = \boldsymbol{b}$, ισοδύναμα αν το $\boldsymbol{b}$ είναι η εικόνα κάποιου $\boldsymbol{x} \in \mathbb{R}^m$ μέσω της απεικόνισης $T_A : \mathbb{R}^m \to \mathbb{R}^n$, η οποία ορίζεται με τη σχέση

$$T_A(\boldsymbol{x}) = A\boldsymbol{x}.$$

Κατά συνέπεια ένα $n \times m$ μητρώο μπορεί να ιδωθεί σαν μια απεικόνιση του $\mathbb{R}^m$ στο $\mathbb{R}^n$. Επιπλέον αν $\lambda, \mu \in \mathbb{R}$ και $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^m$ τότε

$$T_A(\lambda\boldsymbol{x} + \mu\boldsymbol{y}) = A(\lambda\boldsymbol{x} + \mu\boldsymbol{y}) = \lambda A\boldsymbol{x} + \mu A\boldsymbol{y} = \lambda T_A(\boldsymbol{x}) + \mu T_A(\boldsymbol{y})$$

δηλαδή η απεικόνιση είναι γραμμική. Το αποτέλεσμα αυτό είναι πολύ σημαντικό για την επίλυση συστημάτων. Σε σχέση με την απεικόνιση T το σύστημα $A\boldsymbol{x} = \boldsymbol{b}$ έχει λύση αν και μόνον αν το $\boldsymbol{b} \in \mathbb{R}^n$ είναι η εικόνα κάποιου $\boldsymbol{x} \in \mathbb{R}^m$ μέσω της T_A, δηλαδή το $\boldsymbol{b}$ περιέχεται στο πεδίο τιμών της T_A. Έτσι σε αναλογία με τις συναρτήσεις έχουμε

Ορισμός 6.3. Εάν A είναι ένα $n \times m$ μητρώο ορίζουμε την **εικόνα** (range) του A

$$\operatorname{range} A = \{A\boldsymbol{x} : \boldsymbol{x} \in \mathbb{R}^m\}.$$

Παρατήρηση 6.5. Παρατηρούμε ότι $\operatorname{range} A \subseteq \mathbb{R}^n$ και ότι ένα διάνυσμα $\boldsymbol{y} \in \mathbb{R}^n$ περιέχεται στο $\operatorname{range} A$ αν και μόνον αν υπάρχει $\boldsymbol{x} = \begin{pmatrix}x_1 & x_2 & \cdots & x_m\end{pmatrix}^{\mathrm{T}} \in \mathbb{R}^m$ ώστε $A\boldsymbol{x} = \boldsymbol{y}$, ισοδύναμα

$$\boldsymbol{y} = \boldsymbol{c}_1 x_1 + \boldsymbol{c}_2 x_2 + \cdots + \boldsymbol{c}_m x_m$$

όπου με $\boldsymbol{c}_j$ συμβολίζουμε τις στήλες του A. Κατά συνέπεια $\operatorname{range} A = \operatorname{span}\{\boldsymbol{c}_1, \boldsymbol{c}_2, \ldots, \boldsymbol{c}_m\}$. Επειδή κάθε διάνυγμα διανυσμάτων είναι διανυσματικός υπόχωρος έπεται ότι $\operatorname{range} A$ είναι διανυσματικός υπόχωρος του $\mathbb{R}^n$. Τον υπόχωρο $\operatorname{range} A$ λέμε και **χώρο στηλών** (column space) του A και ενίοτε συμβολίζουμε με $C(A)$. Σημειώνουμε ότι η διάστασή του είναι το πολύ n, δηλαδή $\dim(\operatorname{range} A) \leq n$.

Στην Παρατήρηση 6.3 είδαμε ότι οι λύσεις ενός ομοιογενούς συστήματος αποτελούν διανυσματικό χώρο.

Ορισμός 6.4. Εάν A είναι ένα $n \times m$ μητρώο ορίζουμε τον **μηδενόχωρο** (null space) του A

$$\operatorname{null} A = \{\boldsymbol{x} \in \mathbb{R}^m : A\boldsymbol{x} = \boldsymbol{0}\}.$$

Παρατήρηση 6.6. Μια άμεση συνέπεια του Θεωρήματος 3.4 είναι ότι αν το A είναι ένα τετραγωνικό μητρώο τότε οι ισχυρισμοί

(1) Το A είναι αντιστρέψιμο.

(2) $\operatorname{null} A = \{\boldsymbol{0}\}$.

είναι ισοδύναμοι. Επιπλέον αν ισχύει το (1), ή το (2) τότε

$$A\boldsymbol{x} = A\boldsymbol{y} \Rightarrow A(\boldsymbol{x} - \boldsymbol{y}) = \boldsymbol{0} \Rightarrow \boldsymbol{x} - \boldsymbol{y} = \boldsymbol{0} \Rightarrow \boldsymbol{x} = \boldsymbol{y},$$

δηλαδή ένα αντιστρέψιμο μητρώο συμπεριφέρεται όπως μια ένα-προς-ένα απεικόνιση.

Παρατήρηση 6.7 (Ορισμός). Όμοια, για το $n \times m$ μητρώο A το

$$\operatorname{range} A^{\mathrm{T}} = \{A^{\mathrm{T}}\boldsymbol{y} : \boldsymbol{y} \in \mathbb{R}^n\} = C(A^{\mathrm{T}})$$

είναι διανυσματικός υπόχωρος του $\mathbb{R}^m$, και παράγεται από τις στήλες του A^{T}, ισοδύναμα από τις γραμμές του A. Για τον λόγο αυτό τον υπόχωρο $\operatorname{range} A^{\mathrm{T}}$ λέμε και **χώρο γραμμών** (row space) του A και συμβολίζουμε με $R(A)$. Επίσης τον διανυσματικό υπόχωρο του $\mathbb{R}^n$

$$\operatorname{null} A^{\mathrm{T}} = \{\boldsymbol{y} \in \mathbb{R}^n : A^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}\},$$

λέμε **αριστερό μηδενόχωρο** του A, αφού αν $\boldsymbol{y} \in \operatorname{null} A^{\mathrm{T}}$, τότε

$$\boldsymbol{y}^{\mathrm{T}}A = (A^{\mathrm{T}}\boldsymbol{y})^{\mathrm{T}} = \boldsymbol{0}^{\mathrm{T}}.$$

Στη συνέχεια μέσω παραδειγμάτων παρουσιάζουμε ένα συστηματικό τρόπο εύρεσης των τεσσάρων υποχώρων μητρώου.

Παράδειγμα 6.2. Να βρεθεί ο χώρος στηλών και ο μηδενόχωρος του μητρώου

$$A = \begin{pmatrix} 1 & 1 & 2 & 2 \\ 1 & 0 & -1 & 1 \\ 2 & 1 & 1 & 3 \end{pmatrix}.$$

Ο χώρος στηλών $\operatorname{range} A$ ή $C(A)$ είναι υπόχωρος του $\mathbb{R}^3$ κατά συνέπεια παράγεται από το πολύ τρεις στήλες του A, επομένως οι στήλες του A είναι γραμμικά εξαρτημένες. Επειδή η ανηγμένη κλιμακωτή μορφή R_0 αναδεικνύει τη δομή του A ξεκινάμε με την διαδικασία της αναγωγής του A στο R_0.

$$A \to \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & -1 & -3 & -1 \\ 0 & -1 & -3 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 2 & 2 \\ 0 & -1 & -3 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & -1 & -3 & -1 \\ 0 & 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & -1 & 1 \\ 0 & 1 & 3 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Αν με $\boldsymbol{c}_j$ και $\boldsymbol{c}'_j$, $j = 1, 2, 3, 4$ συμβολίσουμε τις στήλες του A και R_0 αντίστοιχα, βλέπουμε ότι οι $\boldsymbol{c}'_1$ και $\boldsymbol{c}'_2$ είναι γραμμικά ανεξάρτητες ενώ

$$\boldsymbol{c}'_3 = -\boldsymbol{c}'_1 + 3\boldsymbol{c}'_2, \qquad \text{και} \qquad \boldsymbol{c}'_4 = \boldsymbol{c}'_1 + \boldsymbol{c}'_2. \tag{6.9}$$

Αν L είναι το αντιστρέψιμο μητρώο που αναγάγει το A στο R_0, από το σύστημα $A\boldsymbol{x} = \boldsymbol{b}$ γράφοντας αναλυτικά τις ισοδυναμίες

$$A\boldsymbol{x} = \boldsymbol{b} \Leftrightarrow LA\boldsymbol{x} = L\boldsymbol{b} \Leftrightarrow R_0\boldsymbol{x} = L\boldsymbol{b},$$

αφού το L^{-1} υπάρχει, παίρνουμε

$$\begin{aligned}
x_1\boldsymbol{c}_1 + x_2\boldsymbol{c}_2 + x_3\boldsymbol{c}_3 + x_4\boldsymbol{c}_4 &= \boldsymbol{b} \Rightarrow \\
L(x_1\boldsymbol{c}_1 + x_2\boldsymbol{c}_2 + x_3\boldsymbol{c}_3 + x_4\boldsymbol{c}_4) &= L\boldsymbol{b} \Rightarrow \\
x_1L\boldsymbol{c}_1 + x_2L\boldsymbol{c}_2 + x_3L\boldsymbol{c}_3 + x_4L\boldsymbol{c}_4) &= L\boldsymbol{b} \Rightarrow && \text{(από γραμμικότητα)}\\
x_1\boldsymbol{c}'_1 + x_2\boldsymbol{c}'_2 + x_3\boldsymbol{c}'_3 + x_4\boldsymbol{c}'_4 &= L\boldsymbol{b} \Rightarrow && (L\boldsymbol{c}_j = \boldsymbol{c}'_j)\\
(x_1 - x_3 + x_4)\boldsymbol{c}'_1 + (x_2 + 3x_3 + x_4)\boldsymbol{c}'_2 &= L\boldsymbol{b} \Rightarrow && \text{(από την (6.9))}\\
L^{-1}((x_1 - x_3 + x_4)\boldsymbol{c}'_1 + (x_2 + 3x_3 + x_4)\boldsymbol{c}'_2) &= L^{-1}L\boldsymbol{b} \Rightarrow \\
(x_1 - x_3 + x_4)L^{-1}\boldsymbol{c}'_1 + (x_2 + 3x_3 + x_4)L^{-1}\boldsymbol{c}'_2 &= \boldsymbol{b} \Rightarrow \\
(x_1 - x_3 + x_4)\boldsymbol{c}_1 + (x_2 + 3x_3 + x_4)\boldsymbol{c}_2 &= \boldsymbol{b} && (L^{-1}\boldsymbol{c}'_j = \boldsymbol{c}_j).
\end{aligned}$$

Κατά συνέπεια μια βάση για τον χώρο στηλών $C(A)$ ή $\operatorname{range} A$ αποτελούν οι γραμμικά ανεξάρτητες στήλες του A $\boldsymbol{c}_1$ και $\boldsymbol{c}_2$, όπως και του R_0, έτσι

$$\operatorname{range} A = \operatorname{span}\left\{\begin{pmatrix}1\\1\\2\end{pmatrix}, \begin{pmatrix}1\\0\\1\end{pmatrix}\right\} = C(A).$$

Αναζητώντας στη συνέχεια μια βάση για τον $\operatorname{null} A$, θεωρούμε το σύστημα $A\boldsymbol{x} = \boldsymbol{0}$, δηλαδή παίρνουμε $\boldsymbol{b} = \boldsymbol{0}$. Τότε

$$A\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow LA\boldsymbol{x} = L\boldsymbol{0} \Leftrightarrow R_0\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \begin{cases} x_1 - x_3 + x_4 = 0 \\ x_2 + 3x_3 + x_4 = 0 \end{cases}$$

οπότε

$$x_1 = x_3 - x_4 \qquad \text{και} \qquad x_2 = -3x_3 - x_4$$

και η λύση του συστήματος είναι

$$\boldsymbol{x} = \begin{pmatrix}x_1\\x_2\\x_3\\x_4\end{pmatrix} = \begin{pmatrix}x_3 - x_4\\-3x_3 - x_4\\x_3\\x_4\end{pmatrix} = \begin{pmatrix}x_3\\-3x_3\\x_3\\0\end{pmatrix} + \begin{pmatrix}-x_4\\-x_4\\0\\x_4\end{pmatrix} = x_3\begin{pmatrix}1\\-3\\1\\0\end{pmatrix} + x_4\begin{pmatrix}-1\\-1\\0\\1\end{pmatrix}.$$

Έτσι έχουμε, αφού τα δύο διανύσματα είναι γραμμικά ανεξάρτητα,

$$\operatorname{null} A = \operatorname{span}\left\{\begin{pmatrix}1\\-3\\1\\0\end{pmatrix}, \begin{pmatrix}-1\\-1\\0\\1\end{pmatrix}\right\}.$$

Παράδειγμα 6.3 (συνέχεια του Παραδείγματος 6.2). Να βρεθεί ο χώρος γραμμών και ο αριστερός μηδενόχωρος του μητρώου A του Παραδείγματος 6.2.

Πρώτη προσέγγιση. Βρίσκουμε την ανηγμένη κλιμακωτή μορφή του A^{T}.

$$A^{\mathrm{T}} = \begin{pmatrix} 1 & 1 & 2 \\ 1 & 0 & 1 \\ 2 & -1 & 1 \\ 2 & 1 & 3 \end{pmatrix} \to \begin{pmatrix} 1 & 1 & 2 \\ 0 & -1 & -1 \\ 0 & -3 & -3 \\ 0 & -1 & -1 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & -1 & -1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} = R_0'.$$

Δύο από τις τρεις στήλες του R_0' είναι γραμμικά ανεξάρτητες, έτσι επιλέγοντας την πρώτη και την δεύτερη παίρνουμε

$$\operatorname{range} A^{\mathrm{T}} = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right\} = C(A^{\mathrm{T}}) = R(A).$$

Για τον αριστερό μηδενόχωρο του A λύνουμε το σύστημα $A^{\mathrm{T}}\boldsymbol{x} = \mathbf{0}$, ισοδυναμα $R_0'\boldsymbol{x} = \mathbf{0}$, ή

$$\left.\begin{aligned} x_1 + x_3 &= 0 \\ x_2 + x_3 &= 0 \end{aligned}\right\} \Leftrightarrow \begin{cases} x_1 = -t \\ x_2 = -t \\ x_3 = t \end{cases} \qquad t \in \mathbb{R}$$

κατά συνέπεια

$$\operatorname{null} A^{\mathrm{T}} = \operatorname{span}\left\{ \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\}.$$

Δεύτερη προσέγγιση. Σκεφτόμαστε μήπως θα μπορούσαμε να πάρουμε τα αποτελέσματα που μας ενδιαφέρουν χωρίς να επαναλάβουμε για το A^{T} τη διαδικασία που ακολουθήθηκε για το A. Ας ξεκινήσουμε με την προφανή παρατήρηση ότι κάθε γραμμή του A είναι γραμμικός συνδυασμός των γραμμών του R_0, την ανηγμένη κλιμακωτή μορφή του A, και αντιστρόφως. Πράγματι σε κάθε βήμα της απαλοιφής μια γραμμή προκύπτει από το πρόσθεση σε αυτή πολλαπλασίου άλλης γραμμής. Έτσι αν με $\boldsymbol{r}_j$ και $\boldsymbol{r}_j'$, $j = 1, 2, 3$ συμβολίσουμε τις γραμμές του A και R_0 αντίστοιχα, βλέπουμε ότι

$$\begin{aligned} \boldsymbol{r}_1 &= \boldsymbol{r}_1' + \boldsymbol{r}_2', & \qquad \boldsymbol{r}_1' &= \boldsymbol{r}_2, \\ \boldsymbol{r}_2 &= \boldsymbol{r}_1', & \qquad \boldsymbol{r}_2' &= \boldsymbol{r}_1 - \boldsymbol{r}_2, \end{aligned}$$

κατά συνέπεια

$$\operatorname{range} A^{\mathrm{T}} = \operatorname{span}\{\boldsymbol{r}_1', \boldsymbol{r}_2'\} = \operatorname{span}\{\boldsymbol{r}_1, \boldsymbol{r}_2\} = R(A)$$

ή

$$\operatorname{range} A^{\mathrm{T}} = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 1 \\ 2 \\ 2 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \\ -1 \\ 1 \end{pmatrix} \right\} = R(A).$$

Το μητρώο L που αναγάγει το A στο R_0 είναι το

$$L = \begin{pmatrix} 1 & 1 & 0 \\ 0 & -1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ -2 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & -1 & 0 \\ -1 & -1 & 1 \end{pmatrix}$$

(γιατί;). Από την σχέση $LA = R_0$ παρατηρούμε ότι η τελευταία γραμμή του L είναι εκείνη που συνδυάζει τις γραμμές του A και δίνει την μηδενική γραμμή, τρίτη γραμμή, του R_0, έτσι η τελευταία γραμμή του L αποτελεί μια βάση του αριστερού μηδενόχωρου $\operatorname{null} A^{\mathrm{T}}$. Πράγματι

$$(LA)^{\mathrm{T}} = A^{\mathrm{T}}L^{\mathrm{T}} = R_0^{\mathrm{T}} \Leftrightarrow \left(A^{\mathrm{T}}\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} \quad A^{\mathrm{T}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix} \quad A^{\mathrm{T}}\begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ -1 & 3 & 0 \\ 1 & 1 & 0 \end{pmatrix},$$

επομένως

$$\operatorname{null} A^{\mathrm{T}} = \operatorname{span}\left\{ \begin{pmatrix} -1 \\ -1 \\ 1 \end{pmatrix} \right\},$$

όπως βρήκαμε με την πρώτη προσέγγιση.

6.3.1 Οι διαστάσεις των τεσσάρων υποχώρων

Αν $C(A)$ και $R(A)$ είναι αντίστοιχα ο χώρος στηλών και ο χώρος γραμμών μητρώου A, από τον ορισμό της τάξης του A έπεται ότι

$$\operatorname{rank} A = \dim C(A) = \dim(\operatorname{range} A). \tag{6.10}$$
$$\operatorname{rank} A = \dim R(A) = \dim(\operatorname{range} A^{\mathrm{T}}). \tag{6.11}$$

Σχετικά με τη διάσταση καθενός από τους δύο μηδενόχωρους έχουμε το

Θεώρημα 6.3 (Θεμελιώδες Θεώρημα, 1ο μέρος). *Εάν A είναι ένα $n \times m$ μητρώο τάξης r, τότε*

(1) $\dim(\operatorname{null} A) = m - r$.

(2) $\dim(\operatorname{null} A^{\mathrm{T}}) = n - r$.

Απόδειξη. (1) Έστω $k = \dim(\operatorname{null} A)$ και έστω $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k\}$ μια βάση για τον μηδενόχωρο $\operatorname{null} A$. Τότε υπάρχουν γραμμικά ανεξάρτητα διανύσματα $\boldsymbol{w}_{k+1}, \ldots, \boldsymbol{w}_m$ ώστε το σύνολο

$$\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_k, \boldsymbol{w}_{k+1}, \ldots, \boldsymbol{w}_m\}$$

είναι μια βάση για τον $\mathbb{R}^m$. Αν $\boldsymbol{x} \in \mathbb{R}^n$, από την υπόθεση έπεται ότι $\boldsymbol{x} = c_1\boldsymbol{w}_1 + \cdots + c_m\boldsymbol{w}_m$ και

$$\begin{aligned} A\boldsymbol{x} &= A(c_1\boldsymbol{w}_1 + \cdots + c_k\boldsymbol{w}_k + c_{k+1}\boldsymbol{w}_{k+1} + \cdots + c_m\boldsymbol{w}_m) \\ &= A(c_1\boldsymbol{w}_1 + \cdots + c_k\boldsymbol{w}_k) + A(c_{k+1}\boldsymbol{w}_{k+1} + \cdots + c_m\boldsymbol{w}_m) \\ &= c_{k+1}A\boldsymbol{w}_{k+1} + \cdots + c_mA\boldsymbol{w}_m \end{aligned}$$

αφού $c_1\boldsymbol{w}_1 + \cdots + c_k\boldsymbol{w}_k \in \operatorname{null} A$, επομένως $\operatorname{range} A \subseteq \operatorname{span}\{A\boldsymbol{w}_{k+1}, \ldots, A\boldsymbol{w}_m\}$. Επειδή $A\boldsymbol{w}_j \in \operatorname{span} A$, για $j = k+1, \ldots, m$ και το $\operatorname{range} A$ είναι υπόχωρος έπεται ότι $\operatorname{span}\{A\boldsymbol{w}_{k+1}, \ldots, A\boldsymbol{w}_m\} \subseteq \operatorname{range} A$, κατά συνέπεια $\operatorname{range} A = \operatorname{span}\{A\boldsymbol{w}_{k+1}, \ldots, A\boldsymbol{w}_m\}$. Ισχυριζόμαστε ότι το $\{A\boldsymbol{w}_{k+1}, \ldots, A\boldsymbol{w}_m\}$ είναι μια βάση για το $\operatorname{range} A$. Αρκεί να δείξουμε ότι τα διανύσματα $A\boldsymbol{w}_{k+1}, \ldots, A\boldsymbol{w}_m$ είναι γραμμικά ανεξάρτητα. Πράγματι διαμορφώνοντας την εξίσωση

$$c_1 A\boldsymbol{w}_{k+1} + \cdots + c_{m-k} A\boldsymbol{w}_m = \boldsymbol{0}$$

παίρνουμε

$$A(c_1\boldsymbol{w}_{k+1} + \cdots + c_{m-k}\boldsymbol{w}_m) = \boldsymbol{0} \Rightarrow c_1\boldsymbol{w}_{k+1} + \cdots + c_{m-k}\boldsymbol{w}_m \in \operatorname{null} A.$$

Επειδή $\boldsymbol{w}_{k+1}, \ldots, \boldsymbol{w}_m \notin \operatorname{null} A$ έπεται ότι $c_1 = \cdots = c_{m-k} = 0$, κατά συνέπεια τα διανύσματα $A\boldsymbol{w}_{k+1}, \ldots, A\boldsymbol{w}_m$ είναι γραμμικά ανεξάρτητα, όπως θέλαμε να δείξουμε. Έτσι α διανύσματα $A\boldsymbol{w}_{k+1}, \ldots, A\boldsymbol{w}_m$ αποτελούν μια βάση για το $\operatorname{range} A$, κατά συνέπεια μέσω της (6.10) έπεται ότι $r = m - k$, δηλαδή $k = \dim(\operatorname{null} A) = m - r$.

Για την απόδειξη του (2) εφαρμόζουμε το αποτέλεσμα (1) στο $m \times n$ μητρώο $B = A^{\mathrm{T}}$. □

Από το Θεώρημα 6.3 και την (6.11) έπεται ότι

$$\dim(\operatorname{null} A) + \dim(\operatorname{range} A^{\mathrm{T}}) = m - r + r = m = \dim \mathbb{R}^m,$$

κατά συνέπεια, αφού οι $\operatorname{null} A$ και $\operatorname{range} A^{\mathrm{T}}$ είναι υπόχωροι του $\mathbb{R}^m$, καταλήγουμε στο σημαντικό συμπέρασμα $\operatorname{null} A + \operatorname{range} A^{\mathrm{T}} = \mathbb{R}^m$. Στην πραγματικότητα ισχύει το ισχυρότερο αποτέλεσμα

Θεώρημα 6.4 (Θεμελιώδες Θεώρημα, 2ο μέρος). *Εάν A είναι ένα $n \times m$ μητρώο, τότε*

(1) $\mathbb{R}^m = \operatorname{null} A \oplus \operatorname{range} A^{\mathrm{T}}$, *δηλαδή ο $\mathbb{R}^m$ είναι το ευθύ άθροισμα των* $\operatorname{null} A$ *και* $\operatorname{range} A^{\mathrm{T}}$.

(2) $\mathbb{R}^n = \operatorname{null} A^{\mathrm{T}} \oplus \operatorname{range} A$, *δηλαδή ο $\mathbb{R}^n$ είναι το ευθύ άθροισμα των* $\operatorname{null} A^{\mathrm{T}}$ *και* $\operatorname{range} A$.

Απόδειξη. Δίνουμε την απόδειξη του (1), η απόδειξη του (2) είναι ανάλογη. Αρκεί να δείξουμε ότι $\operatorname{null} A \cap \operatorname{range} A^{\mathrm{T}} = \{\boldsymbol{0}\}$. Πράγματι αν $\boldsymbol{x} \in \operatorname{null} A \cap \operatorname{range} A^{\mathrm{T}}$, τότε αφενός $A\boldsymbol{x} = \boldsymbol{0}$ και αφετέρου υπάρχει $\boldsymbol{y} \in \mathbb{R}^m$ με $A^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{x}$. Έτσι αναστρέφοντας παίρνουμε

$$A\boldsymbol{x} = \boldsymbol{0} \qquad \text{και} \qquad (A^{\mathrm{T}}\boldsymbol{y})^{\mathrm{T}} = \boldsymbol{x}^{\mathrm{T}} \Rightarrow \boldsymbol{y}^{\mathrm{T}}A = \boldsymbol{x}^{\mathrm{T}}$$

οπότε πολλαπλασιάζοντας από αριστερά τη δεύτερη σχέση με $\boldsymbol{x}$ έχουμε

$$A\boldsymbol{x} = \boldsymbol{0} \qquad \text{και} \qquad \boldsymbol{y}^{\mathrm{T}}A\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x}$$

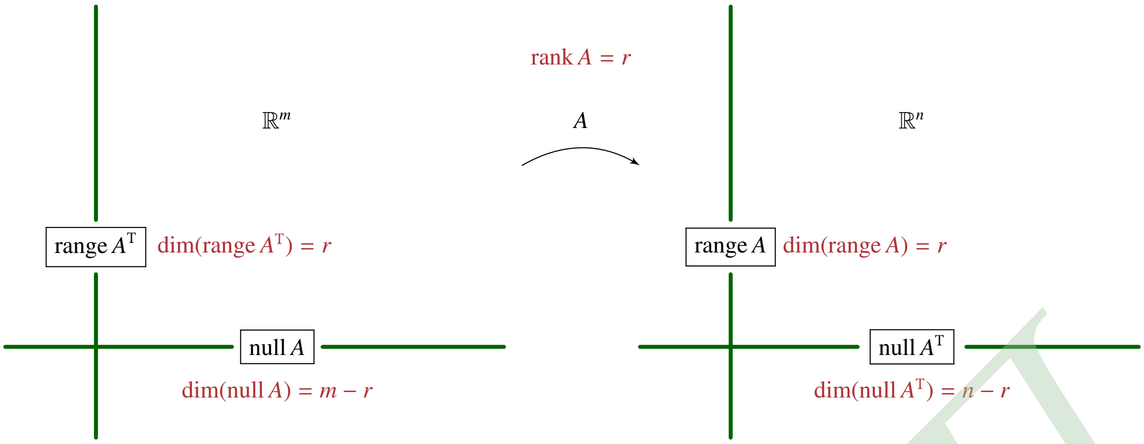
απ΄ όπου συνδυάζοντας έπεται ότι

$$\boldsymbol{y}^{\mathrm{T}}\boldsymbol{0} = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{x} \Rightarrow 0 = \|\boldsymbol{x}\|^2 \Rightarrow \boldsymbol{x} = \boldsymbol{0},$$

από τον ορισμό του μέτρου διανύσματος, που είναι ό,τι θέλαμε να αποδείξουμε. □

6.3.2 Το θεμελιώδες Θεώρημα

Στη συνέχεια δείχνουμε ότι οι υπόχωροι που αποτελούν το κάθε ευθύ άθροισμα του Θεωρήματος 6.4 είναι ορθογώνιοι. Κωδικοποιημένα μπορούμε να αποτυπώσουμε αυτή τη σχέση στο Σχήμα 6.1.



Σχήμα 6.1: Οι τέσσερεις υπόχωροι που παράγονται από μητρώο.

Εξειδικεύοντας την Παρατήρηση 5.1 εάν το A είναι ένα $n \times m$ μητρώο και το $\boldsymbol{x}$ ένα διάνυσμα στο $\mathbb{R}^m$, τότε, με το γνωστό συμβολισμό για τις γραμμές του A, έπεται ότι

$$A\boldsymbol{x} = \left(A_{i*}^{\mathrm{T}}\boldsymbol{x}\right) = \left(\langle A_{i*}, \boldsymbol{x}\rangle\right)$$

κατά συνέπεια έχουμε τις ισοδυναμίες

$$A\boldsymbol{x} = \mathbf{0} \Leftrightarrow \boldsymbol{x} \perp A_{i*}, \quad i = 1, 2, \ldots, n \Leftrightarrow \boldsymbol{x} \perp \operatorname{range} A^{\mathrm{T}}$$

επομένως

$$\boldsymbol{x} \in \operatorname{null} A \Leftrightarrow \boldsymbol{x} \in (\operatorname{range} A^{\mathrm{T}})^{\perp}$$

γεγονός το οποίο ισοδυναμεί (γιατί;) με

$$\operatorname{null} A = (\operatorname{range} A^{\mathrm{T}})^{\perp} \Leftrightarrow (\operatorname{null} A)^{\perp} = \operatorname{range} A^{\mathrm{T}}. \tag{6.12}$$

και εκφράζει ότι $\operatorname{null} A \perp \operatorname{range} A^{\mathrm{T}}$. Όμοια με A^{T} στη θέση του A βρίσκουμε τις ανάλογες σχέσεις

$$\operatorname{null} A^{\mathrm{T}} = (\operatorname{range} A)^{\perp} \Leftrightarrow (\operatorname{null} A^{\mathrm{T}})^{\perp} = \operatorname{range} A. \tag{6.13}$$

Συνδυάζοντας αυτά τα αποτελέσματα με εκείνα του Θεμελιώδους Θεωρήματος 6.4, όπου αποδείξαμε ότι εάν A είναι ένα $n \times m$ μητρώο, τότε

$$\mathbb{R}^m = \operatorname{null} A \oplus \operatorname{range} A^{\mathrm{T}}, \qquad \text{και} \qquad \mathbb{R}^n = \operatorname{null} A^{\mathrm{T}} \oplus \operatorname{range} A,$$

προκύπτει το πλήρες αποτέλεσμα

Θεώρημα 6.5 (Θεμελιώδες Θεώρημα, πλήρης μορφή). *Εάν A είναι ένα $n \times m$ μητρώο, τότε*

$$\begin{aligned} \mathbb{R}^m &= \operatorname{null} A \oplus \operatorname{range} A^{\mathrm{T}}, & \textit{και} \quad & \operatorname{null} A \perp \operatorname{range} A^{T}, \\ \mathbb{R}^n &= \operatorname{null} A^{\mathrm{T}} \oplus \operatorname{range} A, & \textit{και} \quad & \operatorname{null} A^{\mathrm{T}} \perp \operatorname{range} A. \end{aligned}$$

6.4 Η τάξη του γινομένου μητρώων

Παρατήρηση 6.8. Εάν A και B είναι μητρώα συμβατών διαστάσεων ώστε να ορίζεται το AB δημιουργείται το ερώτημα εάν υπάρχει κάποια σχέση μεταξύ των τάξεων των A, B και AB. Μια προσεκτική ανάγνωση της απόδειξης στην Παρατήρηση 6.1 αποκαλύπτει ότι η τάξη του γινομένου μητρώων AB δεν μπορεί να είναι μεγαλύτερη της τάξης του A, ή της τάξης του B. Συζητάμε αυτό το αποτέλεσμα κοιτάζοντας μια απλή περίπτωση, όπου τα A και B είναι 2×2 μητρώα. Υπολογίζοντας ή για την ακρίβεια γράφοντας

$$AB = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{pmatrix} = \begin{pmatrix} \begin{pmatrix} a_{11} & a_{12} \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix} & \begin{pmatrix} a_{11} & a_{12} \end{pmatrix} \begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix} \\ \\ \begin{pmatrix} a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{11} \\ b_{21} \end{pmatrix} & \begin{pmatrix} a_{21} & a_{22} \end{pmatrix} \begin{pmatrix} b_{12} \\ b_{22} \end{pmatrix} \end{pmatrix} = \begin{pmatrix} p & q \\ r & s \end{pmatrix},$$

βλέπουμε ότι

(1) Αν οι γραμμές του A είναι γραμμικά εξαρτημένες, έστω $\begin{pmatrix} a_{21} & a_{22} \end{pmatrix} = \lambda \begin{pmatrix} a_{11} & a_{12} \end{pmatrix}$, τότε

$$AB = \begin{pmatrix} p & q \\ \lambda p & \lambda q \end{pmatrix},$$

δηλαδή και οι γραμμές του AB είναι γραμμικά εξαρτημένες.

(2) Αν οι στήλες του B είναι γραμμικά εξαρτημένες, έστω $\begin{pmatrix} b_{12} & b_{22} \end{pmatrix}^{\mathrm{T}} = \mu \begin{pmatrix} b_{11} & b_{21} \end{pmatrix}^{\mathrm{T}}$, τότε

$$AB = \begin{pmatrix} p & \mu p \\ r & \mu r \end{pmatrix}$$

δηλαδή και οι στήλες του AB είναι γραμμικά εξαρτημένες.

Συνέπεια αυτού του αποτελέσματος είναι η ανισοτική σχέση $\operatorname{rank}(AB) \leq \min\{\operatorname{rank} A, \operatorname{rank} B\}$.

► **Άσκηση 6.1.** Εάν A και B είναι τα μητρώα της Παρατήρησης 6.8 τί παρατηρείτε για το γινόμενο AB αν (i) οι στήλες του A είναι γραμμικά εξαρτημημένες, ή αν (ii) οι γραμμές του B είναι γραμμικά εξαρτημημένες;

► **Άσκηση 6.2.** Εάν A και B είναι μητρώα συμβατών διαστάσεων ώστε να ορίζεται το AB, δείξτε ότι

$$\operatorname{rank}(AB) \leq \min\{\operatorname{rank} A, \operatorname{rank} B\},$$

δείξτε δηλαδή ότι $\operatorname{rank}(AB) \leq \operatorname{rank} A$ και $\operatorname{rank}(AB) \leq \operatorname{rank} B$.

Παράδειγμα 6.4. Ας θεωρήσουμε τα μητρώα

$$Q = \begin{pmatrix} 2 & 1 \\ 3 & -2 \\ 1 & 0 \end{pmatrix}, \qquad S = \begin{pmatrix} 1 & -1 & 2 \\ 3 & 0 & 1 \end{pmatrix}, \qquad A = QS.$$

Το A είναι 3×3, αλλά η τάξη του δεν μπορεί, σύμφωνα με την Παρατήρηση 6.8 και την Άσκηση 6.2, να είναι μεγαλύτερη του δύο, αφού $\operatorname{rank} A \leq \operatorname{rank} Q$ και $\operatorname{rank} A \leq \operatorname{rank} S$. Οι στήλες του A είναι γραμμικοί συνδυασμοί των στηλών του Q, οι οποίες είναι γραμμικά ανεξάρτητες (γιατί;), επομένως για τον χώρο στηλών του A έχουμε

$$C(A) = \operatorname{range} A = \operatorname{span}\left\{\begin{pmatrix}2\\3\\1\end{pmatrix}, \begin{pmatrix}1\\-2\\0\end{pmatrix}\right\}$$

και οι γραμμές του A είναι γραμμικοί συνδυασμοί των γραμμών του S οι οποίες είναι επίσης γραμμικά ανεξάρτητες, συνεπώς για τον χώρο γραμμών του A

$$R(A) = \operatorname{range} A^{\mathrm{T}} = \operatorname{span}\left\{\begin{pmatrix}1\\-1\\2\end{pmatrix}, \begin{pmatrix}3\\0\\1\end{pmatrix}\right\}.$$

Έτσι $\operatorname{rank} A = 2$, γεγονός που περιμέναμε αφού $\operatorname{rank} Q = \operatorname{rank} S = 2$. Από το Θεμελιώδες Θεώρημα έπεται ότι $\dim(\operatorname{null} A) = 3 - 2 = 1$. Επειδή $\operatorname{rank} Q = 2$ έχουμε ότι

$$A\boldsymbol{x} = \mathbf{0} \Leftrightarrow Q^{\mathrm{T}}QS\boldsymbol{x} = Q^{\mathrm{T}}\mathbf{0} \Leftrightarrow S\boldsymbol{x} = \mathbf{0}$$

αφού το 2×2 μητρώο $Q^{\mathrm{T}}Q$ αντιστρέφεται, κατά συνέπεια $\operatorname{null} A = \operatorname{null} S$. Με απαλοιφή βρίσκουμε

$$\begin{pmatrix}1 & -1 & 2\\3 & 0 & 1\end{pmatrix} \to \begin{pmatrix}1 & -1 & 2\\0 & 3 & -5\end{pmatrix} \to \begin{pmatrix}1 & -1 & 2\\0 & 1 & -5/3\end{pmatrix} \to \begin{pmatrix}1 & 0 & 1/3\\0 & 1 & -5/3\end{pmatrix}$$

Έτσι η λύση της $S\boldsymbol{x} = \mathbf{0}$ είναι

$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}-\frac{1}{3}x_3\\\frac{5}{3}x_3\\x_3\end{pmatrix} = \frac{x_3}{3}\begin{pmatrix}-1\\5\\3\end{pmatrix}$$

επομένως

$$\operatorname{null} A = \operatorname{span}\left\{\begin{pmatrix}-1\\5\\3\end{pmatrix}\right\}.$$

Για την εύρεση του αριστερού μηδενόχωρου του A, από την $A = QS$ παίρνουμε $A^{\mathrm{T}} = S^{\mathrm{T}}Q^{\mathrm{T}}$, και όπως πριν αφού $\operatorname{rank} S = 2$ έχουμε $A^{\mathrm{T}}\boldsymbol{x} = \mathbf{0} \Leftrightarrow Q^{\mathrm{T}}\boldsymbol{x} = \mathbf{0}$. Και πάλι με απαλοιφή βρίσκουμε

$$\begin{pmatrix}2 & 3 & 1\\1 & -2 & 0\end{pmatrix} \to \begin{pmatrix}1 & -2 & 0\\2 & 3 & 1\end{pmatrix} \to \begin{pmatrix}1 & -2 & 0\\0 & 7 & 1\end{pmatrix} \to \begin{pmatrix}1 & -2 & 0\\0 & 1 & 1/7\end{pmatrix} \to \begin{pmatrix}1 & 0 & 2/7\\0 & 1 & 1/7\end{pmatrix}.$$

Η λύση λοιπόν της $A^{\mathrm{T}}\boldsymbol{x} = \mathbf{0}$ είναι

$$\begin{pmatrix}x_1\\x_2\\x_3\end{pmatrix} = \begin{pmatrix}-\frac{2}{7}x_3\\-\frac{1}{7}x_3\\x_3\end{pmatrix} = -\frac{x_3}{7}\begin{pmatrix}2\\1\\-7\end{pmatrix}$$

επομένως

$$\operatorname{null} A^{\mathrm{T}} = \operatorname{span}\left\{\begin{pmatrix}2\\1\\-7\end{pmatrix}\right\}.$$

Παρατήρηση 6.9. Για τα μητρώα Q, S και A του Παραδείγματος 6.4 συμπεράναμε ότι $\operatorname{null} A = \operatorname{null} S$ και $\operatorname{null} A^{\mathrm{T}} = \operatorname{null} Q^{\mathrm{T}}$. Γενικεύοντας έχουμε το ακόλουθο αποτέλεσμα. *Εάν για το $n \times m$ μητρώο A και το $m \times n$ μητρώο B ισχύει $\operatorname{rank} A = \operatorname{rank} B = m$, συνεπώς $n \geq m$, και $C = AB$, τότε $\operatorname{null} C = \operatorname{null} B$.*

Πράγματι, έστω $x \in \operatorname{null} B$, τότε

$$C\boldsymbol{x} = AB\boldsymbol{x} = A\boldsymbol{0} = \boldsymbol{0},$$

συνεπώς $\boldsymbol{x} \in \operatorname{null} C$, οπότε $\operatorname{null} B \subseteq \operatorname{null} C$. Για να προκύψει η ισότητα αρκεί να δείξουμε ότι $\operatorname{null} C \subseteq \operatorname{null} B$. Έστω λοιπόν $\boldsymbol{x} \in \operatorname{null} C$, τότε

$$C\boldsymbol{x} = AB\boldsymbol{x} = A\boldsymbol{y} = \boldsymbol{0},$$

όπου $\boldsymbol{y} = B\boldsymbol{x} \in \mathbb{R}^m$. Το τελευταίο σύστημα γράφεται

$$\boldsymbol{a}_1 y_1 + \boldsymbol{a}_2 y_2 + \cdots + \boldsymbol{a}_m y_m = \boldsymbol{0}$$

όπου $\boldsymbol{a}_j$, $j = 1, 2, \ldots, m$ είναι οι στήλες του A. Επειδή $\operatorname{rank} A = m$ οι στήλες είναι γραμμικά ανεξάρτητες κατά συνέπεια $y_1 = y_2 = \cdots = y_m = 0$, οπότε $\boldsymbol{y} = \boldsymbol{0}$, ισοδύναμα $\boldsymbol{x} \in \operatorname{null} B$, ισοδύναμα $\operatorname{null} C \subseteq \operatorname{null} B$ που είναι το ζητούμενο.

► **Άσκηση 6.3.** Εξετάστε αν το αποτέλεσμα που περιγράψαμε στην Παρατήρηση 6.9 ισχύει όταν $\operatorname{rank} A = \operatorname{rank} B = n$ $(n < m)$.

► **Άσκηση 6.4.** Σε σχέση με το Παράδειγμα 6.4 βρείτε το μητρώο A εκτελώντας τον πολλαπλασιασμό και μετά βρείτε, από την 3×3 μορφή του A, τους τέσσερεις υπόχωρους που παράγονται από το A και συγκρίνετε τα αποτελέσματά σας, με αυτά του Παραδείγματος 6.4.

6.5 Ασκήσεις

1. Δίνεται το μητρώο
$$A = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{pmatrix}$$
   (α΄) Δείξτε ότι $\operatorname{rank} A = 3$.
   (β΄) Δείξτε ότι το A αντιστρέφεται και βρείτε το A^{-1}.
   (γ΄) Να βρεθεί ο χώρος στηλών του A.
   (δ΄) Να βρεθεί ο μηδενόχωρος του A.

2. Για το μητρώο
$$A = \begin{pmatrix} 2 & -1 & 1 \\ 4 & 1 & 2 \end{pmatrix}$$
   (α΄) Να βρεθεί μια βάση για τον χώρο στηλών του A.
   (β΄) Να βρεθεί μια βάση για τον χώρο γραμμών του A
   (γ΄) Να βρεθεί μια βάση για τον μηδενόχωρο του A.

3. Θεωρούμε το σύστημα $A\boldsymbol{x} = \boldsymbol{b}$, όπου

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 5 \\ -1 & 3 & 1 \end{pmatrix} \qquad \text{και} \qquad \boldsymbol{b} = \begin{pmatrix} 1 \\ 3 \\ b \end{pmatrix}$$

με b να είναι μια πραγματική παράμετρος.

   (α′) Είναι το μητρώο A αντιστρέψιμο;

   (β′) Να βρεθούν ο χώρος στηλών, ο χώρος γραμμών και ο μηδενόχωρος του A.

   (γ′) Για ποια b το σύστημα έχει λύση ή λύσεις;

   (δ′) Να βρεθεί η γενική λύση του συστήματος για εκείνα τα $\boldsymbol{b}$ για τα οποία είναι συμβιβαστό.

4. Εάν A είναι το 3×4 επαυξημένο μητρώο για το σύστημα

$$\begin{aligned} a_1x + b_1y + c_1z &= d_1 \\ a_2x + b_2y + c_2z &= d_2 \\ a_3x + b_3y + c_3z &= d_3 \end{aligned}$$

με $a_1 \neq 0$, $b_2 \neq 0$ και $c_3 \neq 0$ να βρεθεί η **ανηγμένη κλιμακωτή μορφή** R του A σε κάθε μια από τις περιπτώσεις:

   (α′) Το σύστημα δεν έχει λύση.

   (β′) Το σύστημα έχει μοναδική λύση

   (γ′) Το σύστημα έχει άπειρες λύσεις.

5. **Διάζευξη Fredholm:** Ακριβώς ένα από τα δύο προβλήματα έχει λύση

   Π1: $A\boldsymbol{x} = \boldsymbol{b}$

   Π2: $A^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}$ με $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} \neq 0$.

   Ισοδύναμη διατύπωση: Το σύστημα $A\boldsymbol{x} = \boldsymbol{b}$ έχει λύση αν και μόνο αν για κάθε $\boldsymbol{y}$ με $A^{\mathrm{T}}\boldsymbol{y} = \boldsymbol{0}$ ισχύει $\boldsymbol{y}^{\mathrm{T}}\boldsymbol{b} = 0$ ($\boldsymbol{b} \in (\operatorname{null} A^{\mathrm{T}})^{\perp}$).

6. Επιβεβαιώνοντας ότι καθένα από τα δοσμένα μητρώα A_k, $k = 1, 2$

$$A_1 = \begin{pmatrix} 3 & 6 & 6 \\ 1 & 2 & 2 \\ 4 & 8 & 8 \end{pmatrix} \qquad A_2 = \begin{pmatrix} 2 & 4 & 6 & -4 \\ -1 & -2 & -3 & 2 \end{pmatrix}$$

είναι τάξης ένα βρείτε διανύσματα $\boldsymbol{u}$ και $\boldsymbol{v}$ για την κάθε περίπτωση ώστε $A_k = \boldsymbol{u}\boldsymbol{v}^{\mathrm{T}}$.

7. Εάν μια βάση για το χώρο στηλών του A είναι η $\{\boldsymbol{u}, \boldsymbol{v}\}$ και μια βάση για το χώρο γραμμών του A είναι η $\{\boldsymbol{w}_1, \boldsymbol{w}_2\}$, όπου

$$\boldsymbol{u} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}, \qquad \boldsymbol{v} = \begin{pmatrix} 2 \\ 1 \\ 1 \end{pmatrix}, \qquad \boldsymbol{w}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{w}_2 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}$$

δείξτε ότι $A = \boldsymbol{u}\boldsymbol{w}_1^{\mathrm{T}} + \boldsymbol{v}\boldsymbol{w}_2^{\mathrm{T}}$ και υπολογίστε το A.

8. Δοθέντων τριών πραγματικών αριθμών a, b, c με $a \neq 0$ επιλέξτε $d \in \mathbb{R}$ ώστε το μητρώο

$$M = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

να έχει τάξη ένα. Στη συνέχεια βρείτε βάσεις για τον χώρο γραμμών, το χώρο στηλών και τον μηδενόχωρο του M.

Κεφάλαιο 7

Προβολές

7.1 Προβολές

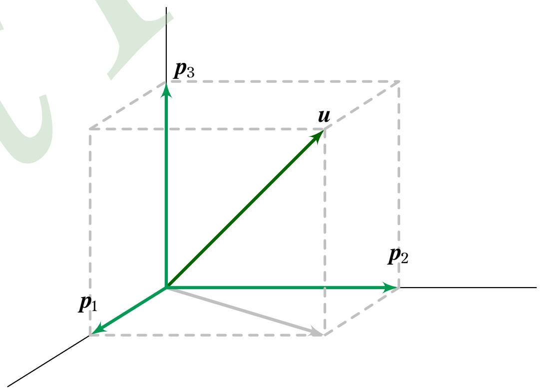
Εκφράζοντας το διάνυσμα $\boldsymbol{u} = \begin{pmatrix} \zeta & \eta & \xi \end{pmatrix}^{\mathrm{T}}$ μέσω της συνήθους βάσης

$$\boldsymbol{u} = \zeta \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + \eta \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} + \xi \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} = \zeta \boldsymbol{e_1} + \eta \boldsymbol{e_2} + \xi \boldsymbol{e_3}$$

θα λέμε τα διανύσματα

$$\boldsymbol{p}_1 = \zeta \boldsymbol{e_1} = \begin{pmatrix} \zeta \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{p}_2 = \eta \boldsymbol{e_2} = \begin{pmatrix} 0 \\ \eta \\ 0 \end{pmatrix}, \qquad \boldsymbol{p}_3 = \xi \boldsymbol{e_3} = \begin{pmatrix} 0 \\ 0 \\ \xi \end{pmatrix}$$

προβολές (projections) του $\boldsymbol{u}$, αντίστοιχα, στον x-άξονα, y-άξονα, z-άξονα, ή στους υποχώρους του $\mathbb{R}^3$, $\operatorname{span}\{\boldsymbol{e}_1\}$, $\operatorname{span}\{\boldsymbol{e}_2\}$, και $\operatorname{span}\{\boldsymbol{e}_3\}$ αντίστοιχα.



Σχήμα 7.1: Συνιστώσες και προβολές στο $\mathbb{R}^3$

Επειδή

$$\langle \boldsymbol{u}, \boldsymbol{e}_1 \rangle = \zeta, \qquad \langle \boldsymbol{u}, \boldsymbol{e}_2 \rangle = \eta, \qquad \langle \boldsymbol{u}, \boldsymbol{e}_3 \rangle = \xi,$$

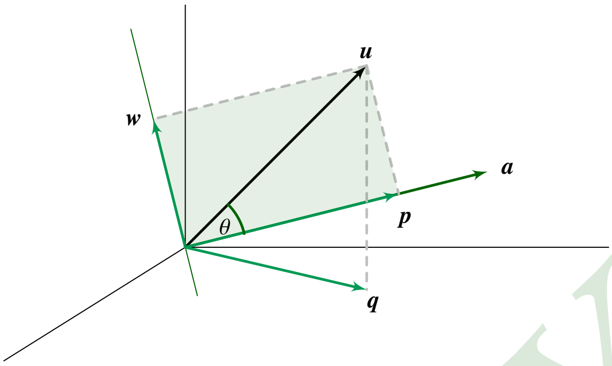
όπου $\langle \cdot, \cdot \rangle$ είναι το σύνηθες εσωτερικό γινόμενο, παίρνουμε όπως στο Θεώρημα 5.7

$$\boldsymbol{u} = \langle \boldsymbol{u}, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 + \langle \boldsymbol{u}, \boldsymbol{e}_2 \rangle \boldsymbol{e}_2 + \langle \boldsymbol{u}, \boldsymbol{e}_3 \rangle \boldsymbol{e}_3,$$

απ΄ όπου συμπεραίνουμε ότι

$$\boldsymbol{p}_1 = \langle \boldsymbol{u}, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1, \qquad \boldsymbol{p}_2 = \langle \boldsymbol{u}, \boldsymbol{e}_2 \rangle \boldsymbol{e}_2, \qquad \boldsymbol{p}_3 = \langle \boldsymbol{u}, \boldsymbol{e}_3 \rangle \boldsymbol{e}_3. \tag{7.1}$$

Αν τώρα $\boldsymbol{a}$ είναι ένα μη μηδενικό διάνυσμα και $\boldsymbol{u}$ είναι ένα διάνυσμα μας ενδιαφέρει να βρούμε την προβολή του $\boldsymbol{u}$ επί της ευθείας η οποία περιέχει το $\boldsymbol{a}$, βλέπε Σχήμα 7.2.



Σχήμα 7.2: Προβολή διανύσματος επί ευθείας και επί επιπέδου.

Αν $\boldsymbol{p}$ είναι η προβολή του $\boldsymbol{u}$ επί της ευθείας δια του $\boldsymbol{a}$, τότε

$$\boldsymbol{u} = \boldsymbol{p} + \boldsymbol{w} \qquad \text{με} \qquad \boldsymbol{p} = \lambda \boldsymbol{a},$$

όπου το $\boldsymbol{w}$ ανήκει στο ορθογώνιο συμπλήρωμα του $\boldsymbol{a}^{\perp}$ στο επίπεδο που ορίζουν τα $\boldsymbol{a}$ και $\boldsymbol{u}$, συνεπώς $\langle \boldsymbol{a}, \boldsymbol{w} \rangle = 0$, και λ είναι μια πραγματική σταθερά. Από τη σχέση ορθογωνιότητας και τη μορφή του $\boldsymbol{p}$ υπολογίζουμε

$$\begin{aligned} \langle \boldsymbol{a}, \boldsymbol{u} \rangle &= \langle \boldsymbol{a}, \boldsymbol{p} + \boldsymbol{w} \rangle \\ &= \langle \boldsymbol{a}, \boldsymbol{p} \rangle + \langle \boldsymbol{a}, \boldsymbol{w} \rangle \\ &= \langle \boldsymbol{a}, \boldsymbol{p} \rangle \\ &= \lambda \langle \boldsymbol{a}, \boldsymbol{a} \rangle \end{aligned}$$

απ΄ όπου προκύπτει η τιμή του λ, έτσι από την $\boldsymbol{p} = \lambda \boldsymbol{a}$ τελικά βρίσκουμε

$$\boldsymbol{p} = \frac{\langle \boldsymbol{a}, \boldsymbol{u} \rangle}{\langle \boldsymbol{a}, \boldsymbol{a} \rangle} \boldsymbol{a} = \frac{\boldsymbol{a}^{\mathrm{T}} \boldsymbol{u}}{\boldsymbol{a}^{\mathrm{T}} \boldsymbol{a}} \boldsymbol{a} \tag{7.2}$$

από τον ορισμό του συνήθους εσωτερικού γινομένου.[1] Σημειώνουμε ότι οι εκφράσεις των προβολών στην (7.1) είναι ειδικές περιπτώσεις της (7.2).

Στο Σχήμα 7.2 βλέπουμε, επίσης, την προβολή $\boldsymbol{q}$ του διανύσματος $\boldsymbol{u}$ επί του xy-επιπέδου, δηλαδή επί του υπόχωρου $\operatorname{span}\{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ του $\mathbb{R}^3$, η οποία, σύμφωνα με την (7.2), είναι

$$\boldsymbol{q} = \boldsymbol{p}_1 + \boldsymbol{p}_2 = \frac{\boldsymbol{e}_1 \cdot \boldsymbol{u}}{\boldsymbol{e}_1 \cdot \boldsymbol{e}_1} \boldsymbol{e}_1 + \frac{\boldsymbol{e}_2 \cdot \boldsymbol{u}}{\boldsymbol{e}_2 \cdot \boldsymbol{e}_2} \boldsymbol{e}_2.$$

ή, επειδή $\boldsymbol{e}_1 \cdot \boldsymbol{e}_1 = \boldsymbol{e}_2 \cdot \boldsymbol{e}_2 = 1$,

$$\boldsymbol{q} = \langle \boldsymbol{u}, \boldsymbol{e}_1 \rangle \boldsymbol{e}_1 + \langle \boldsymbol{u}, \boldsymbol{e}_2 \rangle \boldsymbol{e}_2.$$

---

[1]Δίνουμε μια πιο ¨γεωμετρική¨ απόδειξη της (7.2). Αν θ είναι η γωνία μεταξύ των $\boldsymbol{a}$ και $\boldsymbol{u}$, τότε

$$\|\boldsymbol{p}\| = \|\boldsymbol{u}\| \cos\theta$$

κατά συνέπεια (γιατί;)

$$\boldsymbol{p} = \|\boldsymbol{u}\| \cos\theta \frac{\boldsymbol{a}}{\|\boldsymbol{a}\|} = \frac{\|\boldsymbol{a}\| \|\boldsymbol{u}\| \cos\theta}{\|\boldsymbol{a}\|^2} \boldsymbol{a},$$

η οποία συμφωνεί με την (7.2).

Ορισμός 7.1. Έστω X ένας διανυσματικός χώρος με εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$. Εάν W είναι ένας πεπερασμένης διάστασης υπόχωρος του X και $\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ είναι μια ορθοκανονική βάση του W και $\boldsymbol{u}$ είναι ένα διάνυσμα του X ορίζουμε την **ορθογώνια προβολή του $\boldsymbol{u}$ επί του** W το διάνυσμα

$$\operatorname{proj}_W \boldsymbol{u} := \langle \boldsymbol{u}, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1 + \langle \boldsymbol{u}, \boldsymbol{w}_2 \rangle \boldsymbol{w}_2 + \cdots + \langle \boldsymbol{u}, \boldsymbol{w}_n \rangle \boldsymbol{w}_n. \tag{7.3}$$

Το διάνυσμα

$$\boldsymbol{u} - \operatorname{proj}_W \boldsymbol{u} = \boldsymbol{u} - \langle \boldsymbol{u}, \boldsymbol{w}_1 \rangle \boldsymbol{w}_1 - \langle \boldsymbol{u}, \boldsymbol{w}_2 \rangle \boldsymbol{w}_2 - \cdots - \langle \boldsymbol{u}, \boldsymbol{w}_n \rangle \boldsymbol{w}_n \tag{7.4}$$

θα το λέμε **ορθογώνια στο W συνιστώσα του $\boldsymbol{u}$.**

Ο ορισμός του $\boldsymbol{u} - \operatorname{proj}_W \boldsymbol{u}$ ως ορθογώνια στο W συνιστώσα του $\boldsymbol{u}$ διακαιολογείται από τα Θεωρήματα 5.8 και 5.2, αφού $\boldsymbol{u} - \operatorname{proj}_W \boldsymbol{u} \in W^{\perp}$.

Θεώρημα 7.1 (Βέλτιστη προσέγγιση). *Εάν W είναι ένας πεπερασμένης διάστασης υπόχωρος ενός διανυσματικού χώρου X με εσωτερικό γινόμενο και επαγόμενη νόρμα $\|\cdot\|$, και εάν $\boldsymbol{u} \in X$, τότε για κάθε $\boldsymbol{w} \in W$ με $\boldsymbol{w} \neq \operatorname{proj}_W \boldsymbol{u}$ ισχύει*

$$\|\boldsymbol{u} - \operatorname{proj}_W \boldsymbol{u}\| < \|\boldsymbol{u} - \boldsymbol{w}\|$$

δηλαδή το διάνυσμα $\operatorname{proj}_W \boldsymbol{u}$ είναι η βέλτιστη προσέγγιση του $\boldsymbol{u}$ από το W.

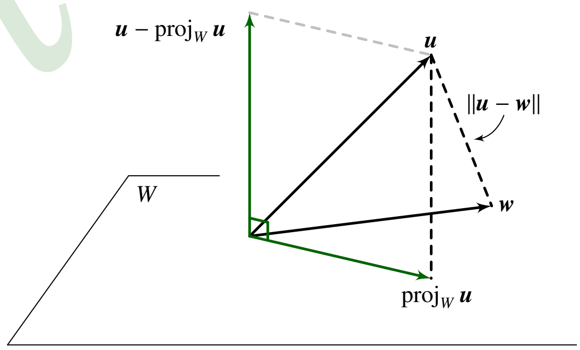
Απόδειξη. Γράφοντας

$$\boldsymbol{u} - \boldsymbol{w} = (\operatorname{proj}_W \boldsymbol{u} - \boldsymbol{w}) + (\boldsymbol{u} - \operatorname{proj}_W \boldsymbol{u})$$

επειδή $\operatorname{proj}_W \boldsymbol{u} - \boldsymbol{w}$ είναι σαν διαφορά διανυσμάτων του W διάνυσμα του W θα είναι ορθογώνιο στο $\boldsymbol{u} - \operatorname{proj}_W \boldsymbol{u} \in W^{\perp}$, κατά συνέπεια από το Πυθαγόρειο Θεώρημα θα έχουμε ότι

$$\|\boldsymbol{u} - \boldsymbol{w}\|^2 = \|\operatorname{proj}_W \boldsymbol{u} - \boldsymbol{w}\|^2 + \|\boldsymbol{u} - \operatorname{proj}_W \boldsymbol{u}\|^2.$$

Αφού $\boldsymbol{w} \neq \operatorname{proj}_W \boldsymbol{u}$ θα είναι $\|\operatorname{proj}_W \boldsymbol{u} - \boldsymbol{w}\| > 0$, επομένως από την τελευταία ισότητα έπεται



Σχήμα 7.3: Προσέγγιση διανύσματος από υπόχωρο.

ότι

$$\|\boldsymbol{u} - \boldsymbol{w}\|^2 > \|\boldsymbol{u} - \operatorname{proj}_W \boldsymbol{u}\|^2.$$

απ΄ όπου έπεται το ζητούμενο. □

Παράδειγμα 7.1. Στο χώρο $C[-1,1]$, των συνεχών συναρτήσεων στο διάστημα $[-1,1]$, με εσωτερικό γινόμενο

$$\langle f, g\rangle = \int_{-1}^{1} f(x)g(x)\,\mathrm{d}x \tag{7.5}$$

να βρεθεί η βέλτιστη προσέγγιση της $f(x) = x^3$ από πολυώνυμα βαθμού το πολύ δύο. Η βέλτιστη προσέγγιση της f είναι η προβολή της f επί του υπόχωρου $\mathbb{P}_2[-1,1]$. Μια βάση για τον $\mathbb{P}_2$, η συνήθης, είναι η $\{1, x, x^2\}$, αλλά αυτή δεν είναι ορθοκανονική.
Βήμα 1. Βρίσκουμε μια ορθοκανονική, ως προς το εσωτερικό γινόμενο (7.5) βάση στον $\mathbb{P}_2$. Εφαρμόζουμε τον αλγόριθμο Gram-Schmidt στη συνήθη βάση.

$$\begin{aligned}
&q_0 = 1 && \\
&\|q_0\|^2 = \int_{-1}^{1} 1\,\mathrm{d}x = 2 && p_0(x) = \frac{1}{\sqrt{2}}\\
&q_1(x) = x - \langle x, p_0\rangle p_0(x), && \\
&\langle x, p_0\rangle = \int_{-1}^{1} x\frac{1}{\sqrt{2}}\,\mathrm{d}x = 0 && q_1(x) = x\\
&\|q_1\|^2 = \int_{-1}^{1} x^2\,\mathrm{d}x = \frac{2}{3} && p_1(x) = \sqrt{\frac{3}{2}}x\\
&q_2(x) = x^2 - \langle x^2, p_0\rangle p_0(x) - \langle x^2, p_1\rangle p_1(x) && \\
&\langle x^2, p_0\rangle = \int_{-1}^{1} x^2\frac{1}{\sqrt{2}}\,\mathrm{d}x = \frac{\sqrt{2}}{3} && \\
&\langle x^2, p_1\rangle = \int_{-1}^{1} x^2\sqrt{\frac{3}{2}}x\,\mathrm{d}x = 0 && q_2(x) = x^2 - \frac{1}{3}\\
&\|q_2\|^2 = \int_{-1}^{1}\left(x^2 - \frac{1}{3}\right)^2\mathrm{d}x = \frac{8}{45} && p_2(x) = \sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right).
\end{aligned}$$

Έτσι μια ορθοκανονική βάση για το $\mathbb{P}_2[-1,1]$ αποτελείται από τα πολυώνυμα p_0, p_1 και p_2 τα οποία είναι τα τρία πρώτα πολυώνυμα Legendre, βλέπε Παράρτημα Β: Ορθογώνια πολυώνυμα, παράγραφο 5.8.
Βήμα 2. Η βέλτιστη προσέγγιση της f από πολυώνυμο βαθμού το πολύ δύο είναι

$$\operatorname{proj}_{\mathbb{P}_2} f = \langle f, p_0\rangle p_0 + \langle f, p_1\rangle p_1 + \langle f, p_2\rangle p_2.$$

Υπολογίζουμε λοιπόν

$$\begin{aligned}
\langle f, p_0\rangle &= \int_{-1}^{1} x^3\frac{1}{\sqrt{2}}\,\mathrm{d}x = 0\\
\langle f, p_1\rangle &= \int_{-1}^{1} x^3\sqrt{\frac{3}{2}}x\,\mathrm{d}x = \sqrt{\frac{3}{2}}\int_{-1}^{1} x^4\,\mathrm{d}x = \frac{\sqrt{6}}{5}\\
\langle f, p_2\rangle &= \int_{-1}^{1} x^3\sqrt{\frac{45}{8}}\left(x^2 - \frac{1}{3}\right)\mathrm{d}x = \sqrt{\frac{45}{8}}\int_{-1}^{1}\left(x^5 - \frac{1}{3}x^3\right)\mathrm{d}x = 0
\end{aligned}$$

επομένως

$$\operatorname{proj}_{\mathbb{P}_2} f(x) = \frac{\sqrt{6}}{5}\sqrt{\frac{3}{2}}x = \frac{3}{5}x,$$

το οποίο δίνει την βέλτιστη προσέγγιση της $f(x) = x^3$ από πολυώνυμα του $\mathbb{P}_2[-1,1]$, δηλαδή
$$\min_{a,b,c} \|x^3 - a - bx - cx^2\| = \left\| x^3 - \frac{3}{5}x \right\|.$$

Παρατήρηση 7.1. Το Θεώρημα 7.1 μπορεί να διατυπωθεί και ως ένα πρόβλημα βελτιστοποίησης, συγκεκριμένα αν X είναι ένας διανυσματικός χώρος με εσωτερικό γινόμενο και $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ είναι ένα ορθοκανονικό σύνολο διανυσμάτων του X, τότε για κάθε διάνυσμα $\boldsymbol{u}$ του X ισχύει

$$\min_{\lambda_1,\ldots,\lambda_n} \|\boldsymbol{u} - \lambda_1\boldsymbol{w}_1 - \cdots - \lambda_n\boldsymbol{w}_n\| = \|\boldsymbol{u} - \langle \boldsymbol{u}, \boldsymbol{w}_1\rangle \boldsymbol{w}_1 - \cdots - \langle \boldsymbol{u}, \boldsymbol{w}_n\rangle \boldsymbol{w}_n\|. \tag{7.6}$$

Ακολουθεί μια απόδειξη διαφορετική από αυτή που δόθηκε στο Θεώρημα 7.1. Υπολογίζοντας

$$\begin{aligned}
\|\boldsymbol{u} - \lambda_1\boldsymbol{w}_1 - \cdots - \lambda_n\boldsymbol{w}_n\|^2 &= \langle \boldsymbol{u} - \sum_{i=1}^n \lambda_i \boldsymbol{w}_i, \boldsymbol{u} - \sum_{i=1}^n \lambda_i \boldsymbol{w}_i\rangle \\
&= \langle \boldsymbol{u}, \boldsymbol{u}\rangle - 2\sum_{i=1}^n \lambda_i \langle \boldsymbol{u}, \boldsymbol{w}_i\rangle + \sum_{i,j=1}^n \lambda_i\lambda_j \langle \boldsymbol{w}_i, \boldsymbol{w}_j\rangle \\
&= \|\boldsymbol{u}\|^2 - 2\sum_{i=1}^n \lambda_i \langle \boldsymbol{u}, \boldsymbol{w}_i\rangle + \sum_{i=1}^n \lambda_i^2 \\
&= \|\boldsymbol{u}\|^2 + \sum_{i=1}^n (\lambda_i^2 - 2\lambda_i \langle \boldsymbol{u}, \boldsymbol{w}_i\rangle) \\
&= \|\boldsymbol{u}\|^2 + \sum_{i=1}^n (\lambda_i^2 - 2\lambda_i \langle \boldsymbol{u}, \boldsymbol{w}_i\rangle + \langle \boldsymbol{u}, \boldsymbol{w}_i\rangle^2 - \langle \boldsymbol{u}, \boldsymbol{w}_i\rangle^2) \\
&= \|\boldsymbol{u}\|^2 - \sum_{i=1}^n \langle \boldsymbol{u}, \boldsymbol{w}_i\rangle^2 + \sum_{i=1}^n (\lambda_i - \langle \boldsymbol{u}, \boldsymbol{w}_i\rangle)^2
\end{aligned}$$

βλέπουμε αμέσως, από την τελευταία ισότητα, ότι το ελάχιστο της ποσότητας $\|\boldsymbol{u} - \lambda_1\boldsymbol{w}_1 - \cdots - \lambda_n\boldsymbol{w}_n\|$ συμβαίνει όταν $\lambda_i = \langle \boldsymbol{u}, \boldsymbol{w}_i\rangle$ γιά κάθε $i = 1, 2, \ldots, n$.

Παρατήρηση 7.2. Κατά την απόδειξη της (7.6) είδαμε ότι

$$\|\boldsymbol{u} - \lambda_1\boldsymbol{w}_1 - \cdots - \lambda_n\boldsymbol{w}_n\|^2 = \|\boldsymbol{u}\|^2 - 2\sum_{i=1}^n \lambda_i \langle \boldsymbol{u}, \boldsymbol{w}_i\rangle + \sum_{i=1}^n \lambda_i^2$$

κατά συνέπεια η προς ελαχιστοποίηση ποσότητα είναι μια πολυωνυμική συνάρτηση των λ_i, με $i = 1, 2, \ldots, n$ έστω $E(\lambda_1, \lambda_2, \ldots, \lambda_n)$. Εφαρμόζουμε στη συνέχεια μεθόδους του Απειροστικού λογισμού για την εύρεση του ελαχίστου της E. Βρίσκουμε τα κρίσιμα σημεία της E, στη περίπτωση αυτή, τα σημεία για τα οποία

$$\frac{\partial E}{\partial \lambda_1} = \frac{\partial E}{\partial \lambda_2} = \cdots = \frac{\partial E}{\partial \lambda_n} = 0.$$

Έτσι για $j = 1, 2, \ldots, n$

$$\frac{\partial E}{\partial \lambda_j} = \frac{\partial E}{\partial \lambda_j}\Big(\|\boldsymbol{u}\|^2 - 2\sum_{i=1}^n \lambda_i \langle \boldsymbol{u}, \boldsymbol{w}_i\rangle + \sum_{i=1}^n \lambda_i^2 \Big) = -2\langle \boldsymbol{u}, \boldsymbol{w}_j\rangle + 2\lambda_j$$

συνεπώς υπάρχει ένα μόνο κρίσιμο σημείο $(\lambda_1, \lambda_2, \ldots, \lambda_n)$ με $\lambda_j = \langle \boldsymbol{u}, \boldsymbol{w}_j\rangle$ για $j = 1, 2, \ldots, n$. Επειδή

$$\frac{\partial^2 E}{\partial \lambda_j^2} = 2 \qquad j = 1, 2, \ldots, n$$

η συνάρτηση E είναι κυρτή προς κάθε κατεύθυνση, επομένως στο κρίσιμο σημείο η E παίρνει την ελάχιστη τιμή της γεγονός που αποδεικνύει την (7.6).

Παρατήρηση 7.3. Στο Παράδειγμα 7.1 υπολογίσαμε το ελάχιστο της ποσότητας $\|x^3 - a - bx - cx^2\|$. Όπως στην Παρατήρηση 7.2 θεωρούμε την συνάρτηση

$$E(a,b,c) = \|x^3 - a - bx - cx^2\|^2 = \int_{-1}^{1} (x^3 - a - bx - cx^2)^2 \,\mathrm{d}x$$

και βρίσκουμε τα κρίσιμα σημεία της E.Για τον υπολογισμό των μερικών παραγώγων μπορούμε να χρησιμοποιήσουμε το αποτέλεσμα της Ανάλυσης[2]

$$\frac{\partial}{\partial \lambda} \int_{-1}^{1} (x^3 - a - bx - cx^2)^2 \,\mathrm{d}x = \int_{-1}^{1} \frac{\partial}{\partial \lambda} (x^3 - a - bx - cx^2)^2 \,\mathrm{d}x \tag{7.7}$$

όπου $\lambda \in \{a,b,c\}$[3]. Υπολογίζοντας

$$\begin{aligned}
\frac{\partial E}{\partial a} &= \int_{-1}^{1} -2(x^3 - a - bx - cx^2)\,\mathrm{d}x = 2a \int_{-1}^{1} -2a\,\mathrm{d}x + 2c \int_{-1}^{1} x^2\,\mathrm{d}x = 4a + \frac{4}{3}c = 0\\
\frac{\partial E}{\partial b} &= \int_{-1}^{1} -2x(x^3 - a - bx - cx^2)\,\mathrm{d}x = -2 \int_{-1}^{1} x^4\,\mathrm{d}x + 2b \int_{-1}^{1} x^2\,\mathrm{d}x = -\frac{4}{5} + \frac{4}{5}b = 0\\
\frac{\partial E}{\partial c} &= \int_{-1}^{1} -2x^2(x^3 - a - bx - cx^2)\,\mathrm{d}x = 2a \int_{-1}^{1} x^2\,\mathrm{d}x + 2c \int_{-1}^{1} x^4\,\mathrm{d}x = \frac{4}{3}a + \frac{4}{5}c = 0
\end{aligned}$$

βρίσκουμε $a = 0$, $b = 3/5$, και $c = 0$. Υπάρχει λοιπόν ένα κρίσιμο σημείο και επειδή

$$\begin{aligned}
E(0,0,0) &= \int_{-1}^{1} x^6\,\mathrm{d}x = \frac{2}{7}\\
E(0,3/5,0) &= \int_{-1}^{1} \left(x^3 - \frac{3}{5}x\right)^2 \mathrm{d}x = \int_{-1}^{1} \left(x^6 - \frac{6}{5}x^4 + \frac{9}{25}x^2\right)\mathrm{d}x = \frac{2}{7} - \frac{6}{25}
\end{aligned}$$

έπεται ότι στο $(0,3/5,0)$ η συνάρτηση E παίρνει την ελάχιστη τιμή της, όπως εξάλλου είδαμε στο Παράδειγμα 7.1.

---

[2]**Θεώρημα.** *Έστω ότι το ολοκλήρωμα*

$$F(t) = \int_a^b \phi(x,t)\,\mathrm{d}x$$

υπάρχει για κάθε $t \in [p,q]$. *Αν η μερική παράγωγος* $\partial\phi/\partial t$ *είναι συνεχής στο* $Q = \{(x,t) : a \le x \le b,\ \ p \le t \le q\}$, *τότε η* $F'(t)$ *υπάρχει για κάθε* $t \in (p,q)$ *επιπλέον*

$$F'(t) = \int_a^b \frac{\partial}{\partial t}\phi(x,t)\,\mathrm{d}x.$$

Βλέπε T. M. Apostol *Mathematical Analysis*, 2nd edition Addison-Wesley, Reading Massachussets, 1974.

[3]Αν δυσπιστείτε για την ισχύ της (7.7) γράψτε

$$\begin{aligned}
\int_{-1}^{1} (x^3 - a - bx - cx^2)^2\,\mathrm{d}x &= \int_{-1}^{1} (x^6 + a^2 + b^2x^2 + c^2x^4 - 2ax^3 - 2bx^4 - 2cx^5 + 2abx + 2acx^2 + 2bcx^3)\,\mathrm{d}x\\
&= \int_{-1}^{1} x^6\,\mathrm{d}x + a^2 \int_{-1}^{1} \mathrm{d}x + (b^2 + 2ac) \int_{-1}^{1} x^2\,\mathrm{d}x + (c^2 - 2b) \int_{-1}^{1} x^4\,\mathrm{d}x\\
&= \frac{2}{7} + 2a^2 + \frac{2}{3}b^2 + \frac{2}{5}c^2 + \frac{4}{3}ac - \frac{4}{5}b.
\end{aligned}$$

7.2 Προσεγγίσεις ελαχίστων τετραγώνων

Για ένα σύστημα $A\boldsymbol{x} = \boldsymbol{b}$ ένα διάνυσμα $\boldsymbol{x}$ θα είναι προσεγγιστική λύση αν η ποσότητα $\|A\boldsymbol{x} - \boldsymbol{b}\|$ είναι μικρή. Σημειώνουμε ότι $\|A\boldsymbol{x} - \boldsymbol{b}\| = 0$ αν και μόνο αν το $\boldsymbol{x}$ είναι λύση του συστήματος, κατά συνέπεια αν το σύστημα δεν έχει λύση θα θέλαμε να γνωρίζουμε εκείνο το διάνυσμα $\boldsymbol{x}$ για το οποίο η $\|A\boldsymbol{x} - \boldsymbol{b}\|$ είναι η ελάχιστη δυνατή. Γράφοντας

$$A\boldsymbol{x} - \boldsymbol{b} = \begin{pmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{pmatrix}$$

έχουμε ότι

$$\|A\boldsymbol{x} - \boldsymbol{b}\|^2 = e_1^2 + e_2^2 + \cdot + e_n^2,$$

κατά συνέπεια ένα διάνυσμα $\boldsymbol{x}$ το οποίο ελαχιστοποιεί την $\|A\boldsymbol{x} - \boldsymbol{b}\|$ ελαχιστοποιεί το άθροισμα $e_1^2 + e_2^2 + \cdot + e_n^2$. Ένα τέτοιο διάνυσμα θα το λέμε **λύση των ελαχίστων τετραγώνων** και την ποσότητα $\|A\boldsymbol{x} - \boldsymbol{b}\|$ θα τη λέμε **σφάλμα των ελαχίστων τετραγώνων** .

Πρόβλημα: Να βρεθεί η λύση ελαχίστων τετραγώνων για το σύστημα $A\boldsymbol{x} = \boldsymbol{b}$.

Αν W είναι ο χώρος στηλών του μητρώου A, τότε

$$\boldsymbol{b} = \operatorname{proj}_W \boldsymbol{b} + \boldsymbol{b} - \operatorname{proj}_W \boldsymbol{b}, \qquad \text{και} \qquad \operatorname{proj}_W \boldsymbol{b} \perp \boldsymbol{b} - \operatorname{proj}_W \boldsymbol{b},$$

έτσι

$$\boldsymbol{b} - A\boldsymbol{x} = \operatorname{proj}_W \boldsymbol{b} - A\boldsymbol{x} + \boldsymbol{b} - \operatorname{proj}_W \boldsymbol{b}, \qquad \text{και} \qquad \operatorname{proj}_W \boldsymbol{b} - A\boldsymbol{x} \perp \boldsymbol{b} - \operatorname{proj}_W \boldsymbol{b}, \tag{7.8}$$

αφού $A\boldsymbol{x} \in W$, επομένως

$$\|\boldsymbol{b} - A\boldsymbol{x}\|^2 = \|\operatorname{proj}_W \boldsymbol{b} - A\boldsymbol{x}\|^2 + \|\boldsymbol{b} - \operatorname{proj}_W \boldsymbol{b}\|^2.$$

Από την τελευταία σχέση προκύπτει ότι το σφάλμα ελαχιστοποιείται όταν $A\boldsymbol{x} = \operatorname{proj}_W \boldsymbol{b}$, ισοδύναμα η λύση των ελαχίστων τετραγώνων, η οποία υπάρχει πάντα (γιατί;), είναι η $\widehat{\boldsymbol{x}}$ για την οποία

$$A\widehat{\boldsymbol{x}} = \operatorname{proj}_W \boldsymbol{b}. \tag{7.9}$$

Έτσι από την (7.8) παίρνουμε

$$\boldsymbol{b} - A\widehat{\boldsymbol{x}} = \boldsymbol{b} - \operatorname{proj}_W \boldsymbol{b}, \qquad \text{με} \qquad \boldsymbol{b} - \operatorname{proj}_W \boldsymbol{b} \in W^{\perp} = (\operatorname{range} A)^{\perp} = \operatorname{null} A^{\mathrm{T}},$$

βλέπε Θεώρημα 6.5, οπότε

$$A^{\mathrm{T}}(\boldsymbol{b} - A\widehat{\boldsymbol{x}}) = A^{\mathrm{T}}(\boldsymbol{b} - \operatorname{proj}_W \boldsymbol{b}) = \boldsymbol{0}$$

επομένως

$$A^{\mathrm{T}}A\widehat{\boldsymbol{x}} = A^{\mathrm{T}}\boldsymbol{b}. \tag{7.10}$$

Σημειώνουμε ότι για την εύρεση του $\widehat{\boldsymbol{x}}$ δεν χρειάζεται να βρούμε την προβολή του b επί του W και στη συνέχεια να λύσουμε το σύστημα $A\boldsymbol{x} = \operatorname{proj}_W \boldsymbol{b}$. Η λύση των ελαχίστων τετραγώνων προκύπτει από την (7.10). Την εξίσωση στην (7.10) λέμε **κανονική εξίσωση**.

Παρατήρηση 7.4. Αν το μητρώο $A^{\mathrm{T}}A$ είναι αντιστρέψιμο από την κανονική εξίσωση παίρνουμε

$$\widehat{\boldsymbol{x}} = (A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b},$$

απ΄ όπου πολλαπλασιάζοντας από αριστερά με A προκύπτει μέσω της (7.9) ότι

$$\operatorname{proj}_W \boldsymbol{b} = A\widehat{\boldsymbol{x}} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{b}.$$

Η τελευταία σχέση μας λέει ότι για το μητρώο A το $A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}$ προβάλει επί του χώρου στηλών του A. Θα επανέλθουμε στο θέμα αυτό στην επόμενη παράγραφο.

■ **Εφαρμογή 7.1.** Με τη μέθοδο των ελαχίστων τετραγώνων να βρεθεί η ευθεία που προσεγγίζει καλύτερα τα σημεία $(x_1, y_1), (x_2, y_2) \dots, (x_n, y_n)$ του επιπέδου.

Έστω $y = \alpha + \beta x$ η ζητούμενη ευθεία. Αν τα δοσμένα σημεία βρίσκονταν επάνω στην ευθεία θα είχαμε

$$\begin{aligned} y_1 &= \alpha + \beta x_1 \\ y_2 &= \alpha + \beta x_2 \\ &\ \vdots \\ y_n &= \alpha + \beta x_n \end{aligned} \qquad \text{ισοδύναμα} \qquad \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} \begin{pmatrix} \alpha \\ \beta \end{pmatrix} = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix},$$

ή $A\boldsymbol{x} = \boldsymbol{b}$ με τις προφανείς αντιστοιχίες. Το σύστημα δεν έχει λύση σχεδόν ποτέ, εκτός αν τα δεδομένα είναι συνευθειακά, επομένως βρίσκουμε την λύση των ελαχίστων τετραγώνων. Υπολογίζοντας

$$A^{\mathrm{T}}A = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} 1 & x_1 \\ 1 & x_2 \\ \vdots & \vdots \\ 1 & x_n \end{pmatrix} = \begin{pmatrix} n & \sum x_j \\ \sum x_j & \sum x_j^2 \end{pmatrix}$$

και

$$A^{\mathrm{T}}b = \begin{pmatrix} 1 & 1 & \dots & 1 \\ x_1 & x_2 & \dots & x_n \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} \sum y_j \\ \sum x_j y_j \end{pmatrix}$$

η κανονική εξίσωση (7.10) γίνεται

$$\begin{pmatrix} n & \sum x_j \\ \sum x_j & \sum x_j^2 \end{pmatrix} \begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \begin{pmatrix} \sum y_j \\ \sum x_j y_j \end{pmatrix}. \tag{7.11}$$

Οι στήλες του A είναι γραμμικά ανεξάρτητες, εκτός αν $x_1 = x_2 = \cdots = x_n$ (γιατί;) κατά συνέπεια αν συμβολίσουμε με $\boldsymbol{a}_1$ και $\boldsymbol{a}_2$ τις στήλες από την ανισότητα Cauchy-Bunyakovsky-Schwarz έχουμε

$$\langle \boldsymbol{a}_1, \boldsymbol{a}_2 \rangle^2 < \langle \boldsymbol{a}_1, \boldsymbol{a}_1 \rangle \langle \boldsymbol{a}_2, \boldsymbol{a}_2 \rangle \Leftrightarrow \left(\sum_{j=1}^{n} x_j \right)^2 < \left(\sum_{j=1}^{n} 1^2 \right) \left(\sum_{j=1}^{n} x_j^2 \right) = n \sum_{j=1}^{n} x_j^2$$

επομένως η ορίζουσα του μητρώου των συντελεστών στη κανονική εξίσωση είναι διάφορη του μηδενός[4], συνεπώς

$$\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix} = \frac{1}{n \sum x_j^2 - (\sum x_j)^2} \begin{pmatrix} \sum x_j^2 & -\sum x_j \\ -\sum x_j & n \end{pmatrix} \begin{pmatrix} \sum y_j \\ \sum x_j y_j \end{pmatrix}.$$

Έτσι τελικά οι συντελεστές στην ευθεία $y = \hat{\alpha} + \hat{\beta} x$ των ελαχίστων τετραγώνων είναι

$$\hat{\alpha} = \frac{\sum x_j^2 \sum y_j - \sum x_j y_j \sum x_j}{n \sum x_j^2 - (\sum x_j)^2}, \qquad \hat{\beta} = \frac{-\sum x_j \sum y_j + n \sum x_j y_j \sum x_j}{n \sum x_j^2 - (\sum x_j)^2}. \tag{7.12}$$

► **Άσκηση 7.1.** Δείξτε ότι η ευθεία των ελαχίστων τετραγώνων διέρχεται από το κέντρο μάζας $(\bar{x}, \bar{y})$ των σημείων. Θυμίζουμε ότι

$$\bar{x} = \frac{1}{n} \sum x_j, \qquad \bar{y} = \frac{1}{n} \sum y_j.$$

7.3 Το μητρώο της ορθογώνιας προβολής

Η εξοικείωσή μας με τα στοιχειώδη μητρώα επιτρέπει να δούμε ότι οι προβολές ενός διανύσματος $\boldsymbol{u} \in \mathbb{R}^3$, επί των ευθειών x-άξονα, y-άξονα, z-άξονα υλοποιούνται ως

$$\boldsymbol{p}_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{u}, \qquad \boldsymbol{p}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{u}, \qquad \boldsymbol{p}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \boldsymbol{u} \tag{7.13}$$

αντίστοιχα, ή

$$\boldsymbol{p}_1 = \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \boldsymbol{u} = P_1 \boldsymbol{u}, \qquad \boldsymbol{p}_2 = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{e}_2 & \boldsymbol{0} \end{pmatrix} \boldsymbol{u} = P_2 \boldsymbol{u}, \qquad \boldsymbol{p}_3 = \begin{pmatrix} \boldsymbol{0} & \boldsymbol{0} & \boldsymbol{e}_3 \end{pmatrix} \boldsymbol{u} = P_3 \boldsymbol{u}$$

και επί του xy-επιπέδου ως

$$\boldsymbol{q} = \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{0} \end{pmatrix} \boldsymbol{u} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{u} = Q \boldsymbol{u}.$$

Πρόβλημα: Ποιο είναι το μητρώο P της ορθογώνιας προβολής ενός διανύσματος $\boldsymbol{u}$ του $\mathbb{R}^m$ πρώτον επί ενός μη μηδενικού διανύσματος $\boldsymbol{a}$, και δεύτερον επί ενός υποχώρου W του $\mathbb{R}^m$.

(1) Αναζητάμε μητρώο P ώστε $P\boldsymbol{u} = \boldsymbol{p} = \mathrm{proj}_{\boldsymbol{a}} \boldsymbol{u}$, ισοδύναμα για ποιο P είναι

$$P\boldsymbol{u} = \frac{\boldsymbol{a}^{\mathrm{T}} \boldsymbol{u}}{\boldsymbol{a}^{\mathrm{T}} \boldsymbol{a}} \boldsymbol{a}?$$

Παρατηρούμε ότι

$$(\boldsymbol{a}^{\mathrm{T}} \boldsymbol{u}) \boldsymbol{a} = \boldsymbol{a} (\boldsymbol{a}^{\mathrm{T}} \boldsymbol{u}) = (\boldsymbol{a} \boldsymbol{a}^{\mathrm{T}}) \boldsymbol{u}$$

από τις ιδιότητες του γινομένου μητρώων, κατά συνέπεια

$$\mathrm{proj}_{\boldsymbol{a}} \boldsymbol{u} = P\boldsymbol{u} = \frac{\boldsymbol{a} \boldsymbol{a}^{\mathrm{T}}}{\boldsymbol{a}^{\mathrm{T}} \boldsymbol{a}} \boldsymbol{u}. \tag{7.14}$$

[4]Διαφορετικά, η ορίζουσα του μητρώου των συντελεστών με $\boldsymbol{x} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}^{\mathrm{T}}$ είναι

$$n\|\boldsymbol{x}\|^2 - \Big(\sum_{j=1}^{n} x_j\Big)^2 = \|\boldsymbol{x}\|^2 \Big(n - \Big(\sum_{j=1}^{n} \frac{x_j}{\|\boldsymbol{x}\|}\Big)^2\Big) > 0$$

αφού $x_j \le |x_j| = (x_j^2)^{1/2} < (x_1^2 + x_2^2 + \cdots + x_n^2)^{1/2} = \|\boldsymbol{x}\|$, για κάθε j, όταν τουλάχιστον δύο από τα x_j είναι διάφορα του μηδενός.

(2) Έστω $W = \operatorname{span}\{\boldsymbol{a}_1, \boldsymbol{a}_2 \dots, \boldsymbol{a}_n\}$, όπου $\boldsymbol{a}_1, \boldsymbol{a}_2 \dots, \boldsymbol{a}_n$ είναι γραμμικά ανεξάρτητα διανύσματα στο $\mathbb{R}^m$. Αν A είναι το μητρώο με στήλες τα $\boldsymbol{a}_j$, τότε $\operatorname{range} A = W$, δηλαδή ο υπόχωρος W είναι ο χώρος στηλών του A. Έστω $\boldsymbol{c} \in \mathbb{R}^m$ τέτοιο ώστε $\operatorname{proj}_W \boldsymbol{u} = A\boldsymbol{c}$. Επειδή το $\boldsymbol{u} - A\boldsymbol{c}$ είναι ορθογώνιο στον W έπεται ότι $\boldsymbol{a}_j^{\mathrm{T}}(\boldsymbol{u} - A\boldsymbol{c}) = 0$ για $j = 1, 2, \dots, m$. Κατά συνέπεια

$$A^{\mathrm{T}}(\boldsymbol{u} - A\boldsymbol{c}) = 0 \Rightarrow A^{\mathrm{T}}\boldsymbol{u} = A^{\mathrm{T}}A\boldsymbol{c}$$

Αποδεικνύεται ότι το μητρώο $A^{\mathrm{T}}A$ είναι αντιστρέψιμο, βλέπε Θεώρημα 7.2, κατά συνέπεια

$$(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{u} = \boldsymbol{c} \Rightarrow A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{u} = A\boldsymbol{c}$$

επομένως

$$\operatorname{proj}_W \boldsymbol{u} = P\boldsymbol{u} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\boldsymbol{u}. \tag{7.15}$$

Παρατήρηση 7.5. Αν $A = \boldsymbol{e}_j$, $j = 1, 2, 3$, τότε $A^{\mathrm{T}}A = \begin{pmatrix}1\end{pmatrix}$ το οποίο ταυτίζουμε με τον αριθμό 1 και η (7.15) καταλήγει στην (7.13). Όμοια αν $A = \boldsymbol{a}$, η (7.15) καταλήγει στην (7.14) (γιατί;).

Θεώρημα 7.2. *Εάν το $n \times m$ μητρώο A έχει γραμμικά ανεξάρτητες στήλες, τότε το μητρώο $A^{\mathrm{T}}A$ είναι αντιστρέψιμο.*

Απόδειξη. Το μητρώο $A^{\mathrm{T}}A$ είναι τετραγωνικό, οπότε αρκεί να δείξουμε ότι $\operatorname{null} A^{\mathrm{T}}A = \{\mathbf{0}\}$. Αν $x \in \operatorname{null} A^{\mathrm{T}}A$, τότε

$$A^{\mathrm{T}}A\boldsymbol{x} = \mathbf{0} \Rightarrow \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\mathbf{0} \Rightarrow (A\boldsymbol{x})^{\mathrm{T}}A\boldsymbol{x} = 0 \Rightarrow \|A\boldsymbol{x}\|^2 = 0 \Rightarrow A\boldsymbol{x} = 0.$$

Τώρα αν $A = \begin{pmatrix}\boldsymbol{a}_1 & \boldsymbol{a}_2 & \cdots & \boldsymbol{a}_m\end{pmatrix}$, και $\boldsymbol{x} = \begin{pmatrix}x_1 & x_2 & \cdots & x_n\end{pmatrix}^{\mathrm{T}}$ η σχέση $A\boldsymbol{x} = \mathbf{0}$ είναι ισοδύναμη με την $\boldsymbol{a}_1x_1 + \boldsymbol{a}_2x_2 + \cdots + \boldsymbol{a}_mx_m = \mathbf{0}$ από την οποία έπεται, επειδή οι στήλες είναι γραμμικά ανεξάρτητες, ότι $x_1 = x_2 = \cdots = x_m = 0$, συνεπώς $\boldsymbol{x} = \mathbf{0}$, άρα $\operatorname{null} A = \{\mathbf{0}\}$ που είναι το ζητούμενο. □

Παρατήρηση 7.6. Σε σχέση με το Θεώρημα 7.2 ισχύει ότι τα μητρώα A και $A^{\mathrm{T}}A$ έχουν τον ίδιο μηδενόχωρο, δηλαδή $\operatorname{null} A = \operatorname{null} A^{\mathrm{T}}A$. Πράγματι από τη σχέση

$$A\boldsymbol{x} = \mathbf{0} \Rightarrow A^{\mathrm{T}}A\boldsymbol{x} = A^{\mathrm{T}}\mathbf{0} = \mathbf{0}$$

έπεται ότι $\operatorname{null} A \subseteq \operatorname{null} A^{\mathrm{T}}A$, ενώ από την

$$A^{\mathrm{T}}A\boldsymbol{x} = \mathbf{0} \Rightarrow \boldsymbol{x}^{\mathrm{T}}A^{\mathrm{T}}A\boldsymbol{x} = \boldsymbol{x}^{\mathrm{T}}\mathbf{0} \Rightarrow (A\boldsymbol{x})^{\mathrm{T}}A\boldsymbol{x} = 0 \Rightarrow \|A\boldsymbol{x}\|^2 = 0 \Rightarrow A\boldsymbol{x} = 0$$

έπεται ότι $\operatorname{null} A^{\mathrm{T}}A \subseteq \operatorname{null} A$.

Ορισμός 7.2 (Θεώρημα). Αν W είναι ένας υπόχωρος του $\mathbb{R}^m$, $\{\boldsymbol{a}_1, \boldsymbol{a}_2, \dots, \boldsymbol{a}_n\}$ είναι μια βάση για τον W ($n \le m$), και A είναι το $m \times n$ μητρώο με στήλες τα διανύσματα $\boldsymbol{a}_j$, $j = 1, 2, \dots, n$, το $m \times m$ μητρώο

$$P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} \tag{7.16}$$

είναι το μητρώο της ορθογώνιας προβολής επί του W.

Παρατήρηση 7.7. Εάν P είναι μητρώο ορθογώνιας προβολής επί ενός υπόχωρου W του $\mathbb{R}^m$, τότε από την (7.14) ή την (7.15), ανάλογα την περίπτωση, προκύπτει ότι αφενός το P είναι συμμετρικό, δηλαδή $P^{\mathrm{T}} = P$ και αφετέρου ισχύει ότι $P^2 = P$. Αν $\boldsymbol{u}$ είναι διάνυσμα του $\mathbb{R}^m$, τότε

$$\boldsymbol{u} = P\boldsymbol{u} + \boldsymbol{u} - P\boldsymbol{u} = P\boldsymbol{u} + (I - P)\boldsymbol{u}$$

ισοδύναμα

$$\mathbb{R}^m = \text{range}(P) + \text{range}(I - P), \tag{7.17}$$

όπου I είναι το ταυτοτικό μητρώο. Το μητρώο $I - P$ είναι επίσης μητρώο ορθογώνιας προβολής, αφού

$$(I - P)^{\mathrm{T}} = I^{\mathrm{T}} - P^{\mathrm{T}} = I - P$$

και

$$\begin{aligned}(I - P)^2 &= (I - P)(I - P)\\ &= I^2 - IP - PI + P^2\\ &= I - 2P + P\\ &= I - P.\end{aligned}$$

Συγκεκριμένα προβάλλει κάθε διάνυσμα επί του ορθογωνίου συμπληρώματος $W^{\perp}$ του W, ισοδύναμα $\text{range}(I - P) = \text{range}(P)^{\perp}$. Πράγματι αν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα του $\mathbb{R}^m$, τότε

$$\begin{aligned}\langle P\boldsymbol{u}, (I - P)\boldsymbol{v}\rangle &= \langle P^2\boldsymbol{u}, (I - P)\boldsymbol{v}\rangle && (P^2\boldsymbol{u} = P\boldsymbol{u})\\ &= \langle P\boldsymbol{u}, P^{\mathrm{T}}(\boldsymbol{v} - P\boldsymbol{v})\rangle && (P^{\mathrm{T}} = P)\\ &= \langle P\boldsymbol{u}, P(\boldsymbol{v} - P\boldsymbol{v})\rangle\\ &= \langle P\boldsymbol{u}, P\boldsymbol{v} - P^2\boldsymbol{v}\rangle\\ &= \langle P\boldsymbol{u}, P\boldsymbol{v}\rangle - \langle P\boldsymbol{u}, P\boldsymbol{v}\rangle && (P^2\boldsymbol{v} = P\boldsymbol{v})\\ &= 0\end{aligned}$$

γεγονός που αποδεικνύει τον ισχυρισμό μας. Επειδή

$$\begin{gathered}\mathbb{R}^m = \text{range}(P) \oplus \text{range}(P)^{\perp}\\ \mathbb{R}^m = \text{range}(P) \oplus \text{null}(P^{\mathrm{T}}) = \text{range}(P) \oplus \text{null}(P)\\ \text{range}(P)^{\perp} = \text{null}(P^{\mathrm{T}}) = \text{null}(P),\end{gathered}$$

έχουμε τελικά ότι $\text{range}(I - P) = \text{null}(P)$.

Σημείωση 7.1. Μπορούμε ανεξάρτητα από το Θεμελιώδες Θεώρημα, με χρήση των ιδιοτήτων του μητρώου προβολής, να δείξουμε ότι $\text{range}(I - P) = \text{null}(P) = \text{range}(P)^{\mathrm{T}}$. Πράγματι

$$\begin{aligned}\boldsymbol{u} \in \text{range}(I - P) &\Rightarrow \boldsymbol{u} = (I - P)\boldsymbol{x} = \boldsymbol{x} - P\boldsymbol{x} && \text{για κάποιο } \boldsymbol{x} \in \mathbb{R}^m\\ &\Rightarrow P\boldsymbol{u} = P(\boldsymbol{x} - P\boldsymbol{x}) = P\boldsymbol{x} - P^2\boldsymbol{x} = P\boldsymbol{x} - P\boldsymbol{x} = \mathbf{0}\\ &\Rightarrow \boldsymbol{u} \in \text{null}(P)\end{aligned}$$

επομένως $\text{range}(I - P) \subseteq \text{null}(P)$. Όμοια

$$\begin{aligned}\boldsymbol{u} \in \text{null}(P) &\Rightarrow P\boldsymbol{u} = \mathbf{0}\\ &\Rightarrow \boldsymbol{u} = \boldsymbol{u} - P\boldsymbol{u} = (I - P)\boldsymbol{u}\\ &\Rightarrow \boldsymbol{u} \in \text{range}(I - P)\end{aligned}$$

οπότε $\text{null}(P) \subseteq \text{range}(I - P)$, κατά συνέπεια $\text{range}(I - P) = \text{null}(P)$. Επίσης από την ισότητα

$$\langle \boldsymbol{u}, P\boldsymbol{x}\rangle = \langle P^{\mathrm{T}}\boldsymbol{u}, \boldsymbol{x}\rangle = \langle P\boldsymbol{u}, \boldsymbol{x}\rangle,$$

όπου $\boldsymbol{x}$ είναι τυχαίο διάνυσμα στον $\mathbb{R}^m$, έπεται, θέτοντας την παραπάνω ισότητα ίση με μηδέν, μέσω της Πρότασης 5.1 ότι $\boldsymbol{u} \in \text{range}(P)^{\perp}$ αν και μόνο αν $P\boldsymbol{u} = \mathbf{0}$, ισοδύναμα $\text{range}(P)^{\perp} = \text{null}(P)$.

Παράδειγμα 7.2. Το μητρώο

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 0 \\ -1 & 1 \end{pmatrix}$$

έχει γραμμικά ανεξάρτητες στήλες. Υπολογίζοντας

$$A^{\mathrm{T}}A = \begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 2 & 0 \\ -1 & 1 \end{pmatrix} = \begin{pmatrix} 6 & 0 \\ 0 & 2 \end{pmatrix}$$

βλέπουμε, όπως εξασφαλίζει το Θεώρημα 7.2, ότι το $A^{\mathrm{T}}A$ είναι αντιστρέψιμο και

$$(A^{\mathrm{T}}A)^{-1} = \frac{1}{12}\begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix}.$$

Αντίθετα το 3×3 μητρώο AA^{T} δεν αντιστρέφεται αφού $\operatorname{rank}(AA^{\mathrm{T}}) \le \operatorname{rank} A = 2$. Έτσι το μητρώο

$$\begin{aligned} P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} &= \frac{1}{6}\begin{pmatrix} 1 & 1 \\ 2 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix} \\ &= \frac{1}{3}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix} \end{aligned}$$

είναι η προβολή επί του χώρου στηλών του A, ισοδύναμα για κάθε $\boldsymbol{u} \in \mathbb{R}^3$ υπάρχουν σταθερές λ και μ ώστε

$$P\boldsymbol{u} = A\begin{pmatrix} \lambda \\ \mu \end{pmatrix}.$$

Ισοδύναμα αν $\boldsymbol{u} = \begin{pmatrix} x & y & z \end{pmatrix}$, τότε

$$\frac{1}{3}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \mu\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda + \mu \\ 2\lambda \\ -\lambda + \mu \end{pmatrix}$$

απ΄ όπου επιλύοντας βρίσκουμε

$$\lambda = \frac{1}{6}(x + 2y - z), \qquad \mu = \frac{1}{2}(x + z).$$

Παράδειγμα 7.3. Θεωρούμε το επίπεδο W με εξίσωση $ax + by + cz = 0$ στο $\mathbb{R}^3$. Να βρεθεί η προβολή του τυχαίου διανύσματος $\boldsymbol{\xi}$ του $\mathbb{R}^3$ επί του W καθώς και το μητρώο προβολής επί του W.

Διαιρώντας την εξίσωση με $(a^2 + b^2 + c^2)^{1/2}$ μπορούμε να υποθέσουμε ότι το διάνυσμα

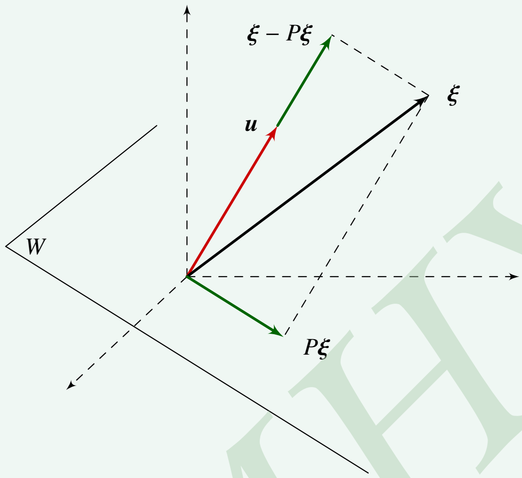
$$\boldsymbol{u} = \begin{pmatrix} a & b & c \end{pmatrix}^{\mathrm{T}}$$

είναι μοναδιαίο, θυμίζουμε δε ότι το $\boldsymbol{u}$ είναι ορθογώνιο στο επίπεδο W, βλέπε Παράδειγμα 5.8. Για τη λύση εργαζόμαστε γεωμετρικά. Αν με $P\boldsymbol{\xi}$ συμβολίσουμε την προβολή του $\boldsymbol{\xi}$ επί του W, βλέπε Σχήμα 7.4, το διάνυσμα $\boldsymbol{\xi} - P\boldsymbol{\xi}$ είναι η προβολή του $\boldsymbol{\xi}$ στο $\boldsymbol{u}$, κατά συνέπεια

$$\boldsymbol{\xi} - P\boldsymbol{\xi} = \langle \boldsymbol{\xi}, \boldsymbol{u} \rangle \boldsymbol{u}$$

επομένως

$$P\boldsymbol{\xi} = \boldsymbol{\xi} - \langle \boldsymbol{\xi}, \boldsymbol{u} \rangle \boldsymbol{u}.$$



Σχήμα 7.4: Προβολή διανύσματος επί του επιπέδου $ax + by + cz = 0$.

Για δε το μητρώο της προβολής P έχουμε

$$\begin{aligned} P\boldsymbol{\xi} &= I\boldsymbol{\xi} - \boldsymbol{u}\langle \boldsymbol{\xi}, \boldsymbol{u} \rangle \\ &= I\boldsymbol{\xi} - \boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}\boldsymbol{\xi} \\ &= (I - \boldsymbol{u}\boldsymbol{u}^{\mathrm{T}})\boldsymbol{\xi}, \end{aligned}$$

ισοδύναμα $P = I - \boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$.

► **Άσκηση 7.2.** Για το μητρώο $P = I - \boldsymbol{u}\boldsymbol{u}^{\mathrm{T}}$ του Παραδείγματος 7.3 δείξτε ότι $P^{\mathrm{T}} = P$ και $P^2 = P$.

Σημείωση 7.2. Για ένα μητρώο προβολής P η ιδιότητα $P^2 = P$ είναι χαρακτηριστική, μάλιστα μπορεί να αποδειχθεί ότι ένα μητρώο P είναι μητρώο προβολής επί ενός υποχώρου, του range P αν και μόνον αν $P^2 = P$, βλέπε για παράδειγμα [4] (Section 41, Theorem 1). Η ιδιότητα $P^{\mathrm{T}} = P$ χαρακτηρίζει την ορθογωνιότητα. Για παράδειγμα το μητρώο

$$P_\lambda = \begin{pmatrix} 0 & 0 \\ \lambda & 1 \end{pmatrix}, \qquad \lambda \in \mathbb{R}$$

ικανοποιεί την σχέση $P_\lambda^2 = P_\lambda$ κατά συνέπεια είναι μητρώο προβολής, αλλά $P_\lambda^{\mathrm{T}} \neq P_\lambda$ αν $\lambda \neq 0$. Παρατηρούμε ότι

$$\begin{pmatrix} 0 & 0 \\ \lambda & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 0 \\ \lambda x + y \end{pmatrix},$$

επομένως

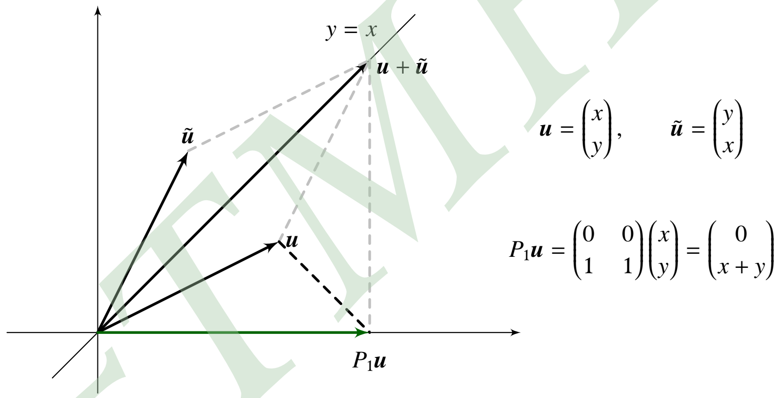
$$\operatorname{range} P_\lambda = \left\{\begin{pmatrix}0\\x\end{pmatrix} : x \in \mathbb{R}\right\} = \operatorname{span}\left\{\begin{pmatrix}0\\1\end{pmatrix}\right\},$$
$$\operatorname{null} P_\lambda = \left\{\begin{pmatrix}x\\-\lambda x\end{pmatrix} : x \in \mathbb{R}\right\} = \operatorname{span}\left\{\begin{pmatrix}1\\-\lambda\end{pmatrix}\right\},$$

άρα, όπως περιμένουμε άλλωστε, $\mathbb{R}^2 = (\operatorname{range} P_\lambda) \oplus (\operatorname{null} P_\lambda)$. Ελέγχουμε αν $\operatorname{range} P_\lambda \perp \operatorname{null} P_\lambda$, όπως συμβαίνει στην περίπτωση της ορθογώνιας προβολής. Έχουμε λοιπόν

$$\begin{pmatrix}0\\x\end{pmatrix} \cdot \begin{pmatrix}y\\-\lambda y\end{pmatrix} = -\lambda xy = 0, \qquad \forall x \quad \forall y$$

αν και μόνο αν $\lambda = 0$. Συνεπώς ο μηδενόχωρος $\operatorname{null} P$ δεν είναι το ορθογώνιο συμπλήρωμα του $\operatorname{range} P$ αν $\lambda \neq 0$. Στο Σχήμα 7.5 δείχνεται η υλοποίηση της προβολής στην περίπτωση όπου $\lambda = 1$.

Ο χαρακτηρισμός ορθογώνια για μια προβολή P εκφράζει ακριβώς το γεγονός ότι ο χώρος στηλών του P, $\operatorname{range} P$, και ο μηδενόχωρος του P, $\operatorname{null} P = \operatorname{range}(I - P)$, είναι ορθογώνια συμπληρώματα το ένα του άλλου.



Σχήμα 7.5: Υλοποίηση της μη ορθογώνιας προβολής $P_1\boldsymbol{u}$, ως ορθογώνιας προβολής επί του x-άξονα του διανύσματος $\boldsymbol{u} + \tilde{\boldsymbol{u}}$. Τα σημεία (x, y) και (y, x) είναι συμμετρικά ως προς την ευθεία $y = x$.

▶ **Άσκηση 7.3.** Να βρεθεί γεωμετρικά η προβολή του τυχαίου $\boldsymbol{u} \in \mathbb{R}^2$ μέσω του μητρώου P_λ της Σημείωσης 7.2 για γενικό λ. **Υπόδειξη:** Από το $\boldsymbol{u} = \begin{pmatrix}x & y\end{pmatrix}^{\mathrm{T}}$ κατασκευάστε το $\boldsymbol{u}^\lambda = \begin{pmatrix}\lambda x & y\end{pmatrix}^{\mathrm{T}}$.

Θεώρημα 7.3. *Μια προβολή P, ένα μητρώο δηλαδή με την ιδιότητα $P^2 = P$, είναι ορθογώνια αν και μόνο αν $P^{\mathrm{T}} = P$.*

Απόδειξη. Δείχνουμε πρώτα ότι $P^{\mathrm{T}} = P \Rightarrow \operatorname{range} P \perp \operatorname{null} P$. Για διανύσματα $\boldsymbol{u}, \boldsymbol{v}$, όπως ακριβώς στην Παρατήρηση 7.7 έχουμε ότι $\langle P\boldsymbol{u}, (I - P)\boldsymbol{v}\rangle = 0$, που είναι το ζητούμενο.
Στη συνέχεια δείχνουμε ότι $\operatorname{range} P \perp \operatorname{null} P \Rightarrow P^{\mathrm{T}} = P$. Για διανύσματα $\boldsymbol{u}, \boldsymbol{v}$ από την υπόθεση έχουμε ότι

$$\langle P\boldsymbol{u}, \boldsymbol{v} - P\boldsymbol{v}\rangle = \langle \boldsymbol{u} - P\boldsymbol{u}, P\boldsymbol{v}\rangle = 0$$

επομένως από τις ιδιότητες του εσωτερικού γινομένου παίρνουμε

$$\langle P\boldsymbol{u}, \boldsymbol{v}\rangle - \langle P\boldsymbol{u}, P\boldsymbol{v}\rangle = \langle \boldsymbol{u}, P\boldsymbol{v}\rangle - \langle P\boldsymbol{u}, P\boldsymbol{v}\rangle = 0 \Rightarrow \langle P\boldsymbol{u}, \boldsymbol{v}\rangle = \langle \boldsymbol{u}, P\boldsymbol{v}\rangle.$$

Επειδή από τον ορισμό του εσωτερικού γινομένου είναι

$$\langle P\boldsymbol{u}, \boldsymbol{v}\rangle = \langle \boldsymbol{u}, P^{\mathrm{T}}\boldsymbol{v}\rangle$$

από την τελευταία συνεπαγωγή παίρνουμε

$$\langle \boldsymbol{u}, P\boldsymbol{v}\rangle = \langle \boldsymbol{u}, P^{\mathrm{T}}\boldsymbol{v}\rangle \Rightarrow \langle \boldsymbol{u}, P\boldsymbol{v} - P^{\mathrm{T}}\boldsymbol{v}\rangle = 0.$$

Η τελευταία ισότητα ισχύει για κάθε $\boldsymbol{u}$, άρα και για $\boldsymbol{u} = P\boldsymbol{v} - P^{\mathrm{T}}\boldsymbol{v}$, οπότε

$$\langle P\boldsymbol{v} - P^{\mathrm{T}}\boldsymbol{v}, P\boldsymbol{v} - P^{\mathrm{T}}\boldsymbol{v}\rangle = 0$$

για κάθε $\boldsymbol{v}$, απ΄ όπου προκύπτει ότι $P\boldsymbol{v} - P^{\mathrm{T}}\boldsymbol{v} = 0$, ισοδύναμα $P = P^{\mathrm{T}}$. □

7.4 Ασκήσεις

1. Εάν για ένα $n \times n$ μητρώο A ισχύει $A^2 = A$ και $\operatorname{rank} A = n$, δείξτε ότι $A = I$.

2. Έστω A ένα $n \times 3$ μητρώο του οποίου οι στήλες είναι ορθογώνιες.

   (α΄) Εάν η κάθε στήλη έχει μέτρο 4, ποιο είναι το μητρώο $A^{\mathrm{T}}A$;

   (β΄) Εάν οι στήλες έχουν μέτρα 1, 2 και 3 αντίστοιχα, ποιο είναι το μητρώο $A^{\mathrm{T}}A$;

3. **Δυϊκό του Θεωρήματος 7.2.** Εάν το $n \times m$ μητρώο A έχει γραμμικά ανεξάρτητες γραμμές, τότε το μητρώο AA^{T} είναι αντιστρέψιμο.

4. Να βρεθεί το μητρώο ορθογώνιας προβολής P επί του υπόχωρου W ο οποίος παράγεται από τα διανύσματα

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}.$$

5. Δίνονται τα διανύσματα

$$\boldsymbol{a}_1 = \begin{pmatrix} -1 & 2 & -1 \end{pmatrix}^{\mathrm{T}}, \qquad \boldsymbol{a}_2 = \begin{pmatrix} 2 & 2 & 2 \end{pmatrix}^{\mathrm{T}}.$$

   (α΄) Να βρεθούν τα μητρώα ορθογώνιας προβολής P_1 και P_2 αντίστοιχα επί των ευθειών δια των $\boldsymbol{a}_1$ και $\boldsymbol{a}_2$.

   (β΄) Δείξτε ότι αν $\boldsymbol{u} \in \operatorname{range} P_1$ και $\boldsymbol{v} \in \operatorname{range} P_2$, τότε $\boldsymbol{u} \perp \boldsymbol{v}$.

   (γ΄) Υπολογίστε το γινόμενο P_1P_2 και εξηγήστε γιατί είναι αυτό που είναι.

6. Αν $\boldsymbol{p}$ είναι ένα μοναδιαίο διάνυσμα στον $\mathbb{R}^n$ και λ είναι μια πραγματική σταθερά θεωρούμε το μητρώο $P = I - \lambda \boldsymbol{p}\boldsymbol{p}^{\mathrm{T}}$, όπου I είναι το $n \times n$ ταυτοτικό μητρώο. Για ποια λ είναι το P

   (α΄) Συμμετρικό.

   (β΄) Ορθογώνιο.

   (γ΄) Μητρώο ορθογώνιας προβολής.

Κεφάλαιο 8

Ορίζουσες

Την ορίζουσα ενός 2×2 μητρώου συναντήσαμε σε προηγούμενα κεφάλαια, βλέπε Παράδειγμα 2.15, ή Σημείωση 4.1, ή Παρατήρηση 3.6. Γενικά η ορίζουσα (determinant) είναι μία συνάρτηση ορισμένη στο χώρο των τετραγωνικών μητρώων $\mathbb{M}^{n,n}(\mathbb{F})$ με τιμές στο σώμα των σταθερών $\mathbb{F}$, όπου $\mathbb{F}=\mathbb{R}$, ή $\mathbb{F}=\mathbb{C}$, σχηματικά

$$\det : \mathbb{M}^{n,n} \to \mathbb{F}, \qquad \mathbb{M}^{n,n} \ni A \mapsto \det A \in \mathbb{F}.$$

Αν $A = \begin{pmatrix} a_{ij} \end{pmatrix}$ είναι ένα $n\times n$ μητρώο, γράφουμε

$$\det A = |A| = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix}.$$

Θα συμβολίζουμε την i γραμμή του μητρώου με r_i, έτσι $r_i = \begin{pmatrix} a_{i1} & a_{i2} & \cdots & a_{in} \end{pmatrix}$. Υπάρχουν διάφοροι τρόποι να ορισθεί η ορίζουσα μητρώου. Επιλέγουμε έναν επαγωγικό ο οποίος την ίδια στιγμή δίνει και την αριθμητική τιμή της ορίζουσας μητρώου. Ξεκινάμε με την 2×2 περίπτωση, όπου με ευκολία επιβεβαιώνουμε τις ιδιότητες της ορίζουσας οι οποίες την χαρακτηρίζουν μοναδικά.

8.1 Η 2×2 ορίζουσα

Ορισμός 8.1. Για το μητρώο

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

ορίζουμε ως **ορίζουσα** (determinant) του A το βαθμωτό μέγεθος $ad - bc$ και γράφουμε

$$\det A = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = ad - bc.$$

Την ορίζουσα του A συμβολίζουμε και με $|A|$.

Είδαμε ότι η ορίζουσα ενός μητρώου A παίζει έναν ιδιαίτερο ρόλο και δίνει διάφορες πληροφορίες για το μητρώο. Για παράδειγμα, είδαμε ότι αν $\det A \neq 0$, τότε υπάρχει το αντίστροφο μητρώο A^{-1}, βλέπε Άσκηση 2.8 και Παράδειγμα 3.7, επιπλέον

$$A^{-1} = \frac{1}{\det A}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}.$$

Στο Παράδειγμα 2.15 δείξαμε επίσης ότι για 2×2 μητρώα A και B ισχύει $\det(AB) = (\det A)(\det B)$, έτσι, αφού

$$\det I = \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = (1)(1) - (0)(0) = 1,$$

εφόσον $\det A \neq 0$ βρίσκουμε

$$\begin{aligned} AA^{-1} = I &\Rightarrow \det(AA^{-1}) = \det I \\ &\Rightarrow (\det A)(\det A^{-1}) = 1 \\ &\Rightarrow \det A^{-1} = \frac{1}{\det A}. \end{aligned}$$

Συμπεραίνουμε λοιπόν ότι η ορίζουσα ενός 2×2, αλλά και γενικότερα ενός $n \times n$ μητρώου όπως θα δούμε, δίνει πληροφορία για την αντιστρεψιμότητα του μητρώου, ισοδύναμα για την γραμμική ανεξαρτησία των στηλών του μητρώου, ισοδύναμα για την τάξη του μητρώου.

Ιδιότητες της ορίζουσας

(ο1) Η ορίζουσα του ταυτοτικού μητρώου είναι ίση με 1.

(ο2) Η ορίζουσα μητρώου αλλάζει πρόσημο εάν εναλλάξουμε τις γραμμές του μητρώου.

$$\begin{vmatrix} c & d \\ a & b \end{vmatrix} = cb - da = -(ad - bc) = -\begin{vmatrix} a & b \\ c & d \end{vmatrix}.$$

(ο3) Η ορίζουσα είναι γραμμική ως προς την πρώτη γραμμή του μητρώου.
Πράγματι

$$\begin{aligned} \begin{vmatrix} \lambda a + \mu p & \lambda b + \mu q \\ c & d \end{vmatrix} &= (\lambda a + \mu p)d - (\lambda b + \mu q)c \\ &= \lambda(ad - bc) + \mu(pd - qc) \\ &= \lambda\begin{vmatrix} a & b \\ c & d \end{vmatrix} + \mu\begin{vmatrix} p & q \\ c & d \end{vmatrix}. \end{aligned}$$

(ο4) Η ορίζουσα μητρώου είναι ίση με την ορίζουσα του ανάστροφου μητρώου.
Πράγματι αν

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

τότε υπολογίζουμε

$$\det A^{\mathrm{T}} = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = ad - bc = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \det A.$$

Παρατήρηση 8.1. Από τις ιδιότητες (ο2) και (ο3) έπεται ότι η ορίζουσα είναι γραμμική ως προς την δεύτερη γραμμή του μητρώου, αφού

$$\begin{aligned}\begin{vmatrix} a & b \\ \lambda c + \mu p & \lambda d + \mu q \end{vmatrix} &= -\begin{vmatrix} \lambda c + \mu p & \lambda d + \mu q \\ a & b \end{vmatrix} \\ &= -\left(\lambda \begin{vmatrix} c & d \\ a & b \end{vmatrix} + \mu \begin{vmatrix} p & q \\ a & b \end{vmatrix}\right) \\ &= \lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \mu \begin{vmatrix} a & b \\ p & q \end{vmatrix}.\end{aligned}$$

Συνεπώς η ορίζουσα είναι γραμμική ως προς την πρώτη ή δεύτερη γραμμή του μητρώου κρατώντας την άλλη γραμμή σταθερή. Επίσης από την ιδιότητα (ο4) έπεται ότι οποιοδήποτε συμπέρασμα αφορά στις γραμές μητρώου ισχύει και για τις στήλες. Για παράδειγμα εναλλάσοντας δύο στήλες αλλάζει το πρόσημο

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \begin{vmatrix} a & c \\ b & d \end{vmatrix} = -\begin{vmatrix} b & d \\ a & c \end{vmatrix} = -\begin{vmatrix} b & a \\ d & c \end{vmatrix}.$$

και όσον αφορά στη γραμμικότητα

$$\begin{aligned}\begin{vmatrix} \lambda a + \mu p & b \\ \lambda c + \mu q & d \end{vmatrix} &= \begin{vmatrix} \lambda a + \mu p & \lambda c + \mu q \\ b & d \end{vmatrix} \\ &= \lambda \begin{vmatrix} a & c \\ b & d \end{vmatrix} + \mu \begin{vmatrix} p & q \\ b & d \end{vmatrix} \\ &= \lambda \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \mu \begin{vmatrix} p & b \\ q & d \end{vmatrix}.\end{aligned}$$

Συνοψίζοντας αν $\boldsymbol{a}$, $\boldsymbol{b}$ και $\boldsymbol{c}$ είναι διανύσματα στο $\mathbb{R}^2$, τότε

$$\det\big(\lambda \boldsymbol{a} + \mu \boldsymbol{b} \;\; \boldsymbol{c}\big) = \lambda \det\big(\boldsymbol{a} \;\; \boldsymbol{c}\big) + \mu \det\big(\boldsymbol{b} \;\; \boldsymbol{c}\big), \qquad \det\big(\boldsymbol{a} \;\; \lambda \boldsymbol{b} + \mu \boldsymbol{c}\big) = \lambda \det\big(\boldsymbol{a} \;\; \boldsymbol{b}\big) + \mu \det\big(\boldsymbol{a} \;\; \boldsymbol{c}\big), \tag{8.1}$$

$$\det\begin{pmatrix} \lambda \boldsymbol{a}^{\mathrm{T}} + \mu \boldsymbol{b}^{\mathrm{T}} \\ \boldsymbol{c}^{\mathrm{T}} \end{pmatrix} = \lambda \det\begin{pmatrix} \boldsymbol{a}^{\mathrm{T}} \\ \boldsymbol{c}^{\mathrm{T}} \end{pmatrix} + \mu \det\begin{pmatrix} \boldsymbol{a}^{\mathrm{T}} \\ \boldsymbol{c}^{\mathrm{T}} \end{pmatrix}, \qquad \det\begin{pmatrix} \boldsymbol{a}^{\mathrm{T}} \\ \lambda \boldsymbol{b}^{\mathrm{T}} + \mu \boldsymbol{c}^{\mathrm{T}} \end{pmatrix} = \lambda \det\begin{pmatrix} \boldsymbol{a}^{\mathrm{T}} \\ \boldsymbol{b}^{\mathrm{T}} \end{pmatrix} + \mu \det\begin{pmatrix} \boldsymbol{a}^{\mathrm{T}} \\ \boldsymbol{c}^{\mathrm{T}} \end{pmatrix}, \tag{8.2}$$

γεγονός που εκφράζει ότι η ορίζουσα είναι γραμμική ως προς κάθε στήλη ξεχωριστά, ή ως προς κάθε γραμμή ξεχωριστά.

Παρατήρηση 8.2. Αναρωτιόμαστε αν είναι σωστό, και προκύπτει από τις ιδιότητες της ορίζουσας που αναφέραμε, ότι $\det(A+B) = \det A + \det B$. Ας το ελέγξουμε. Για

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

και λ μια σταθερά, από τη γραμμικότητα της ορίζουσας ως προς κάθε γραμμή ξεχωριστά, υπολογίζουμε

$$\det(\lambda A) = \begin{vmatrix} \lambda a & \lambda b \\ \lambda c & \lambda d \end{vmatrix} = \lambda \begin{vmatrix} a & b \\ \lambda c & \lambda d \end{vmatrix} = \lambda^2 \begin{vmatrix} a & b \\ c & d \end{vmatrix} = \lambda^2 \det A. \tag{8.3}$$

Έτσι για $\lambda = 2$ έχουμε

$$\det(A+A) = \det(2A) = 4\det A \neq 2\det A = \det A + \det A,$$

κατά συνέπεια

$$\det(A+B) \neq \det A + \det B.$$

Θεώρημα 8.1 (Ιδιότητες της ορίζουσας). *Έστω A ένα 2×2 μητρώο.*

($O1$) *Η ορίζουσα του ταυτοτικού μητρώου είναι ίση με* 1.

($O2$) *Εάν το μητρώο A' προκύπτει από το A με εναλλαγή των γραμμών του, τότε η ορίζουσα αλλάζει πρόσημο, δηλαδή* $\det A' = -\det A$.

($O3$) *Η ορίζουσα μητρώου ως συνάρτηση είναι γραμμική ξεχωριστά ως προς κάθε γραμμή κρατώντας την άλλη γραμμή σταθερή.*

($O4$) *Η ορίζουσα μητρώου είναι ίση με την ορίζουσα του ανάστροφου μητρώου.*

($O5$) *Εάν οι δύο γραμμές του A είναι ίσες, τότε* $\det A = 0$.

($O6$) *Εάν οι δύο γραμμές του A είναι γραμμικά εξαρτημένες, τότε* $\det A = 0$.

($O7$) *Εάν σε μια γραμμή του A προστεθεί το πολλαπλάσιο άλλης γραμμής, η τιμή της ορίζουσας δεν αλλάζει.*

Η αντίστοιχη εκδοχή για στήλη ή στήλες καθεμιάς από τις ιδιότητες ($O2$) – ($O3$) *και* ($O5$) – ($O7$) *ισχύει επίσης.*

Απόδειξη. Οι ιδιότητες (O1) – (O4) είναι αντίστοιχα οι (o1) – (o4). Η απόδειξη των υπολοίπων είναι τετριμμένη εάν χρησιμοποιήσουμε τον ορισμό της ορίζουσας. Δίνουμε, ωστόσο, μια διαφορετική απόδειξη η οποία χρησιμοποιεί τις ιδιότητες (O2) και (O3) και είναι ανεξάρτητη του μεγέθους.

(O5) Εάν οι γραμμές i και j του A, $i \neq j$, είναι ίσες, τότε εναλλάσοντάς τες το μητρώο παραμένει το ίδιο αλλά από την (O2) η ορίζουσα αλλάζει πρόσημο, έτσι θα έχουμε

$$\det A = -\det A$$

κατά συνέπεια $\det A = 0$.

(O6) Εάν η γραμμή j του A είναι πολλαπλάσιο της γραμμής i, $i \neq j$, έστω $r_j = \lambda r_i$, και A' είναι το μητρώο A στο οποίο η j γραμμή έχει αντικατασταθεί από την i, τότε το A' έχει δύο γραμμές ίσες, άρα $\det A' = 0$. Επίσης από γραμμικότητα, ιδιότητα (O3), παίρνουμε

$$\det A = \lambda \det A' = 0$$

που είναι το ζητούμενο.

(O7) Έστω A'' να είναι το μητρώο A στο οποίο η j γραμμή έχει αντικατασταθεί από την $r_j + \lambda r_i$ και έστω A' να είναι όπως στο (O6), τότε και πάλι από τη γραμμικότητα της ορίζουζας βρίσκουμε

$$\det A'' = \det A + \lambda \det A' = \det A,$$

αφού $\det A' = 0$, που είναι ό,τι θέλουμε να αποδείξουμε.

Συνέπεια της (O4) είναι ότι οτιδήποτε ισχύει για γραμμές ισχύει και για στήλες, ειδικά η για στήλη εκδοχή της (O3) είναι το περιεχόμενο της Παρατήρησης 8.1. □

Παράδειγμα 8.1. Αν

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = k$$

να βρεθούν, δίχως να εκτελεστούν πράξεις, οι τιμές των οριζουσών

(α′) $\begin{vmatrix} \lambda c & \lambda d \\ \lambda a & \lambda b \end{vmatrix}$ (β′) $\begin{vmatrix} a & b \\ \lambda a + \mu c & \lambda b + \mu d \end{vmatrix}$ (γ′) $\begin{vmatrix} a - \lambda c & b - \lambda d \\ \lambda a + c & \lambda b + d \end{vmatrix}$

όπου λ και μ είναι μια πραγματικές σταθερές.

(α′) Από τις ιδιότητες της ορίζουσας υπολογίζουμε

$$\begin{aligned} \begin{vmatrix} \lambda c & \lambda d \\ \lambda a & \lambda b \end{vmatrix} &= \lambda \begin{vmatrix} c & d \\ \lambda a & \lambda b \end{vmatrix} && \text{(από την (Ο3))} \\ &= \lambda^2 \begin{vmatrix} c & d \\ a & b \end{vmatrix} && \text{(από την (Ο3))} \\ &= \lambda^2(-1) \begin{vmatrix} a & b \\ c & d \end{vmatrix} && \text{(από την (Ο2))} \\ &= -\lambda^2 k. \end{aligned}$$

(β′)

$$\begin{aligned} \begin{vmatrix} a & b \\ \lambda a + \mu c & \lambda b + \mu d \end{vmatrix} &= \begin{vmatrix} a & b \\ \lambda a & \lambda b \end{vmatrix} + \begin{vmatrix} a & b \\ \mu c & \mu d \end{vmatrix} && \text{(από την (Ο2))} \\ &= \lambda \begin{vmatrix} a & b \\ a & b \end{vmatrix} + \mu \begin{vmatrix} a & b \\ c & d \end{vmatrix} && \text{(από την (Ο2))} \\ &= \lambda 0 + \mu k && \text{(από την (Ο6))} \\ &= \mu k. \end{aligned}$$

(γ′)

$$\begin{aligned} \begin{vmatrix} a - \lambda c & b - \lambda d \\ \lambda a + c & \lambda b + d \end{vmatrix} &= \begin{vmatrix} a & b \\ \lambda a + c & \lambda b + d \end{vmatrix} - \lambda \begin{vmatrix} c & d \\ \lambda a + c & \lambda b + d \end{vmatrix} \\ &= \lambda \begin{vmatrix} a & b \\ a & b \end{vmatrix} + \begin{vmatrix} a & b \\ c & d \end{vmatrix} - \lambda^2 \begin{vmatrix} c & d \\ a & b \end{vmatrix} - \lambda \begin{vmatrix} c & d \\ c & d \end{vmatrix} \\ &= \begin{vmatrix} a & b \\ c & d \end{vmatrix} + \lambda^2 \begin{vmatrix} a & b \\ c & d \end{vmatrix} \\ &= (1 + \lambda^2) k. \end{aligned}$$

8.2 Η $n \times n$ ορίζουσα

Ορισμός 8.2. Αν $A = (a_{ij})$ είναι ένα $n \times n$ μητρώο, με A_{ij} συμβολίζουμε το $(n-1) \times (n-1)$ μητρώο το οποίο προκύπτει από το A διαγράφοντας την i γραμμή και την j στήλη.

Παράδειγμα 8.2. Αν

$$A = \begin{pmatrix} 2 & -1 & 3 \\ 1 & 4 & -2 \\ 0 & -3 & 2 \end{pmatrix}$$

τότε

$$A_{11} = \begin{pmatrix} 4 & -2 \\ -3 & 2 \end{pmatrix}, \qquad A_{22} = \begin{pmatrix} 2 & 3 \\ 0 & 2 \end{pmatrix}, \qquad A_{23} = \begin{pmatrix} 2 & -1 \\ 0 & -3 \end{pmatrix}, \qquad A_{32} = \begin{pmatrix} 2 & 3 \\ 1 & -2 \end{pmatrix}.$$

Ορισμός 8.3. Αν $n \in \mathbb{N}$ ορίζουμε τη συνάρτηση $\det : \mathbb{M}^{n,n}(\mathbb{F}) \to \mathbb{F}$ επαγωγικά ως

(1) $n = 1$ και $A = (a_{11})$, τότε

$$\det A = a_{11}.$$

(2) $n > 1$ και $A = (a_{ij})$, τότε

$$\det A = (-1)^{1+1} a_{11} \det A_{11} + (-1)^{1+2} a_{12} \det A_{12} + \cdots + (-1)^{1+n} a_{1n} \det A_{1n}. \tag{8.4}$$

Από τον ορισμό λοιπόν έπεται ότι

1. Αν $n = 2$, οπότε

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix},$$

από το (1) υπολογίζουμε

$$\begin{aligned} \det A &= (-1)^{1+1} a_{11} \det(a_{22}) + (-1)^{1+2} a_{12} \det(a_{21}) \\ &= a_{11} a_{22} - a_{12} a_{21} \end{aligned}$$

δηλαδή στην περίπτωση $n = 2$ η συνάρτηση det συμφωνεί με την ορίζουσα τάξης δύο.

2. Αν $n = 3$, οπότε

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix},$$

υπολογίζουμε

$$\begin{aligned} \det A &= (-1)^{1+1} a_{11} \det A_{11} + (-1)^{1+2} a_{12} \det A_{12} + (-1)^{1+3} a_{13} \det A_{13} \\ &= a_{11} \begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12} \begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13} \begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix}, \end{aligned}$$

και το τελικό αποτέλεσμα προκύπτει από την περίπτωση $n = 2$. Κατά συνέπεια η έκφραση στην (8.4) γενικεύει την ορίζουσα τάξης δύο και μοιάζει να είναι ένας λογικός ορισμός για την ορίζουσα κάθε τάξης. Είναι όντως έτσι αλλά πρέπει να πειστούμε γι αυτό. Θέλουμε η ορίζουσα να ικανοποιεί τις ιδιότητες που ικανοποιεί η ορίζουσα τάξης δύο. Προς αυτή την κατεύθυνση, σαν πρώτο βήμα, έχουμε το

Θεώρημα 8.2. *Η συνάρτηση* $\det : \mathbb{M}^{n,n}(\mathbb{F}) \to \mathbb{F}$ *στον Ορισμό 8.3 ικανοποιεί τις ιδιότητες*

(1) *Αν I είναι το ταυτοτικό μητρώο, τότε* $\det I = 1$.

(2) *Η συνάρτηση* det *είναι γραμμική ως προς κάθε γραμμή ξεχωριστά.*

Στην ιδιότητα (2) εννοούμε ότι αν με r_j συμβολίσουμε την j-γραμμή του μητρώου A, ως διάνυσμα γραμμή, ώστε $A = \begin{pmatrix} r_1 & r_2 & \dots & r_n \end{pmatrix}^{\mathrm{T}}$, τότε

$$\det\begin{pmatrix} r_1 \\ \vdots \\ \lambda r + \mu r' \\ \vdots \\ r_n \end{pmatrix} = \lambda \det\begin{pmatrix} r_1 \\ \vdots \\ r \\ \vdots \\ r_n \end{pmatrix} + \mu \det\begin{pmatrix} r_1 \\ \vdots \\ r' \\ \vdots \\ r_n \end{pmatrix}$$

Απόδειξη του θεωρήματος. Αποδεικνύουμε με επαγωγή ότι η συνάρτηση det ικανοποιεί τις ιδιότητες (1), και και (2). Για $n = 2$ έχουμε ήδη δείξει ότι οι ιδιότητες αυτές ικανοποιούνται. Υποθέτουμε ότι ισχύουν για n και δείχνουμε ότι ισχύουν και για $n+1$. Για την (1) από την (8.4) υπολογίζουμε

$$\det I_{n+1} = 1 \det I_n = 1$$

από την υπόθεση της επαγωγής. Η (2) ικανοποιείται σχεδόν αυτόματα είτε στην περίπτωση όπου ο γραμμικός συνδυασμός είναι στην πρώτη γραμμή από την επιμεριστική ιδιότητα, είτε στην περίπτωση όπου ο γραμμικός συνδυασμός εμφανίζεται στη δεύτερη, ή τρίτη, ή $n + 1$ γραμμή, από την υπόθεση της επαγωγής. Κατά συνέπεια οι δύο αυτές ιδιότητες ισχύουν για κάθε φυσικό αριθμό n. □

Μια άμεση συνέπεια της ιδιότητας της γραμμικότητας ως προς κάθε γραμμή χωριστά μας δίνει το

▼ **Πόρισμα 8.1.** *Αν A είναι ένα τετραγωνικό μητρώο μια γραμμή του οποίου περιέχει μόνο το μηδέν, τότε η ορίζουσα του μητρώου είναι ίση με μηδέν.*

Απόδειξη. Πράγματι αν συμβολίσουμε με 0 τη μηδενική γραμμή, τότε

$$\det\begin{pmatrix} r_1 \\ \vdots \\ 0 \\ \vdots \\ r_n \end{pmatrix} = \det\begin{pmatrix} r_1 \\ \vdots \\ r - r \\ \vdots \\ r_n \end{pmatrix} = \det\begin{pmatrix} r_1 \\ \vdots \\ r \\ \vdots \\ r_n \end{pmatrix} - \det\begin{pmatrix} r_1 \\ \vdots \\ r \\ \vdots \\ r_n \end{pmatrix} = 0$$

που είναι το ζητούμενο. □

Στη συνέχεια θα αποδείξουμε ότι στον τύπο 8.3, ορισμό της ορίζουσας, το ανάπτυγμα ως προς την πρώτη γραμμή μπορεί να αντικατασταθεί από το ανάπτυγμα ως προς οποιαδήποτε γραμμή. Ως πρώτο βήμα αποδεικκνύουμε μια ειδική περίπτωση.

Λήμμα 8.1. *Έστω B ένα $n \times n$ μητρώο με $n \geq 2$ του οποίου η i-γραμμή είναι ίση με $e_k = \boldsymbol{e}_k^{\mathrm{T}}$, $1 \leq k \leq n$, όπου $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_n\}$ είναι η κανονική βάση στον $\mathbb{R}^n$, τότε*

$$\det B = (-1)^{i+k} \det B_{ik}.$$

Απόδειξη. Δίνουμε την απόδειξη με επαγωγή στο n. Για $n = 2$ το αποτέλεσμα ισχύει αφού το B είναι ένα από τα μητρώα

$$\begin{pmatrix} 1 & 0 \\ b_{21} & b_{22} \end{pmatrix}, \qquad \begin{pmatrix} 0 & 1 \\ b_{21} & b_{22} \end{pmatrix}, \qquad \begin{pmatrix} b_{11} & b_{12} \\ 1 & 0 \end{pmatrix}, \qquad \begin{pmatrix} b_{11} & b_{12} \\ 0 & 1 \end{pmatrix}.$$

Υποθέτουμε λοιπόν ότι για $n \geq 3$ το αποτέλεσμα ισχύει και δείχνουμε ότι ισχύει και για $n + 1$. Έστω ότι το $(n+1) \times (n+1)$ μητρώο B έχει την i-γραμμή ίση με e_k. Αν $i = 1$ το αποτέλεσμα έπεται αμέσως από την (8.4), οπότε θεωρούμε την περίπτωση όπου $1 < i \leq n + 1$, οπότε

$$\det B = \sum_{j=1}^{n+1} (-1)^{1+j} b_{1j} \det B_{1j}. \tag{8.5}$$

Μας ενδιαφέρει η $(i-1)$-γραμμή του $n \times n$ μητρώου B_{1j} η οποία προέρχεται από την i-γραμμή του B, $e_k = \boldsymbol{e}_k^{\mathrm{T}}$ από την οποία έχει διαγραφεί η j-στήλη.

Αν $j < k$ η γραμμή περιέχει μηδέν παντού εκτός από την $k-1$-στήλη στην οποία υπάρχει το 1.

Αν $j = k$ η γραμμή περιέχει μηδέν παντού.

Αν $j > k$ η γραμμή περιέχει μηδέν παντού εκτός από την k-στήλη στην οποία υπάρχει το 1.

Έτσι από την υπόθεση της επαγωγής έχουμε

$$\det B_{1j} = \begin{cases} (-1)^{(i-1)+(k-1)} \det B^{1j}_{(i-1)(k-1)}, & j < k \\ 0, & j = k \\ (-1)^{(i-1)+k} \det B^{1j}_{(i-1)k}, & j > k \end{cases}$$

όπου με $B^{1j}_{(i-1)l}$ συμβολίζουμε το μητρώο που προκύπτει από το B_{1j} διαγράφοντας την $(i-1)$-γραμμή και την l-στήλη. Μπορούμε έτσι να ξαναγράψουμε την (8.5) ως

$$\begin{aligned} \det B &= \sum_{j<k} (-1)^{1+j} b_{1j} (-1)^{(i-1)+(k-1)} \det B^{1j}_{(i-1)(k-1)} + \sum_{j>k} (-1)^{1+j} b_{1j} (-1)^{(i-1)+k} \det B^{1j}_{(i-1)k} \\ &= \sum_{j<k} (-1)^{j+i+k-1} b_{1j} \det B^{1j}_{(i-1)(k-1)} + \sum_{j>k} (-1)^{j+i+k} b_{1j} \det B^{1j}_{(i-1)k} \\ &= (-1)^{i+k} \Big[\sum_{j<k} (-1)^{j-1} b_{1j} \det B^{1j}_{(i-1)(k-1)} + \sum_{j>k} (-1)^{j} b_{1j} \det B^{1j}_{(i-1)k} \Big] \\ &= (-1)^{i+k} \Big[\sum_{j<k} (-1)^{1+j} b_{1j} \det B^{1j}_{(i-1)(k-1)} + \sum_{j>k} (-1)^{1+(j-1)} b_{1j} \det B^{1j}_{(i-1)k} \Big]. \end{aligned}$$

Συνειδητοποιώντας ότι το $B^{1j}_{(i-1)(k-1)}$, ή το $B^{1j}_{(i-1)k}$ είναι το $(n-2) \times (n-2)$ μητρώο το οποίο προκύπτει από το B διαγράφοντας την πρώτη και την i γραμμή και τις j και k στήλες με $j \neq k$, βλέπουμε ότι η έκφραση στην αγκύλη είναι η ορίζουσα του B_{ik}, που είναι ό,τι θέλουμε να δείξουμε. □

Θεώρημα 8.3 (το ανάπτυγμα της ορίζουσας ως προς γραμμή). *Εάν $A = \left(a_{ij}\right)$ είναι ένα $n \times n$ μητρώο με $n \geq 2$, τότε για κάθε i με $1 \leq i \leq n$ η ορίζουσα του A δίνεται από τη σχέση*

$$\det A = (-1)^{i+1} a_{i1} \det A_{i1} + (-1)^{i+2} a_{i2} \det A_{i2} + \cdots + (-1)^{i+n} a_{in} \det A_{in}. \tag{8.6}$$

Απόδειξη. Αν r_i είναι η i-γραμμή του A και $e_k = \boldsymbol{e}_k^{\mathrm{T}}$, $1 \leq k \leq n$, όπου $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \dots, \boldsymbol{e}_n\}$ είναι η κανονική βάση στον $\mathbb{R}^n$, τότε

$$r_i = a_{i1} e_1 + a_{i2} e_2 + \cdots + a_{in} e_n$$

και από την γραμμικότητα της ορίζουσας έχουμε

$$\det A = a_{i1} \det \begin{pmatrix} r_1 \\ \vdots \\ e_1 \\ \vdots \\ r_n \end{pmatrix} + a_{i2} \det \begin{pmatrix} r_1 \\ \vdots \\ e_2 \\ \vdots \\ r_n \end{pmatrix} + \cdots + a_{in} \det \begin{pmatrix} r_1 \\ \vdots \\ e_n \\ \vdots \\ r_n \end{pmatrix}$$

οπότε από το Λήμμα 8.1 έπεται το ζητούμενο. □

Θεώρημα 8.4. *Η συνάρτηση* $\det : \mathbb{M}^{n,n}(\mathbb{F}) \to \mathbb{F}$ *στον Ορισμό 8.3 ικανοποιεί τις ιδιότητες*

(i) *Αν I είναι το ταυτοτικό μητρώο, τότε* $\det I = 1$.

(ii) *Η συνάρτηση* $\det$ *είναι γραμμική ως προς κάθε γραμμή ξεχωριστά.*

(iii) *Αν το μητρώο A' προκύπτει από το A με εναλλαγή δύο γραμμών, τότε* $\det A' = -\det A$.

Επιπλέον η $\det$ *είναι η* **μοναδική** *συνάρτηση* $\mathbb{M}^{n,n}(\mathbb{F}) \mapsto \mathbb{F}$ *η οποία χαρακτηρίζεται από τις παραπάνω ιδιότητες.*

Απόδειξη. Η ισχύς των ιδιοτήτων (i) και (ii) είναι το περιεχόμενο του Θεωρήματος 8.2. Αποδεικνύουμε με επαγωγή ότι η συνάρτηση det ικανοποιεί και την ιδιότητα (iii). Για $n = 2$ το αποτέλεσμα ισχύει, είναι η (O2). Υποθέτουμε λοιπόν ότι η ιδιότητα ισχύει για n και δείχνουμε ότι ισχύει και για $n+1$. Αν εναλλάξουμε την j-γραμμή με την k-γραμμή με $k > j \geq 2$ το συμπέρασμα έπεται από την υπόθεση της επαγωγής αφού η εναλλαγή εμφανίζεται στις ορίζουσες τάξης n στο ανάπτυγμα (8.4). Απομένει λοιπόν να δείξουμε ότι η ιδιότητα ισχύει αν γίνει εναλλαγή της πρώτης με κάποια άλλη γραμμή. Εάν A' είναι το μητρώο που προκύπτει από το A εναλλάσοντας την πρώτη με την k-γραμμή, τότε μέσω του Θεωρήματος 8.3 εκφράζοντας την ορίζουσα ως προς μια γραμμή $j \neq 1, k$ έχουμε από την υπόθεση της επαγωγής ότι

$$\det A_{jl} = -\det A'_{jl}, \qquad l = 1, 2, \dots n+1$$

οπότε

$$\det A = \sum_{l=1}^{n+1} (-1)^{j+l} \det A_{jl} = \sum_{l=1}^{n+1} (-1)^{j+l} (-\det A'_{jl}) = -\det A'.$$

Μοναδικότητα. Δείχνουμε ότι αν det και det′ είναι δύο συναρτήσεις ορισμένες στο $\mathbb{M}^{n,n}(\mathbb{F})$ με τιμές στο $\mathbb{F}$, οι οποίες ικανοποιούν τις (i), (ii), και (iii), τότε $\det = \det'$, δηλαδή $\det A = \det' A$

για κάθε $A \in \mathbb{M}^{n,n}(\mathbb{F})$. Όπως στο Θεώρημα 8.1, το οποίο στην πραγματικότητα αποδείξαμε για κάθε διάσταση n (γιατί;) έχουμε ότι οι ιδιότητες (ii) και (iii) συνεπάγονται τις

(iv) Αν το μητρώο A έχει δύο ίσες γραμμές, τότε $\det A = 0$.

(v) Αν $\operatorname{rank} A < n$, τότε $\det A = 0$.

(vi) Αν το μητρώο A' προκύπτει από το A προσθέτοντας το πολλαπλάσιο μιας γραμμής σε άλλη γραμμή, τότε $\det A = \det A'$.

Κατά συνέπεια οι $\det$ και $\det'$ ικανοποιούν επί πλέον τις (iv), (v), και (vi). Έστω A ένα $n \times n$ μητρώο. Αν $\operatorname{rank} A < n$, τότε από την (v) έπεται ότι $\det A = 0$ και $\det' A = 0$, συνεπώς $\det A = \det' A$. Αν $\operatorname{rank} A = n$, τότε το A είναι γραμμοϊσοδύναμο με ένα διαγώνιο μητρώο με, έστω, $a_1, a_2, \ldots, a_n$ στη κύρια διαγώνιο. Τότε από γραμμικότητα και την (vi) έχουμε

$$\det A = a_1 a_2 \cdots a_n \det I, \qquad \text{και} \qquad \det' A = a_1 a_2 \cdots a_n \det' I$$

και επειδή $\det I = \det' I = 1$, έπεται ότι $\det A = \det' A$ και σε αυτή την περίπτωση. Συνεπώς $\det = \det'$. □

► **Άσκηση 8.1.** Δείξτε ότι η ιδιότητα (iii) στο Θεώρημα 8.4, δεδομένης της (ii), είναι, για την ακρίβεια ισοδύναμη με κάθε μια από τις (iv), (v), (vi) δείχνοντας τις συνεπαγωγές (iii) ⇒ (iv) ⇒ (v) ⇒ (vi) ⇒ (iii). Συμπεράνετε ότι η (iii) στην εκφώνηση του Θεωρήματος μπορεί να αντικατασταθεί με μία από τις παραπάνω τρεις ιδιότητες.

Στο Θεώρημα 8.3 είδαμε ότι η ορίζουσα μητρώου μπορεί να εκφραστεί ως ανάπτυγμα ως προς οποιαδήποτε γραμμή. Στο επόμενο θεώρημα δείχνουμε ότι μπορεί να εκφραστεί επίσης και ως προς οποιαδήποτε στήλη.

Θεώρημα 8.5 (το ανάπτυγμα της ορίζουσας ως προς στήλη). *Εάν $A = \left(a_{ij}\right)$ είναι ένα $n \times n$ μητρώο με $n \geq 2$, τότε για κάθε j με $1 \leq j \leq n$ η ορίζουσα του A δίνεται από τη σχέση*

$$\det A = (-1)^{1+j} a_{1j} \det A_{1j} + (-1)^{2+j} a_{2j} \det A_{2j} + \cdots + (-1)^{n+j} a_{nj} \det A_{nj}. \tag{8.7}$$

Απόδειξη. Δείχνουμε με επαγωγή ότι η $\det$ στην (8.7) ικανοποιεί τις ιδιότητες (i) – (iii) του Θεωρήματος 8.4, οπότε από μοναδικότητα η έκφραση (8.7) είναι η ορίζουσα του μητρώου A. Αρκεί να δείξουμε μόνο την περίπτωση $n = 2$ (γιατί;), η οποία είναι ουσιατικά η ιδιότητα (Ο4) αφού το ανάπτυγμα ως προς μια στήλη του μητρώου είναι το ανάπτυγμα ως προς την αντίστοιχη γραμμή του ανάστροφου μητρώου. □

Ορισμός 8.4. Για το μητρώο $A = \left(a_{ij}\right)$, τον όρο $(-1)^{i+j} \det A_{ij}$ στο ανάπτυγμα της ορίζουσας λέμε **συμπαράγοντα** (cofactor) του στοιχείου a_{ij}. Την έκφραση (8.6) λέμε ανάπτυγμα σε συμπαράγοντες ως προς την i-γραμμή του A και εκείνη στην (8.7) λέμε ανάπτυγμα ως προς την j-στήλη του A. Το ανάπτυγμα της ορίζουσας σε συμπαράγοντες

$$\det A = \sum_{j=1}^{n} (-1)^{i+j} \det A_{ij} = \sum_{i=1}^{n} (-1)^{i+j} \det A_{ij} \tag{8.8}$$

κατά μήκος μιας γραμμής, ή στήλης λέγεται **τύπος του Laplace**.

Παράδειγμα 8.3. Να υπολογισθεί η ορίζουσα του μητρώου

$$A = \begin{pmatrix} 2 & -1 & 3 & 1 \\ 1 & 0 & 2 & -2 \\ 4 & 0 & 1 & 3 \\ 2 & 0 & -1 & 3 \end{pmatrix}.$$

Είναι λογικό να επιλέξουμε να αναπτύξουμε ως προς την γραμμή ή την στήλη η οποία περιέχει τα περισσότερα μηδενικά. Επιλέγοντας την δεύτερη στήλη υπολογίζουμε

$$\begin{aligned}
\det A &= (-1)^{1+2}(-1)\det A_{12} + (-1)^{2+2}(0)\det A_{22} + (-1)^{3+2}(0)\det A_{32} + (-1)^{4+2}(0)\det A_{42} \\
&= \det A_{12} \\
&= \det \begin{pmatrix} 1 & 2 & -2 \\ 4 & 1 & 3 \\ 2 & -1 & 3 \end{pmatrix} \\
&= (-1)^{1+1}\det\begin{pmatrix} 1 & 3 \\ -1 & 3 \end{pmatrix} + (-1)^{1+2}(2)\det\begin{pmatrix} 4 & 3 \\ 2 & 3 \end{pmatrix} + (-1)^{1+3}(-2)\det\begin{pmatrix} 4 & 1 \\ 2 & -1 \end{pmatrix} \\
&= 6 - 2(6) - 2(-6) \\
&= 6.
\end{aligned}$$

Παράδειγμα 8.4. Να υπολογισθεί η ορίζουσα του μητρώου

$$A = \begin{pmatrix} 2 & 2 & 3 & 1 \\ 0 & 3 & 2 & 1 \\ 0 & 0 & 4 & 3 \\ 0 & 0 & 0 & 5 \end{pmatrix}.$$

Αναπτύσοντας ως προς την πρώτη στήλη υπολογίζουμε

$$\begin{aligned}
\det A &= (-1)^{1+1}(2)\det A_{11} \\
&= 2\det\begin{pmatrix} 3 & 2 & 1 \\ 0 & 4 & 3 \\ 0 & 0 & 5 \end{pmatrix} = (2)(3)\det\begin{pmatrix} 4 & 3 \\ 0 & 5 \end{pmatrix} = (2)(3)(4)(5).
\end{aligned}$$

Κατά συνέπεια η ορίζουσα του άνω τριγωνικού μητρώου είναι ίση με το γινόμενο των στοιχείων της (κύριας) διαγωνίου.

▲ **Πρόταση 8.1.** *Η ορίζουσα ενός άνω, ή κάτω τριγωνικού, ή διαγωνίου μητρώου είναι ίση με το γινόμενο των στοιχείων της διαγωνίου.*

Απόδειξη. Εάν $A = \left(a_{ij}\right)$ είναι ένα άνω, ή κάτω τριγωνικό, ή διαγώνο μητρώο, τότε αναπτύσοντας, ανάλογα, ως προς την πρώτη στήλη (στην περίπτωση άνω τριγωνικού), ή ως προς την πρώτη γραμμή (στην περίπτωση κάτω τριγωνικού) έχουμε

$$\det A = a_{11}\det A_{11}. \tag{8.9}$$

Θέτουμε $A^{(1)} = A_{11}$, $A^{(2)} = A_{11}^{(1)}$ και γενικά $A^{(k+1)} = A_{11}^{(k)}$, για $k = 1, \ldots, n-1$ αφού $A^{(n-1)} = \left(a_{nn}\right)$.

Κάθε μητρώο $A^{(k)}$ είναι του ιδίου τύπου με το A, κατά συνέπεια

$$\det A = a_{11}a_{22}\det A^{(2)} = a_{11}a_{22}\cdots a_{kk}\det A^{(k)}, \tag{8.10}$$

απ' όπου έπεται το συμπέρασμα για $k = n-1$, αφού $\det A^{(n-1)} = a_{nn}$. □

8.3 Ιδιότητες της ορίζουσας

Θεώρημα 8.6. *Εάν A είναι ένα τετραγωνικό μητρώο, τότε $\det A = \det A^{\mathrm{T}}$.*

Απόδειξη. Αποδεικνύουμε το ζητούμενο με επαγωγή. Για την ακρίβεια αποδεικνύουμε το αρχικό βήμα της επαγωγής. Το συμπέρασμα ισχύει για 2×2 μητρώα, είναι η ιδιότητα (O4). Αν το A είναι 3×3 και γράψουμε $A^{\mathrm{T}} = A' = \left(a'_{ij}\right)$ παρατηρούμε ότι

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \Rightarrow A^{\mathrm{T}} = A' = \begin{pmatrix} a_{11} & a_{21} & a_{31} \\ a_{12} & a_{22} & a_{32} \\ a_{13} & a_{23} & a_{33} \end{pmatrix}$$

οπότε

$$A'_{11} = \begin{pmatrix} a_{22} & a_{32} \\ a_{23} & a_{33} \end{pmatrix} = (A_{11})^{\mathrm{T}}, \qquad A'_{21} = \begin{pmatrix} a_{21} & a_{31} \\ a_{23} & a_{33} \end{pmatrix} = (A_{12})^{\mathrm{T}}, \qquad A'_{31} = \begin{pmatrix} a_{21} & a_{31} \\ a_{22} & a_{32} \end{pmatrix} = (A_{13})^{\mathrm{T}},$$

κατά συνέπεια από την (8.7) υπολογίζουμε

$$\begin{aligned} \det A^{\mathrm{T}} &= a'_{11}\det A'_{11} - a'_{21}\det A'_{21} + a'_{31}\det A'_{31} \\ &= a_{11}\det(A_{11})^{\mathrm{T}} - a_{12}\det(A_{12})^{\mathrm{T}} + a_{13}\det(A_{13})^{\mathrm{T}} \\ &= a_{11}\det A_{11} - a_{12}\det A_{12} + a_{13}\det A_{13} && \text{(από την ιδιότητα (O4))} \\ &= \det A && \text{(από την (8.6))} \end{aligned}$$

που είναι το ζητούμενο. Το γενικό βήμα αποδεικνύεται με τελείως όμοιο τρόπο. □

Λήμμα 8.2. *Εάν E και B είναι $n\times n$ μητρώα και το E είναι είναι ένα στοιχειώδες μητρώο, τότε* $\det(EB) = (\det E)(\det B)$.

Απόδειξη. Θεωρούμε τις σχετικές περιπτώσεις για το E.

1. Το E προκύπτει πολλαπλασιάζοντας μια γραμμή του I με μη μηδενική σταθερά, έστω λ, τότε από γραμμικότητα είναι $\det E = \lambda\det I = \lambda$, και το EB προκύπτει πολλαπλασιάζοντας την αντίστοιχη γραμμή του B με λ, έτσι από γραμμικότητα έχουμε $\det(EB) = \lambda\det B = (\det E)(\det B)$.
2. Το E προκύπτει με εναλλαγή δύο γραμμών του I, τότε $\det E = -1$ και το EB προκύπτει από το B με εναλλαγή των αντίστοιχων δύο γραμμών κατά συνέπεια $\det(EB) = -\det B = (\det E)(\det B)$.
3. Το E προκύπτει προσθέτοντας σε μια γραμμή του I το πολλαπλάσιο μιας άλλης γραμμής. Το ίδιο αποτέλεσμα επιφέρει και ο πολλαπλασιασμός EB στο B, και από το Θεώρημα 8.4 (vi) υπολογίζουμε $\det(EB) = \det B = (\det E)(\det B)$.

Η απόδειξη είναι πλήρης. □

Θεώρημα 8.7. *Εάν A και B είναι $n \times n$ μητρώα, τότε* $\det(AB) = (\det A)(\det B)$.

Απόδειξη. Έστω $\operatorname{rank} A < n$, τότε αφενός $\det A = 0$ και αφετέρου, αφού $\operatorname{rank}(AB) \leq \operatorname{rank} A$, έπεται ότι $\det(AB) = 0 = (\det A)(\det B)$. Αν τώρα $\operatorname{rank} A = n$, τότε το A είναι γραμμοϊσοδύναμο με το ταυτοτικό μητρώο, κατά συνέπεια υπάρχουν στοιχειώδη μητρώα $E_1, E_2, \ldots, E_k$, ώστε $A = E_1 E_2 \cdots E_k I$, επομένως από το Λήμμα 8.2 βρίσκουμε

$$\begin{aligned}
\det(AB) &= \det(E_1 E_2 \cdots E_k B) \\
&= (\det E_1) \det(E_2 \cdots E_k B) \\
&= (\det E_1)(\det E_2) \cdots (\det E_k)(\det B) \\
&= (\det(E_1 E_2 \cdots E_k))(\det B) \\
&= (\det A)(\det B)
\end{aligned}$$

□

▼ **Πόρισμα 8.2.** *Εάν A είναι ένα $n \times n$ αντιστρέψιμο μητρώο τότε* $\det A \neq 0$, *επιπλέον*

$$\det(A^{-1}) = \frac{1}{\det A}.$$

Απόδειξη. Από τη σχέση $AA^{-1} = I$ προκύπτει

$$\det(AA^{-1}) = \det I \Rightarrow (\det A)(\det(A^{-1})) = 1,$$

απ΄ όπου προκύπτει το ζητούμενο. □

Παράδειγμα 8.5. Υπολογίζουμε την 2×2 ορίζουσα με χρήση των ιδιοτήτων που την χαρακτηρίζουν μοναδικά (Θεώρημα 8.4) και τις ιδιότητες που απορρέουν από αυτές. Αν

$$\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \end{pmatrix} \qquad \text{και} \qquad \boldsymbol{e}_2 = \begin{pmatrix} 0 \\ 1 \end{pmatrix},$$

τότε

$$\begin{aligned}
\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} &= \det\big(a\boldsymbol{e}_1 + c\boldsymbol{e}_2 \;\; b\boldsymbol{e}_1 + d\boldsymbol{e}_2\big) \\
&= a \det\big(\boldsymbol{e}_1 \;\; b\boldsymbol{e}_1 + d\boldsymbol{e}_2\big) + c \det\big(\boldsymbol{e}_2 \;\; b\boldsymbol{e}_1 + d\boldsymbol{e}_2\big) \\
&= ab \det\big(\boldsymbol{e}_1 \;\; \boldsymbol{e}_1\big) + ad \det\big(\boldsymbol{e}_1 \;\; \boldsymbol{e}_2\big) + cb \det\big(\boldsymbol{e}_2 \;\; \boldsymbol{e}_1\big) + cd \det\big(\boldsymbol{e}_2 \;\; \boldsymbol{e}_2\big) \\
&= ad \det\big(\boldsymbol{e}_1 \;\; \boldsymbol{e}_2\big) - bc \det\big(\boldsymbol{e}_1 \;\; \boldsymbol{e}_2\big) \\
&= ad \det I - bc \det I \\
&= ad - bc.
\end{aligned}$$

Παρατήρηση 8.3 (Ορίζουσα και οδηγοί). Με τη διαδικασία απαλοιφής υπολογίζουμε

$$\begin{pmatrix} 1 & 3 & -2 \\ 2 & 1 & 1 \\ -1 & 2 & 1 \end{pmatrix} \xrightarrow{r_2 \to -2r_1 + r_2} \begin{pmatrix} 1 & 3 & -2 \\ 0 & -5 & 5 \\ -1 & 2 & 1 \end{pmatrix} \xrightarrow{r_3 \to r_2 + r_3} \begin{pmatrix} 1 & 3 & -2 \\ 0 & -5 & 5 \\ 0 & 5 & -1 \end{pmatrix} \xrightarrow{r_3 \to r_2 + r_3} \begin{pmatrix} 1 & 3 & -2 \\ 0 & -5 & 5 \\ 0 & 0 & 4 \end{pmatrix}.$$

Σε κάθε βήμα της απαλοιφής το νέο μητρώο προκύπτει από το παλαιό προσθέτοντας σε μια γραμμή το πολλαπλάσιο μιας άλλης γραμμής, κατά συνέπεια, από τις ιδιότητες της ορίζουσας, η ορίζουσα του νέου μητρώου είναι ίση με αυτή του παλαιού. Έτσι αν A είναι το αρχικό μητρώο και U το τελικό θα έχουμε ότι

$$\det A = \det U = 1(-5)4 = -20.$$

Γενικεύοντας ισχύει το αποτέλεσμα: *Εάν από το μητρώο A με τη διαδικασία της απαλοιφής, κατά την οποία εκτελούνται μόνο οι δύο από τις τρεις στοιχειώδεις πράξεις οι (i) πρόσθεση σε μια γραμμή του πολλαπλασίου άλλης γραμμής και (ii) μετάθεση γραμμών, καταλήγουμε στο άνω τριγωνικό μητρώο U, τότε*

$$\det A = (-1)^p \det U, \tag{8.11}$$

όπου p είναι το πλήθος των μεταθέσεων γραμμών κατά την απαλοιφή. Έτσι αν $d_1, d_2, \ldots, d_n$ είναι τα στοιχεία της διαγωνίου του U, τότε

$$\det A = \pm(d_1 d_2 \cdots d_n). \tag{8.12}$$

Ειδικά αν $d_k \neq 0$ για κάθε k, τότε η ορίζουσα του μητρώου ισούται με το γινόμενο των οδηγών, ενώ αν κάποιο d_k είναι ίσο με μηδέν, ισοδύναμα το πλήθος των οδηγών είναι μικρότερο του n ($\operatorname{rank} A < n$), τότε η ορίζουσα είναι ίση με μηδέν. Βλέπε Παρατήρηση 3.5.

Ο κανόνας του Cramer

Έστω $A\boldsymbol{x} = \boldsymbol{b}$ ένα τετραγωνικό σύστημα, n εξισώσεων με n αγνώστους, και ας υποθέσουμε ότι το μητρώο A είναι αντιστρέψιμο, τότε η μοναδική λύση του συστήματος δίνεται από

$$\boldsymbol{x} = A^{-1}\boldsymbol{b}.$$

Χρησιμοποιώντας ορίζουσες ο τύπος αυτός για τη λύση γίνεται περισσότερο συγκεκριμένος.

Θεώρημα 8.8 (Ο κανόνας του Cramer). *Έστω A ένα $n \times n$ αντιστρέψιμο μητρώο και έστω $\boldsymbol{x} = \begin{pmatrix} x_1 & x_2 & \cdots & x_n \end{pmatrix}^{\mathrm{T}}$ η λύση του συστήματος $A\boldsymbol{x} = \boldsymbol{b}$. Για $k = 1, 2, \ldots, n$ συμβολίζουμε με A_k το μητρώο που προκύπτει από το A αν αντικαταστήσουμε την k-στήλη του με το διάνυσμα $\boldsymbol{b}$, τότε η λύση του συστήματος δίνεται από τη σχέση*

$$x_k = \frac{\det A_k}{\det A}, \qquad k = 1, 2, \ldots, n. \tag{8.13}$$

Απόδειξη. Αν με $\boldsymbol{c}_k$, $k = 1, 2, \ldots, n$ συμβολίσουμε τις στήλες του A, ώστε $A = \begin{pmatrix} \boldsymbol{c}_1 & \boldsymbol{c}_2 & \cdots & \boldsymbol{c}_n \end{pmatrix}$, και με X_k συμβολίσουμε το μητρώο που προκύπτει από το ταυτοτικό μητρώο I αν αντικαταστήσουμε την k-στήλη του με $\boldsymbol{x}$, δηλαδή

$$X_k = \begin{pmatrix} \boldsymbol{e}_1 & \cdots & \boldsymbol{e}_{k-1} & \boldsymbol{x} & \boldsymbol{e}_{k+1} & \cdots & \boldsymbol{e}_n \end{pmatrix}, \tag{8.14}$$

όπου $\boldsymbol{e}_k$, $k = 1, \ldots, n$ είναι οι στήλες του I, τότε

$$AX_k = \begin{pmatrix} A\boldsymbol{e}_1 & \cdots & A\boldsymbol{e}_{k-1} & A\boldsymbol{x} & A\boldsymbol{e}_{k+1} & \cdots & A\boldsymbol{e}_n \end{pmatrix}$$

ή, αφού $AI = A$,

$$AX_k = \begin{pmatrix} \boldsymbol{c}_1 & \cdots & \boldsymbol{c}_{k-1} & \boldsymbol{b} & \boldsymbol{c}_{k+1} & \cdots & \boldsymbol{c}_n \end{pmatrix} = A_k.$$

Έτσι υπολογίζουμε

$$\det A_k = \det(AX_k) = (\det A)(\det X_k) \tag{8.15}$$

Γράφοντας το μητρώο X_k ως

$$X_k = \begin{pmatrix} \boldsymbol{e}_1 & \cdots & \boldsymbol{e}_{k-1} & \underbrace{x_1\boldsymbol{e}_1 + \cdots + x_k\boldsymbol{e}_k + \cdots + x_n\boldsymbol{e}_n}_{k\text{-στήλη}} & \boldsymbol{e}_{k+1} & \cdots & \boldsymbol{e}_n \end{pmatrix}, \tag{8.16}$$

από τη γραμμικότητα της ορίζουσας και τη γραμμική εξάρτηση των στηλών του X_k υπολογίζουμε

$$\det X_k = x_k \det I = x_k. \tag{8.17}$$

Από τις (8.15), (8.17) έπεται η ζητούμενη σχέση. □

8.4 Ένας τύπος για το αντίστροφο μητρώο

Στην περίπτωση ενός 2×2 μητρώου

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

το αντίστροφο, εφόσον αυτό υπάρχει είδαμε ότι είναι το

$$A^{-1} = \frac{1}{\det A}\begin{pmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{pmatrix}.$$

Θέλουμε να γενικεύσουμε αυτό τον τύπο για ένα $n \times n$ μητρώο. Ας δούμε τι κρύβει ο τύπος της 2×2 περίπτωσης. Αν με $\left(A^{-1}\right)_{ij}$ συμβολίσουμε το στοιχείο του A^{-1} στην i-γραμμή και j-στήλη, και με $\psi_{ij} = (-1)^{i+j} \det A_{ij}$ τον συμπαράγοντα του a_{ij}, βλέπε Ορισμό 8.4, παρατηρούμε, επιστρατεύοντας και την φαντασία μας είναι αλήθεια, ότι

$$\left(A^{-1}\right)_{11} = \frac{a_{22}}{\det A} = (-1)^{1+1}\frac{\det A_{11}}{\det A} = \frac{\psi_{11}}{\det A}, \qquad \left(A^{-1}\right)_{12} = \frac{-a_{12}}{\det A} = (-1)^{2+1}\frac{\det A_{21}}{\det A} = \frac{\psi_{21}}{\det A},$$

$$\left(A^{-1}\right)_{21} = \frac{-a_{12}}{\det A} = (-1)^{1+2}\frac{\det A_{12}}{\det A} = \frac{\psi_{12}}{\det A}, \qquad \left(A^{-1}\right)_{22} = \frac{a_{11}}{\det A} = (-1)^{2+2}\frac{\det A_{22}}{\det A} = \frac{\psi_{22}}{\det A}.$$

Έτσι αν Ψ είναι το μητρώο των συμπαραγόντων, τότε

$$A^{-1} = \frac{1}{\det A}\Psi^{\mathrm{T}}.$$

Αποδεικνύεται ότι ο τύπος αυτός ισχύει και στη γενική $n \times n$ περίπτωση.

Θεώρημα 8.9. *Εάν A είναι ένα $n \times n$ αντιστρέψιμο μητρώο και $\Psi = \left(\psi_{ij}\right)$ είναι το μητρώο των συμπαραγόντων, δηλαδή $\psi_{ij} = (-1)^{i+j} \det A_{ij}$, τότε*

$$A^{-1} = \frac{1}{\det A}\Psi^{\mathrm{T}}. \tag{8.18}$$

Απόδειξη. Αποδεικνύουμε την ισοδύναμη σχέση $(\det A)I = A\Psi^{\mathrm{T}}$, ή

$$\begin{pmatrix} \det A & 0 & \cdots & 0 \\ 0 & \det A & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \det A \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}\begin{pmatrix} \psi_{11} & \psi_{21} & \cdots & \psi_{n1} \\ \psi_{12} & \psi_{22} & \cdots & \psi_{n2} \\ \vdots & \vdots & \ddots & \vdots \\ \psi_{1n} & \psi_{2n} & \cdots & \psi_{nn} \end{pmatrix}. \tag{8.19}$$

Για τα στοιχεία της διαγωνίου θα πρέπει

$$\det A = a_{i1}\psi_{i1} + a_{i2}\psi_{i2} + \cdots + a_{in}\psi_{in}$$

το οποίο ισχύει αφού είναι το ανάπτυγμα της ορίζουσας ως προς την i-γραμμή. Για τα στοιχεία εκτός διαγωνίου θα πρέπει

$$\begin{aligned} 0 &= a_{i1}\psi_{j1} + a_{i2}\psi_{j2} + \cdots + a_{in}\psi_{jn}, && i \neq j \\ &= (-1)^{j+1}a_{i1}\det A_{j1} + (-1)^{j+2}a_{i2}\det A_{j2} + \cdots + (-1)^{j+n}a_{in}\det A_{jn}, && i \neq j. \end{aligned}$$

Ας δούμε μια περίπτωση για συγκεκριμένα i και j. Για $i = 1$ και $j = 2$ το μέλος στα δεξιά της τελευταίας σχέσης είναι

$$(-1)^{2+1}a_{11}\det A_{21} + (-1)^{2+2}a_{12}\det A_{22} + \cdots + (-1)^{2+n}a_{1n}\det A_{2n} = \begin{vmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{11} & a_{12} & \cdots & a_{1n} \\ a_{31} & a_{32} & \cdots & a_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{vmatrix} = 0$$

αφού η πρώτη και η δεύτερη γραμμή είναι ίσες. Όμοια η έκφραση

$$(-1)^{j+1}a_{i1}\det A_{j1} + (-1)^{j+2}a_{i2}\det A_{j2} + \cdots + (-1)^{j+n}a_{in}\det A_{jn}, \qquad i \neq j$$

είναι η ορίζουσα του μητρώου που προκύπτει από το A αντικαθιστώντας την j-γραμμή με την i-γραμμή. Επειδή $i \neq j$ το μητρώο έχει δύο ίσες γραμμές, κατά συνέπεια η ορίζουσά του είναι ίση με μηδέν. Συνεπώς η (8.19) ισχύει. □

8.5 Ο τύπος του Leibniz

Ένας διαφορετικός τρόπος υπολογισμού της ορίζουζας επιτυγχάνεται με εργαλεία της Συνδυαστικής Ανάλυσης. Παραθέτουμε χωρίς απόδειξη το βασικό αποτέλεσμα. Για την απόδειξη αλλά και την προσέγγιση αυτή παραπέμπουμε στα συγγράμματα [6] ή [7].

Αναπτύσσοντας μια ορίζουσα, για παράδειγμα, τάξης 3 βρίσκουμε

$$\begin{aligned} \begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= a_{11}\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33} \end{vmatrix} + a_{13}\begin{vmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{vmatrix} \\ &= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31} \end{aligned}$$

και παρατηρώντας ότι κάθε όρος του αθροίσματος είναι της μορφής $a_{1i}a_{2j}a_{3k}$ με $i, j, k \in \{1, 2, 3\}$ όπου $i \neq j$, $i \neq k$, και $j \neq k$ σε όλους τους δυνατούς συνδυασμούς, μπορούμε να γράψουμε

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \sum_{k=1}^{3!}(-1)^{n_k}a_{1\sigma_k(1)}a_{2\sigma_k(2)}a_{3\sigma_k(3)}$$

όπου $n_k \in \{1, 2\}$ και σ_k είναι μετάθεση του $\{1, 2, 3\}$, $k = 1, 2, \ldots, 6$. Θυμίζουμε ότι μια μετάθεση του συνόλου $\{1, 2, \ldots, n\}$ είναι μια ένα-προς-ένα αντιστοιχία $\sigma : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}$. Στο Παράδειγμα 2.14 δείξαμε ότι οι μεταθέσεις του $\{1, 2, \ldots, n\}$ με την πράξη του πολλαπλασιασμού-σύνθεσης αποτελούν ομάδα, την συμμετρική ομάδα S_n με πλήθος στοιχείων $n!$. Θα λέμε ότι οι τιμές της ταυτοτικής μετάθεσης έχουν τη φυσική διάταξη. Θα λέμε ότι μια μια μετάθεση είναι **άρτια** αν οι τιμές της επανέρχονται στη φυσική διάταξη με άρτιο πλήθος εναλλαγών. Θα λέμε ότι μια μια μετάθεση είναι **περιττή** αν οι τιμές της επανέρχονται στη φυσική διάταξη με περιττό πλήθος εναλλαγών. Για παράδειγμα η 312 είναι άρτια μετάθεση ενώ η 132 είναι περιττή μετάθεση. Ορίζουμε το **πρόσημο** μιας μετάθεσης σ, $\pi(s)$, με τη σχέση

$$\pi(\sigma) = \begin{cases} +1, & \text{αν η } \sigma \text{ είναι άρτια} \\ -1, & \text{αν η } \sigma \text{ είναι περιττή.} \end{cases}$$

Θεώρημα 8.10. *Εάν A είναι ένα $n \times n$ μητρώο και S_n είναι η ομάδα των μεταθέσεων του $\{1, 2, \dots, n\}$, τότε η ορίζουσα του A είναι*

$$\det A = \sum_{\sigma \in S_n} \pi(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)},$$

όπου $\pi(\sigma)$ είναι το πρόσημο της μετάθεσης σ.

8.6 Το εξωτερικό γινόμενο

Αν $\{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ είναι η κανονική βάση στον $\mathbb{R}^3$ μπορούμε να γράψουμε, όπως συνηθίζεται στον Απειροστικό Λογισμό, ή στη Φυσική,

$$\boldsymbol{i} = \boldsymbol{e}_1, \qquad \boldsymbol{j} = \boldsymbol{e}_2, \qquad \boldsymbol{k} = \boldsymbol{e}_3.$$

Ορισμός 8.5. Αν

$$\boldsymbol{u} = \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}, \qquad \boldsymbol{v} = \begin{pmatrix} v_1 \\ v_2 \\ v_3 \end{pmatrix},$$

είναι διανύσματα στον $\mathbb{R}^3$, ορίζουμε το **εξωτερικό γινόμενο** των $\boldsymbol{u}$ και $\boldsymbol{v}$, και το συμβολίζουμε με $\boldsymbol{u} \times \boldsymbol{v}$, το διάνυσμα

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} \boldsymbol{i} - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} \boldsymbol{j} + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} \boldsymbol{k}. \tag{8.20}$$

Από το ανάπτυγμα της 3×3 ορίζουσας βλέπουμε αμέσως ότι μπορούμε να παραστήσουμε το εξωτερικό γινόμενο ως ορίζουσα στην ευκολομνημόνευτη μορφή

$$\boldsymbol{u} \times \boldsymbol{v} = \begin{vmatrix} \boldsymbol{i} & \boldsymbol{j} & \boldsymbol{k} \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix}.$$

Το εξωτερικό γινόμενο το λέμε και **σταυρωτό γινόμενο** (cross product), ή **διανυσματικό γινόμενο** (vector product).

Παρατήρηση 8.4. Από την (8.20) και τις ιδιότητες της ορίζουσας έπεται αμέσως ότι

(1) Αν τα $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι γραμμικά εξαρτημένα, τότε $\boldsymbol{u} \times \boldsymbol{v} = \boldsymbol{0}$.

(2) $\boldsymbol{u} \times \boldsymbol{v} = -\boldsymbol{v} \times \boldsymbol{u}$.

(3) $\boldsymbol{u} \times \mathbf{0} = \mathbf{0}$.

Επίσης άμμεση συνέπεια του ορισμού των διανυσμάτων $\boldsymbol{i}$, $\boldsymbol{j}$ και $\boldsymbol{k}$ και του εξωτερικού γινομένου είναι το αποτέλεσμα

▲ **Πρόταση 8.2.** *Εάν, όπως ορίσαμε,*

$$\boldsymbol{i} = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{j} = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{k} = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix},$$

τότε

$$\boldsymbol{i} \times \boldsymbol{j} = \boldsymbol{k}, \qquad \boldsymbol{j} \times \boldsymbol{k} = \boldsymbol{i}, \qquad \boldsymbol{k} \times \boldsymbol{i} = \boldsymbol{j}.$$

Απόδειξη. Αφήνεται σαν άσκηση. □

Ένα γενικότερο αποτέλεσμα παρέχεται από την

▲ **Πρόταση 8.3.** *Εάν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα στον $\mathbb{R}^3$, τότε*

$$\boldsymbol{u} \times \boldsymbol{v} \perp \boldsymbol{u} \qquad \textit{και} \qquad \boldsymbol{u} \times \boldsymbol{v} \perp \boldsymbol{v},$$

δηλαδή το $\boldsymbol{u} \times \boldsymbol{v}$ είναι ορθογώνιο στο $\boldsymbol{u}$ και στο $\boldsymbol{v}$, κατά συνέπεια είναι ορθογώνιο στο επίπεδο που παράγεται από τα $\boldsymbol{u}$ και $\boldsymbol{v}$ οποτεδήποτε τα $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι γραμμικά ανεξάρτητα.

Απόδειξη. Δείχνουμε ότι $(\boldsymbol{u}\times\boldsymbol{v})\cdot\boldsymbol{u} = 0$. Πράγματι από την (8.20) και τον ορισμό του εσωτερικού γινομένου υπολογίζουμε

$$\begin{aligned}(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{u} &= \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix} u_1 - \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix} u_2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix} u_3 \\ &= \begin{vmatrix} u_1 & u_2 & u_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix},\end{aligned}$$

από τον ορισμό της ορίζουσας, συνεπώς $(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{u} = 0$ αφού η ορίζουσα έχει δυο γραμμές ίσες. Η απόδειξη ότι $(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{v} = 0$ είναι ανάλογη, ενώ από τη γραμμικότητα του εσωτερικού γινομένου έπεται ότι

$$(\boldsymbol{u} \times \boldsymbol{v}) \cdot (\lambda \boldsymbol{u} + \mu \boldsymbol{v}) = \lambda(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{u} + \mu(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{v} = 0$$

για κάθε $\lambda \in \mathbb{R}$ και για κάθε $\mu \in \mathbb{R}$. □

Εάν $\boldsymbol{w} \in \mathbb{R}^3$ παρατηρούμε ότι

$$(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w} = \begin{vmatrix} w_1 & w_2 & w_3 \\ u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = -\begin{vmatrix} u_1 & u_2 & u_3 \\ w_1 & w_2 & w_3 \\ v_1 & v_2 & v_3 \end{vmatrix} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix}$$

Ορισμός 8.6. Εάν $\boldsymbol{u}$, $\boldsymbol{v}$ και $\boldsymbol{w}$ είναι διανύσματα στον $\mathbb{R}^3$ το $(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w}$ λέμε **τριπλό**

γινόμενο (triple product) των $\boldsymbol{u}$, $\boldsymbol{v}$ και $\boldsymbol{w}$, έτσι

$$(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w} = \begin{vmatrix} u_1 & u_2 & u_3 \\ v_1 & v_2 & v_3 \\ w_1 & w_2 & w_3 \end{vmatrix} = \begin{vmatrix} u_1 & v_1 & w_1 \\ u_2 & v_2 & w_2 \\ u_3 & v_3 & w_3 \end{vmatrix}.$$

Από τον ορισμό του τριπλού γινομένου απορρέει αμέσως ότι τα διανύσματα $\boldsymbol{u}$, $\boldsymbol{v}$ και $\boldsymbol{w}$ είναι γραμμικά εξαρτημένα, ισοδύναμα είναι συνεπίπεδα, περιέχονται δηλαδή στο ίδιο επίπεδο, αν και μόνο αν

$$(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w} = 0.$$

► **Άσκηση 8.2.** Το εξωτερικό γινόμενο προσδίδει στον $\mathbb{R}^3$ επιπλέον δομή.

(α΄) Εάν $\boldsymbol{a}$, $\boldsymbol{b}$, και $\boldsymbol{c}$ είναι διανύσματα στον $\mathbb{R}^3$ δείξτε ότι

$$\boldsymbol{a} \times (\boldsymbol{b} + \boldsymbol{c}) = \boldsymbol{a} \times \boldsymbol{b} + \boldsymbol{a} \times \boldsymbol{c}, \qquad \text{και} \qquad (\boldsymbol{a} + \boldsymbol{b}) \times \boldsymbol{c} = \boldsymbol{a} \times \boldsymbol{c} + \boldsymbol{b} \times \boldsymbol{c}$$

ισχύει δηλαδή ο επιμεριστικός νόμος.

(β΄) Δείξτε ότι ο χώρος $\mathbb{R}^3$ εφοδιασμένος με το εξωτερικό γινόμενο, $(\mathbb{R}^3, +, \cdot, \times)$, είναι μια μη προσεταιριστική άλγεβρα[1], δηλαδή δεν ισχύει πάντα η ισότητα

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) = (\boldsymbol{a} \times \boldsymbol{b}) \times \boldsymbol{c}.$$

(γ΄) Δείξτε ότι ικανοποιείται η **ταυτότητα του Jacobi**

$$\boldsymbol{a} \times (\boldsymbol{b} \times \boldsymbol{c}) + \boldsymbol{b} \times (\boldsymbol{c} \times \boldsymbol{a}) + \boldsymbol{c} \times (\boldsymbol{a} \times \boldsymbol{b}) = \boldsymbol{0}$$

για όλα τα διανύσματα $\boldsymbol{a}$, $\boldsymbol{b}$, και $\boldsymbol{c}$.

8.7 Όγκος και ορίζουσα

Αν τα τρία διανύσματα δεν είναι συνεπίπεδα ορίζουν ένα παραλληλεπίπεδο στο $\mathbb{R}^3$. Το βαθμωτό μέγεθος $(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w}$ σχετίζεται με τον όγκο του παραλληλεπιπέδου. Ας δούμε κάποιες απλές περιπτώσεις.

- Αν $\boldsymbol{u} = \boldsymbol{e}_1$, $\boldsymbol{v} = \boldsymbol{e}_2$, $\boldsymbol{w} = \boldsymbol{e}_3$, τότε
$$(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w} = |I| = 1$$
όπου I είναι το ταυτοτικό μητρώο, και το $(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w}$ δίνει ακριβώς τον όγκο του μοναδιαίου κύβου.

- Αν λ, μ, ν είναι θετικές σταθερές και $\boldsymbol{u} = \lambda \boldsymbol{e}_1$, $\boldsymbol{v} = \mu \boldsymbol{e}_2$, $\boldsymbol{w} = \nu \boldsymbol{e}_3$, τότε
$$(\boldsymbol{u} \times \boldsymbol{v}) \cdot \boldsymbol{w} = \begin{vmatrix} \lambda & 0 & 0 \\ 0 & \mu & 0 \\ 0 & 0 & \nu \end{vmatrix} = \lambda\mu\nu$$
που και πάλι είναι ο όγκος του ορθογωνίου παραλληλεπιπέδου που ορίζουν τα τρία διανύσματα.

[1]Στην παράγραφο 4.7 ορίσαμε την δομή Άλγεβρα και αναφέραμε ότι ο πολλαπλασιασμός μεταξύ διανυσμάτων είναι προσεταιριστικός, νόμος (A1). Μπορεί ωστόσο ο προσεταιριστικός νόμος να παραληφθεί από τον ορισμό της Άλγεβρας και πολλοί συγγραφείς το ακολουθούν. Σε μια τέτοια περίπτωση όταν ισχύει η (A1) η Άλγεβρα λέγεται προσεταιριστική, ενώ αν δεν ισχύει η (A1) η Άλγεβρα λέγεται μη προσεταιριστική.

▲ **Πρόταση 8.4.** *Εάν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα στον $\mathbb{R}^3$, τότε*

$$\|\boldsymbol{u}\times\boldsymbol{v}\|^2 = \|\boldsymbol{u}\|^2\|\boldsymbol{v}\|^2 - \langle\boldsymbol{u},\boldsymbol{v}\rangle^2, \tag{8.21}$$

όπου $\langle\cdot,\cdot\rangle$ είναι το σύνηθες εσωτερικό γινόμενο.

Απόδειξη. Από την (8.20) υπολογίζουμε

$$\begin{aligned}\|\boldsymbol{u}\times\boldsymbol{v}\|^2 &= \begin{vmatrix} u_2 & u_3 \\ v_2 & v_3 \end{vmatrix}^2 + \begin{vmatrix} u_1 & u_3 \\ v_1 & v_3 \end{vmatrix}^2 + \begin{vmatrix} u_1 & u_2 \\ v_1 & v_2 \end{vmatrix}^2 \\ &= (u_2v_3 - u_3v_3)^2 + (u_1v_3 - u_3v_1)^2 + (u_1u_2 - u_2v_1)^2 \\ &= (u_1^2 + u_2^2 + u_3^2)(v_1^2 + v_2^2 + v_3^2) - (u_1v_1 + u_2v_2 + u_3v_3)^2\end{aligned}$$

που είναι το ζητούμενο. □

▼ **Πόρισμα 8.3.** *Εάν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα στον $\mathbb{R}^3$, τότε*

$$\|\boldsymbol{u}\times\boldsymbol{v}\| = \|\boldsymbol{u}\|\|\boldsymbol{v}\|\sin\theta,$$

όπου θ είναι η γωνία μεταξύ των $\boldsymbol{u}$ και $\boldsymbol{v}$, $0 \le \theta \le \pi$.

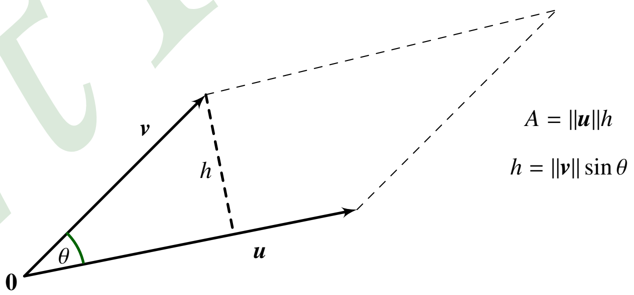
Απόδειξη. Από την (8.21) και τον ορισμό του εσωτερικού γινομένου υπολογίζουμε

$$\begin{aligned}\|\boldsymbol{u}\times\boldsymbol{v}\|^2 &= \|\boldsymbol{u}\|^2\|\boldsymbol{v}\|^2 - \|\boldsymbol{u}\|^2\|\boldsymbol{v}\|^2\cos^2\theta \\ &= \|\boldsymbol{u}\|^2\|\boldsymbol{v}\|^2\sin^2\theta,\end{aligned}$$

απ' όπου έπεται το ζητούμενο παίρνοντας τετραγωνικές ρίζες αφού $\sin\theta \ge 0$. □

Παρατήρηση 8.5 (Γεωμετρική σημασία του Πορίσματος 8.3). Εάν $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι διανύσματα στον $\mathbb{R}^3$ και με $A(\boldsymbol{u},\boldsymbol{v})$ συμβολίσουμε το εμβαδόν του παραλληλογράμμου με πλευρές τα $\boldsymbol{u}$ και $\boldsymbol{v}$, τότε, όπως βλέπουμε και στο Σχήμα 8.1,

$$A(\boldsymbol{u},\boldsymbol{v}) = \|\boldsymbol{u}\times\boldsymbol{v}\|. \tag{8.22}$$



Σχήμα 8.1: Εμβαδόν ορθογωνίου

Ορισμός 8.7. Θα λέμε ότι η τριάδα $\boldsymbol{u}$-$\boldsymbol{v}$-$\boldsymbol{w}$ διανυσμάτων του $\mathbb{R}^3$ αποτελεί ένα **θετικά προσανατολισμένο σύστημα** αν

$$\det\begin{pmatrix}\boldsymbol{u} & \boldsymbol{v} & \boldsymbol{w}\end{pmatrix} > 0.$$

Στην περίπτωση αυτή το σύστημα $\boldsymbol{u}$-$\boldsymbol{v}$-$\boldsymbol{w}$ έχει τον προσανατολισμό του τρισορθογωνίου συστήματος $\boldsymbol{e}_1$-$\boldsymbol{e}_2$-$\boldsymbol{e}_3$.

Παρατήρηση 8.6. Αν τα διανύσματα $\boldsymbol{u}$ και $\boldsymbol{v}$ είναι γραμμικά ανεξάρτητα, τότε το σύστημα $\boldsymbol{u}$-$\boldsymbol{v}$-$\boldsymbol{u}\times\boldsymbol{v}$ είναι θετικά προσανατολισμένο.

Πράγματι γράφοντας $\boldsymbol{u}\times\boldsymbol{v} = \begin{pmatrix} p_1 & p_2 & p_3 \end{pmatrix}^{\mathrm{T}}$ έχουμε

$$\begin{aligned}\det\begin{pmatrix}\boldsymbol{u} & \boldsymbol{v} & \boldsymbol{u}\times\boldsymbol{v}\end{pmatrix} &= \begin{vmatrix} u_1 & v_1 & p_1 \\ u_2 & v_2 & p_2 \\ u_3 & v_3 & p_3 \end{vmatrix} \\ &= p_1\begin{vmatrix} u_2 & v_2 \\ u_3 & v_3 \end{vmatrix} - p_2\begin{vmatrix} u_1 & v_1 \\ u_3 & v_3 \end{vmatrix} + p_3\begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix} \\ &= \begin{vmatrix} u_2 & v_2 \\ u_3 & v_3 \end{vmatrix}^2 + \begin{vmatrix} u_1 & v_1 \\ u_3 & v_3 \end{vmatrix}^2 + \begin{vmatrix} u_1 & v_1 \\ u_2 & v_2 \end{vmatrix}^2\end{aligned}$$

από την (8.20).

▲ **Πρόταση 8.5.** *Αν $\boldsymbol{u}$, $\boldsymbol{v}$, $\boldsymbol{w}$ είναι διανύσματα στον $\mathbb{R}^3$, τότε ο όγκος $V(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w})$ του στερεού που παράγεται από τα τρία αυτά διανύσματα είναι*

$$V(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) = |\det\begin{pmatrix}\boldsymbol{u} & \boldsymbol{v} & \boldsymbol{w}\end{pmatrix}|,$$

όπου $|\cdot|$ είναι η απόλυτη τιμή.

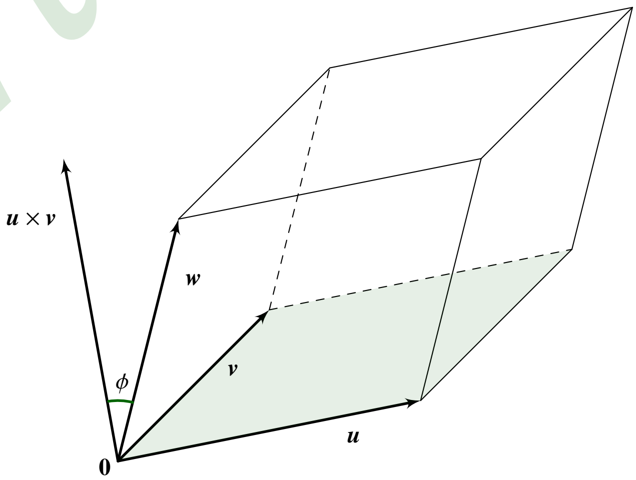
Απόδειξη. Έστω ότι το σύστημα $\boldsymbol{u}$-$\boldsymbol{v}$-$\boldsymbol{w}$ είναι θετικά προσανατολισμένο. Αν ϕ είναι η γωνία μεταξύ των $\boldsymbol{w}$ και $\boldsymbol{u}\times\boldsymbol{v}$, βλέπε Σχήμα 8.2, τότε $0 \le \phi < \pi/2$ (γιατί;). Έτσι υπολογίζουμε

$$\begin{aligned}V(\boldsymbol{u},\boldsymbol{v},\boldsymbol{w}) &= A(\boldsymbol{u},\boldsymbol{v})\|\boldsymbol{w}\|\cos\phi \\ &= \|\boldsymbol{u}\times\boldsymbol{v}\|\|\boldsymbol{w}\|\cos\phi \\ &= (\boldsymbol{u}\times\boldsymbol{v})\cdot\boldsymbol{w} \\ &= \det\begin{pmatrix}\boldsymbol{u} & \boldsymbol{v} & \boldsymbol{w}\end{pmatrix}.\end{aligned}$$

Αν το σύστημα $\boldsymbol{u}$-$\boldsymbol{v}$-$\boldsymbol{w}$ δεν είναι θετικά προσανατολισμένο, τότε το $\boldsymbol{v}$-$\boldsymbol{u}$-$\boldsymbol{w}$ είναι θετικά προσανατολισμένο αφού

$$\det\begin{pmatrix}\boldsymbol{v} & \boldsymbol{u} & \boldsymbol{w}\end{pmatrix} = -\det\begin{pmatrix}\boldsymbol{u} & \boldsymbol{v} & \boldsymbol{w}\end{pmatrix}.$$

□



Σχήμα 8.2: Το παραλληλεπίπεδο που παράγεται από τα διανύσματα $\boldsymbol{u}$, $\boldsymbol{v}$, και $\boldsymbol{w}$

Παράδειγμα 8.6. Αν $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{u}_3$ είναι γραμμικά ανεξάρτητα διανύσματα στον $\mathbb{R}^3$, και $T : \mathbb{R}^3 \to \mathbb{R}^3$ είναι ένας γραμμικός μετασχηματισμός να υπολογισθεί ο όγκος του παραλληλεπιπέδου που παράγεται από τα διανύσματα $T\boldsymbol{u}_1$, $T\boldsymbol{u}_2$, και $T\boldsymbol{u}_3$.
Από την Πρόταση 8.5 έχουμε

$$V(T\boldsymbol{u}_1, T\boldsymbol{u}_2, T\boldsymbol{u}_3) = \left|\det\begin{pmatrix}T\boldsymbol{u}_1 & T\boldsymbol{u}_2 & T\boldsymbol{u}_3\end{pmatrix}\right| \tag{8.23}$$

Τα διανύσματα $\boldsymbol{u}_1$, $\boldsymbol{u}_2$, $\boldsymbol{u}_3$ αποτελούν μια βάση για τον $\mathbb{R}^3$, κατά συνέπεια

$$\begin{aligned} T\boldsymbol{u}_1 &= a_{11}\boldsymbol{u}_1 + a_{12}\boldsymbol{u}_2 + a_{13}\boldsymbol{u}_3 \\ T\boldsymbol{u}_2 &= a_{21}\boldsymbol{u}_1 + a_{22}\boldsymbol{u}_2 + a_{23}\boldsymbol{u}_3 \\ T\boldsymbol{u}_3 &= a_{31}\boldsymbol{u}_1 + a_{32}\boldsymbol{u}_2 + a_{33}\boldsymbol{u}_3 \end{aligned}$$

για κάποιες σταθερές a_{ij}. Έτσι

$$T\boldsymbol{u}_i = \begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3\end{pmatrix}\begin{pmatrix}a_{i1}\\ a_{i2}\\ a_{i3}\end{pmatrix}, \qquad i = 1, 2, 3$$

επομένως

$$\begin{pmatrix}T\boldsymbol{u}_1 & T\boldsymbol{u}_2 & T\boldsymbol{u}_3\end{pmatrix} = \begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3\end{pmatrix}\begin{pmatrix}a_{11} & a_{21} & a_{31}\\ a_{12} & a_{22} & a_{32}\\ a_{13} & a_{23} & a_{33}\end{pmatrix}.$$

Εάν A_T είναι το μητρώο στα δεξιά της τελευταίας σχέσης, από τις ιδιότητες της ορίζουσας παίρνουμε

$$\det\begin{pmatrix}T\boldsymbol{u}_1 & T\boldsymbol{u}_2 & T\boldsymbol{u}_3\end{pmatrix} = \det\begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3\end{pmatrix}(\det A_T)$$

απ΄ όπου, μέσω της (8.23), έπεται ότι

$$V(T\boldsymbol{u}_1, T\boldsymbol{u}_2, T\boldsymbol{u}_3) = |\det A_T| V(\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3) \tag{8.24}$$

Σημειώνουμε ότι το μητρώο A_T είναι το μητρώο που παριστάνει τον μετασχηματισμό T ως προς την βάση $\mathcal{B} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$, με την ένοια ότι αν

$$\boldsymbol{x} = c_1\boldsymbol{u}_1 + c_2\boldsymbol{u}_2 + c_3\boldsymbol{u}_3 = \begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3\end{pmatrix}\begin{pmatrix}c_1\\ c_2\\ c_3\end{pmatrix},$$

τότε

$$T\boldsymbol{x} = \begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3\end{pmatrix}A_T\begin{pmatrix}c_1\\ c_2\\ c_3\end{pmatrix},$$

για κάθε $\boldsymbol{x} \in \mathbb{R}^3$. Η j-στήλη του A_T, $j = 1, 2, 3$, περιέχει τους συντελεστές του $T\boldsymbol{u}_j$ ως προς τη βάση $\mathcal{B} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$, δηλαδή είναι το διάνυσμα συντεταγμένων $(T\boldsymbol{u}_j)_{\mathcal{B}}$ του $T\boldsymbol{u}_j$ ως προς την βάση $\mathcal{B}$, έτσι

$$A_T = \begin{pmatrix}(T\boldsymbol{u}_1)_{\mathcal{B}} & (T\boldsymbol{u}_2)_{\mathcal{B}} & (T\boldsymbol{u}_3)_{\mathcal{B}}\end{pmatrix},$$

βλέπε Παρατήρηση 4.9.

8.8 Ασκήσεις

1. Να βρεθεί η τιμή του k για την οποία

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ 2b_1 - 3a_1 & 2b_2 - 3a_2 & 2b_3 - 3a_3 \\ 3c_1 - 4b_1 & 3c_2 - 4b_2 & 3c_3 - 4b_3 \end{vmatrix} = k \begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix}.$$

2. Εάν

$$\begin{vmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \\ c_1 & c_2 & c_3 \end{vmatrix} = m,$$

υπολογίστε τις ορίζουσες

(α΄) $\begin{vmatrix} b_1 & -b_2 & b_3 \\ a_1 & -a_2 & a_3 \\ c_1 & -c_2 & c_3 \end{vmatrix}$

(β΄) $\begin{vmatrix} a_1 & a_2 & 3a_3 \\ b_1 & b_2 & 3b_3 \\ 3c_1 & 3c_2 & 9c_3 \end{vmatrix}$

(γ΄) $\begin{vmatrix} 2a_1 & 2a_2 & 2a_3 \\ 2b_1 & 2b_2 & 2b_3 \\ 2c_1 & 2c_2 & 2c_3 \end{vmatrix}$

(δ΄) $\begin{vmatrix} b_1 & b_3 & b_2 \\ a_1 & a_3 & a_2 \\ c_1 & c_3 & c_2 \end{vmatrix}$

(ε΄) $\begin{vmatrix} a_1 & c_1 & b_1 \\ a_2 & c_2 & b_2 \\ a_3 & c_3 & b_3 \end{vmatrix}$

(ϛ΄) $\begin{vmatrix} c_1 - a_1 & c_2 - a_2 & c_3 - a_3 \\ b_1 + a_1 & b_2 + a_2 & b_3 + a_3 \\ b_1 - c_1 & b_2 - c_2 & b_3 - c_3 \end{vmatrix}$.

3. Εάν A είναι ένα 3×3 μητρώο και $\det A = -2$ να υπολογισθούν οι ορίζουσες

(α΄) $\det(-3A)$

(β΄) $\det(2A^{-1})$

(γ΄) $\det((2A)^{-1})$

(δ΄) $\det(2A(3A^{-1}))$

(ε΄) $\det(2A(3A)^{-1})$

(ϛ΄) $\det(2A(3A)^{\mathrm{T}})$.

4. Χρησιμοποιώντας τις ιδιότητες της ορίζουσας μόνο, δίχως να αναπτύξετε, υπολογίσετε τις

(α΄) $\begin{vmatrix} 111 & 211 & 311 & 411 \\ 112 & 212 & 312 & 412 \\ 113 & 213 & 313 & 413 \\ 114 & 214 & 314 & 414 \end{vmatrix}$,

(β΄) $\begin{vmatrix} a & b & c & d \\ a+k & b+k & c+k & d+k \\ a+n & b+n & c+n & d+n \\ ab & bc & cd & da \end{vmatrix}$,

όπου a, b, c, d είναι μη μηδενικοί πραγματικοί αριθμοί και k, n ακέραιοι.

5. Δικαιολογήστε δίχως να αναπτύξετε γιατί η τιμή της ορίζουσας είναι μηδέν.

$$\begin{vmatrix} a_1 & a_2 & a_3 & a_4 \\ b_1 & b_2 & 0 & 0 \\ c_1 & c_2 & 0 & 0 \\ d_1 & d_2 & 0 & 0 \end{vmatrix}.$$

6. Να υπολογιστεί η τιμή της ορίζουσας
$$\begin{vmatrix} 0 & 0 & a_1 & a_2 \\ 0 & 0 & b_1 & b_2 \\ c_1 & c_2 & 0 & 0 \\ d_1 & d_2 & 0 & 0 \end{vmatrix}.$$

7. [9]. Εάν $A = \left(a_{ij}\right)$ είναι $n \times n$, $n > 1$, και $a_{ij} = ij$, δείξτε ότι $\det A = 0$.

8. Εάν $A = \left(a_{ij}\right)$ είναι $n \times n$, $n > 2$, και $a_{ij} = i + j$, δείξτε ότι $\det A = 0$.

9. ΣΩΣΤΟ/ΛΑΘΟΣ Εάν A και B είναι $n \times n$ μητρώα με $\det A = \det B$, τότε $\det(A + B) = 2\det A$. Δικαιολογήστε την απάντησή σας.

10. Υπολογίστε την ορίζουσα του $n \times n$ μητρώου
$$\begin{pmatrix} & & 1 \\ & \iddots & \\ 1 & & \end{pmatrix}.$$
Υπόδειξη: Διακρίνετε τις περιπτώσεις n = άρτιο και n = περιττό.

11. Δίνεται το τριδιαγώνιο μητρώο
$$A_2 = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}, \qquad A_n = \begin{pmatrix} 2 & 1 & 0 & 0 & \cdots & 0 & 0 \\ 1 & 2 & 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & 2 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 2 & 1 \\ 0 & 0 & 0 & 0 & \cdots & 1 & 2 \end{pmatrix}, \quad n \geq 3.$$
Δείξτε ότι $\det A_n = 2\det A_{n-1} - \det A_{n-2}$ για $n \geq 4$ και συμπεράνετε ότι $\det A_n = n + 1$.

12. Έστω A ένα $n \times n$ μητρώο τέτοιο ώστε το άθροισμα των στοιχείων κάθε γραμμής του να είναι ίσο με μηδέν. Δείξτε ότι $\det A = 0$.
Υπόδειξη: Αν $\boldsymbol{x}$ είναι το διάνυσμα του $\mathbb{R}^n$ του οποίου όλες οι συνιστώσες είναι ίσες με ένα, τότε $A\boldsymbol{x} =$;

13. Αν f και g είναι πραγματικές συναρτήσεις για τις οποίες υπάρχουν οι πρώτες και δεύτερες παράγωγοι, ορίζουμε
$$h(t) = \begin{vmatrix} f(t) & g(t) \\ f'(t) & g'(t) \end{vmatrix}.$$
Δείξτε ότι
$$h'(t) = \begin{vmatrix} f(t) & g(t) \\ f''(t) & g''(t) \end{vmatrix}.$$

14. Για ένα τετραγωνικό μητρώο $A = \left(a_{ij}\right)$ ορίζουμε το **ίχνος** του (trace), $\operatorname{tr} A$ ως το άθροισμα των στοιχείων της κύριας διαγωνίου του, $\operatorname{tr} A = a_{11} + a_{22} + \cdots + a_{nn}$. Για ένα 2×2 μητώο A δείξτε ότι
$$\det A = \frac{1}{2}\begin{vmatrix} \operatorname{tr} A & 1 \\ \operatorname{tr}(A^2) & \operatorname{tr} A \end{vmatrix} = \frac{1}{2}[(\operatorname{tr} A)^2 - \operatorname{tr}(A^2)].$$

15. Εάν A είναι ένα $n \times n$ αντιστρέψιμο μητρώο και Ψ είναι το μητρώο των συμπαραγόντων του A δείξτε ότι
$$\det \Psi = (\det A)^{n-1}.$$

16. Εάν Q είναι ένα $n \times n$ ορθογώνιο μητρώο δείξτε ότι $\det Q = \pm 1$. Εξετάστε αν ισχύει και το αντίστροφο, δηλαδή αν για ένα τετραγωνικό μητρώο A ισχύει $\det A = \pm 1$, τότε το A είναι ορθογώνιο.

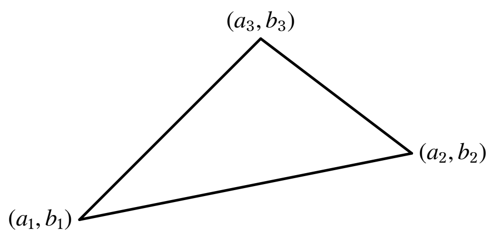
17. Δείξτε ότι κάθε 2×2 πραγματικό ορθογώνιο μητρώο με ορίζουσα ίση με -1 μπορεί να γραφεί ως γινόμενο
$$\begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$
για κάποιο πραγματικό αριθμό θ.

18. Εάν τα σημεία του επιπέδου (x_1, y_1), (x_2, y_2) και (x_3, y_3) είναι συνευθειακά, δείξτε ότι
$$\begin{vmatrix} x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \\ x_3 & y_3 & 1 \end{vmatrix} = 0.$$

19. Αποδείξτε ότι η εξίσωση της ευθείας η οποία περιέχει τα σημεία (x_1, y_1) και (x_2, y_2) δίνεται με τη σχέση
$$\begin{vmatrix} x & y & 1 \\ x_1 & y_1 & 1 \\ x_2 & y_2 & 1 \end{vmatrix} = 0.$$

20. **Η ορίζουσα Vandermonde**. Εάν x_1, x_2, x_3 είναι πραγματικοί αριθμοί δείξτε ότι
$$V_3 = \begin{vmatrix} 1 & x_1 & x_1^2 \\ 1 & x_2 & x_2^2 \\ 1 & x_3 & x_3^2 \end{vmatrix} = (x_3 - x_2)(x_3 - x_1)(x_2 - x_1).$$
Γενικότερα αν $x_1, x_2, \ldots, x_n$ είναι αριθμοί, αποδείξτε με επαγωγή ότι
$$V_n = \begin{vmatrix} 1 & x_1 & \cdots & x_1^{n-1} \\ 1 & x_2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \cdots & \vdots \\ 1 & x_n & \cdots & x_n^{n-1} \end{vmatrix} = \prod_{i<j} (x_j - x_i).$$
Το σύμβολο $\prod$, είναι το αντίστοιχο του $\sum$, και δηλώνει το γινόμενο (product). Έτσι η ποσότητα στο δεξί μέλος της ισότητας είναι το γινόμενο των πρωτοβάθμιων όρων $(x_j - x_i)$ με $i < j$, βλέπε την 3×3 περίπτωση.
Υπόδειξη: Πολλαπλασιάστε κάθε στήλη με x_1 και αφαιρέστε την από την επόμενη στήλη στα δεξιά. Δείξτε ότι
$$V_n(x_1, \ldots, x_n) = (x_n - x_1) \cdots (x_2 - x_1) V_{n-1}(x_2, \ldots, x_n).$$

21. Εάν (a_1, b_1), (a_2, b_2) και (a_3, b_3) είναι σημεία του επιπέδου όπως στο σχήμα, δείξτε ότι



το εμβαδόν του τριγώνου με κορυφές τα σημεία αυτά δίνεται από τη σχέση

$$E = \frac{1}{2}\begin{vmatrix} a_1 & b_1 & 1 \\ a_2 & b_2 & 1 \\ a_3 & b_3 & 1 \end{vmatrix}.$$

Υπόδειξη: Υποθέστε ότι το τρίγωνο βρίσκεται στο πρώτο τεταρτημόριο, ισοδύναμα $b_j > 0,\ j = 1, 2, 3$.

22. Θεωρήστε τρίγωνο με γωνίες α, β, γ και απέναντι πλευρές a, b, c αντίστοιχα.

(α΄) Δείξτε ότι

$$\begin{aligned} b\cos\gamma + c\cos\beta &= a \\ c\cos\alpha + a\cos\gamma &= b \\ a\cos\beta + b\cos\alpha &= c. \end{aligned}$$

(β΄) Αποδείξτε το νόμο του συνημιτόνου $a^2 = b^2 + c^2 - 2bc\cos\alpha$, ισοδύναμα χρησιμοποιήστε τον κανόνα του Cramer για να δείξετε ότι

$$\cos\alpha = \frac{b^2 + c^2 - a^2}{2bc}.$$

Κεφάλαιο 9

Ιδιοτιμές, ιδιοδιανύσματα και διαγωνοποίηση

9.1 Ιδιοτιμές και ιδιοδιανύσματα

Για ένα τετραγωνικό μητρώο A συμβαίνει να υπάρχουν χαρακτηριστικά μη μηδενικά διανύσματα $\boldsymbol{x}$ ώστε

$$A\boldsymbol{x} = \lambda \boldsymbol{x}, \tag{9.1}$$

όπου λ είναι κάποια σταθερά. Για παράδειγμα για

$$A = \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}, \qquad \boldsymbol{x}_1 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}, \qquad \boldsymbol{x}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

έχουμε ότι

$$A\boldsymbol{x}_1 = \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} -3 \\ 1 \end{pmatrix} = \begin{pmatrix} -3 \\ 1 \end{pmatrix} = \boldsymbol{x}_1, \qquad \lambda = 1$$

και

$$A\boldsymbol{x}_2 = \begin{pmatrix} 2 & 3 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} 2 \\ 0 \end{pmatrix} = 2\begin{pmatrix} 1 \\ 0 \end{pmatrix} = 2\boldsymbol{x}_2, \qquad \lambda = 2.$$

Η σχέση (9.1) εκφράζει μια σημαντική, γεωμετρική ιδιότητα αυτών των χαρακτηριστικών, για το μητρώο διανυσμάτων, συγκεκριμένα το αποτέλεσμα του πολλαπλασιασμού από αριστερά ενός τέτοιου διανύσματος με το μητρώο είναι ένα συγγραμικό διάνυσμα, στην κατεύθυνση του αρχικού αν $\lambda > 0$, ή στην αντίθετη αν $\lambda < 0$.

Ορισμός 9.1. Έστω A ένα $n \times n$ μητρώο και έστω ότι για κάποια σταθερά λ υπάρχει μη μηδενικό διάνυσμα $\boldsymbol{x}$ ώστε $A\boldsymbol{x} = \lambda\boldsymbol{x}$. Η σταθερά λ λέγεται **ιδιοτιμή** (eigenvalue) του A και το μη μεδενικό διάνυσμα $\boldsymbol{x}$ λέγεται **ιδιοδιάνυσμα** (eigenvector) του A που αντιστοιχεί στην ιδιοτιμή λ.

Λόγω της γεωμετρικής ιδιότητας που εκφράζει η (9.1) είναι σημαντικό να γνωρίζουμε πώς να βρίσκουμε τις ιδιοτιμές και τα αντίστοιχα ιδιοδιανύσματα ενός τετραγωνικού μητρώου. Γράφοντας τη φαινομενικά παράξενη εξίσωση (9.1) στη μορφή

$$A\boldsymbol{x} - \lambda\boldsymbol{x} = \mathbf{0} \Leftrightarrow (A - \lambda I)\boldsymbol{x} = \mathbf{0}, \tag{9.2}$$

βλέπουμε αμέσως ότι το ζητούμενο σχετίζεται με την επίλυση ενός συστήματος το οποίο εξαρτάται από μια παράμετρο, συνεπώς

1. Η σταθερά λ είναι ιδιοτιμή του A αν και μόνο αν το μητρώο $A - \lambda I$ είναι μη αντιστρέψιμο.
2. Το μη μηδενικό διάνυσμα $\boldsymbol{x}$ είναι ιδιοδιάνυσμα του A το οποίο αντιστοιχεί στην ιδιοτιμή λ αν και μόνο αν είναι λύση του συστήματος $(A - \lambda I)\boldsymbol{x} = \mathbf{0}$.

Κατά συνέπεια, ως αποτέλεσμα, της θεωρίας που έχουμε αναπτύξει διατυπώνουμε το

Θεώρημα 9.1. *Αν A είναι ένα τετραγωνικό μητρώο, τότε*

(1) *Η σταθερά λ είναι ιδιοτιμή του A αν και μόνο αν είναι λύση της εξίσωσης* $\det(A - \lambda I) = 0$.

(2) *Το διάνυσμα $\boldsymbol{x}$ είναι ιδιοδιάνυσμα του A το οποίο αντιστοιχεί στην ιδιοτιμή λ αν και μόνο αν είναι μη μηδενικό στοιχείο του μηδενόχωρου του $A - \lambda I$.*

Παράδειγμα 9.1. Διαμορφώνοντας την εξίσωση $\det(A - \lambda I) = 0$ για το μητρώο στην αρχή του κεφαλαίου υπολογίζουμε

$$\det(A - \lambda I) = \begin{vmatrix} 2-\lambda & 3 \\ 0 & 1-\lambda \end{vmatrix} = (2-\lambda)(1-\lambda) = \lambda^2 - 3\lambda + 2,$$

επομένως οι λύσεις της εξίσωσης $\det(A - \lambda I) = 0$, ισοδύναμα οι ιδιοτιμές του A είναι οι $\lambda_1 = 1$ και $\lambda_2 = 2$, όπως είδαμε στην αρχή του κεφαλαίου. Για την εύρεση των ιδιοδιανυσμάτων διαμορφώνουμε τα συστήματα

$$(A - \lambda I)\boldsymbol{x} = \mathbf{0} \Leftrightarrow \begin{pmatrix} 2-\lambda & 3 \\ 0 & 1-\lambda \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$$

όπου η παράμετρος λ παίρνει τις τιμές $\lambda = 1, 2$. Έτσι επιλύουμε

$$\begin{aligned} \lambda = 1, &\quad \begin{pmatrix} 1 & 3 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} x_1 + 3x_2 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad x_1 = -3s, \quad x_2 = s \\ \lambda = 2, &\quad \begin{pmatrix} 0 & 3 \\ 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 3x_2 \\ -x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad x_1 = t, \quad x_2 = 0 \end{aligned}$$

όπου $s, t \in \mathbb{R}$. Για $s = t = 1$ βρίσκουμε, όπως στην εισαγωγή, τα αντίστοιχα ιδιοδιανύσματα

$$\boldsymbol{x}_1 = \begin{pmatrix} -3 \\ 1 \end{pmatrix}, \qquad \boldsymbol{x}_2 = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

Παρατήρηση 9.1. Παρατηρήστε ότι αν το $\boldsymbol{x}$ είναι ιδιοδιάνυσμα του A το οποίο αντιστοιχεί σε μια ιδιοτιμή, έστω λ, και c είναι μια μη μηδενική σταθερά, από την (9.1) έπεται ότι

$$A\boldsymbol{x} = \lambda\boldsymbol{x} \Rightarrow c(A\boldsymbol{x}) = c(\lambda\boldsymbol{x}) \Rightarrow \begin{cases} (cA)\boldsymbol{x} = (c\lambda)\boldsymbol{x} \\ A(c\boldsymbol{x}) = \lambda(c\boldsymbol{x}) \end{cases}$$

ισοδύναμα, αφενός το $\boldsymbol{x}$ είναι ιδιοδιάνυσμα του cA το οποίο αντιστοιχεί στην ιδιοτιμή $c\lambda$, και αφετέρου η λ είναι ιδιοτιμή του A με αντίστοιχο ιδιοδιάνυσμα το $c\boldsymbol{x}$.

Παρατήρηση 9.2. Ένα μητρώο περιστροφής

$$Q = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}$$

περιστρέφει κάθε διάνυσμα $\boldsymbol{x}$ στο $\mathbb{R}^2$ κατά γωνία θ, κατά συνέπεια δεν μπορεί το $Q\boldsymbol{x}$ να είναι ένα διάνυσμα συγγραμμικό με το $\boldsymbol{x}$, εκτός κι αν $\theta = k\pi$, όπου το k είναι ακέραιος αριθμός, ισοδύναμα δεν μπορεί να ισχύει μια σχέση $Q\boldsymbol{x} = \lambda\boldsymbol{x}$ για κάποια σταθερά λ. Θα μπορούσαμε επομένως να πούμε ότι αυτό το μητρώο, όπως και το αντίστοιχο της ανάκλασης δεν έχει ιδιοτιμές και ιδιοδιανύσματα. Ας δούμε όμως που μας οδηγεί η ανάλυση τόσο της εξίσωσης $\det(Q - \lambda I)$ όσο και του συστήματος $(Q - \lambda I)\boldsymbol{x} = \mathbf{0}$. Ας θεωρήσουμε την περίπτωση $\theta = \pi/4$, τότε

$$Q = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 & -1 \\ 1 & 1 \end{pmatrix}$$

και αν $Q' = \sqrt{2}Q$ ισχύει

$$Q\boldsymbol{x} = \lambda\boldsymbol{x} \Leftrightarrow Q'\boldsymbol{x} = \lambda'\boldsymbol{x} \qquad \lambda' = \sqrt{2}\lambda.$$

Θεωρούμε λοιπόν τα ανάλογα προβλήματα για το μητρώο Q'. Έτσι υπολογίζουμε

$$\det(Q' - \lambda I) = \begin{vmatrix} 1-\lambda & -1 \\ 1 & 1-\lambda \end{vmatrix} = (1-\lambda)^2 + 1,$$

το οποίο έχει μιγαδικές ρίζες τις $\lambda_1 = 1+i$ και $\lambda_2 = 1-i$. Επιλύοντας τα αντίστοιχα συστήματα $(Q' - \lambda I)\boldsymbol{x} = \mathbf{0}$ βρίσκουμε

$$\lambda = 1+i, \qquad \begin{pmatrix} -i & -1 \\ 1 & -i \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} -ix_1 - x_2 \\ x_1 - ix_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad x_1 = i\zeta, \quad x_2 = \zeta$$

$$\lambda = 1-i, \qquad \begin{pmatrix} i & -1 \\ 1 & i \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{pmatrix} 1x_1 - x_2 \\ x_1 + ix_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad x_1 = -i\eta, \quad x_2 = \eta$$

όπου $\zeta, \eta \in \mathbb{C}$. Για $\zeta = \eta = 1$ βρίσκουμε τα αντίστοιχα μιγαδικά ιδιοδιανύσματα

$$\boldsymbol{x}_1 = \begin{pmatrix} i \\ 1 \end{pmatrix}, \qquad \boldsymbol{x}_2 = \begin{pmatrix} -i \\ 1 \end{pmatrix}.$$

Κατά συνέπεια η εξίσωση $Q'\boldsymbol{x} = \lambda x$, για πραγματικό μητρώο Q', ενώ δεν έχει πραγματικές λύσεις λ και αντίστοιχα $\boldsymbol{x} \in \mathbb{R}^2$, έχει μιγαδικές λύσεις λ και αντίστοιχα διανύσματα $\boldsymbol{x} \in \mathbb{C}^2$.

Σχόλιο 9.1. Στην εξίσωση $A\boldsymbol{x} = \lambda\boldsymbol{x}$ η σταθερά λ είναι στοιχείο του σώματος από το οποίο προέρχονται τα στοιχεία του μητρώου. Έτσι αν $A \in \mathbb{M}^{n,n}(\mathbb{R})$, τότε η προηγούμενη εξίσωση έχει έννοια για $\lambda \in \mathbb{R}$. Ωστόσο οι τιμές λ που ικανοποιούν την εξίσωση είναι λύσεις της πολυωνυμικής εξίσωσης $\det(A - \lambda I) = 0$ και σύμφωνα με το Θεμελιώδες θεώρημα της Άλγεβρας κάθε πολυωνυμική εξίσωση βαθμού n έχει ακριβώς n ρίζες. Και είναι ακριβώς αυτό το σημείο, δηλαδή η ύπαρξη και εύρεση ιδιοτιμών, στο οποίο οι μιγαδικοί αριθμοί εμφανίζονται ενεργά και όχι απλά ως μία "εξωτική" επιλογή ως σώμα πάνω στο οποίο ορίζεται ένας διανυσματικός χώρος.

Για ένα 2×2 μητρώο A η έκφραση $\det(A - \lambda I)$ είναι ένα πολυώνυμο του λ βαθμού 2. Πράγματι

$$\begin{aligned} \det(A - \lambda I) &= \begin{vmatrix} a_{11} - \lambda & a_{12} \\ a_{21} & a_{22} - \lambda \end{vmatrix} \\ &= (a_{11} - \lambda)(a_{22} - \lambda) - a_{12}a_{21} \\ &= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21}, \end{aligned}$$

και στην περίπτωση όπου το A είναι ένα 3×3 μητρώο υπολογίζουμε

$$\begin{aligned}\det(A-\lambda I) &= \begin{vmatrix} a_{11}-\lambda & a_{12} & a_{13} \\ a_{21} & a_{22}-\lambda & a_{23} \\ a_{31} & a_{32} & a_{33}-\lambda \end{vmatrix} \\ &= (a_{11}-\lambda)\begin{vmatrix} a_{22}-\lambda & a_{23} \\ a_{32} & a_{33}-\lambda \end{vmatrix} - a_{12}\begin{vmatrix} a_{21} & a_{23} \\ a_{31} & a_{33}-\lambda \end{vmatrix} - a_{13}\begin{vmatrix} a_{21} & a_{22}-\lambda \\ a_{31} & a_{32} \end{vmatrix} \\ &= (a_{11}-\lambda)(a_{22}-\lambda)(a_{33}-\lambda) - (a_{11}-\lambda)a_{23}a_{32} + c_1\lambda + c_0\end{aligned}$$

όπου c_1, c_0 είναι κατάλληλες σταθερές. Μπορούμε επομένως να ισχυριστούμε, και θα αποδείξουμε την ορθότητα του ισχυρισμού μας παρακάτω, ότι για ένα $n\times n$ μητρώο A η ορίζουσα $\det(A-\lambda I)$ είναι ένα πολυώνυμο βαθμού n με μεγιστοβάθμιο όρο τον $(-1)^n\lambda^n$.

Λήμμα 9.1. *Εάν $A = \left(a_{ij}\right)$ είναι ένα $n\times n$ μητρώο, τότε ισχύει*

$$\det(A-\lambda I) = (a_{11}-\lambda)(a_{22}-\lambda)\cdots(a_{nn}-\lambda) + p_{n-2}(\lambda), \tag{9.3}$$

όπου $p_{n-2}(\lambda)$ είναι ένα πολυώνυμο του λ βαθμού $n-2$.

Απόδειξη. Μόλις είδαμε ότι ο ισχυρισμός ισχύει για $n = 2$ και $n = 3$, και μια τυπική απόδειξη του γενικού συμπεράσματος γίνεται με επαγωγή όπου η υπόθεση της επαγωγής χρησιμοποιείται όπως στην απόδειξη της περίπτωσης $n = 3$ μέσω της $n = 2$. Ας δούμε όμως αυτή την έκφραση

$$\det(A-\lambda I) = \begin{vmatrix} a_{11}-\lambda & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22}-\lambda & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}-\lambda \end{vmatrix}.$$

Θέτοντας $A' = A-\lambda I$, το ανάπτυγμα της ορίζουσας ως προς την πρώτη γραμμή δίνει

$$\det A' = (a_{11}-\lambda)\det A'_{11} + \sum_{j=2}^{n}(-1)^{1+j}a_{1j}\det A'_{1j}.$$

Το κάθε μητρώο A'_{1j} με $j\geq 2$ περιέχει τους $n-2$ όρους $a_{kk}-\lambda$ με $k\neq 1$ και $k\neq j$ κατά συνέπεια η ορίζουσα $\det A'_{1j}$, ως άθροισμα γινομένων διαφορετικών μεταξύ τους στοιχείων του A'_{1j}, είναι πολυώνυμο βαθμού το πολύ $n-2$, έτσι

$$\det(A-\lambda I) = (a_{11}-\lambda)\det A'_{11} + q_{n-2}(\lambda),$$

όπου $q_{n-2}(\lambda)$ είναι ένα πολυώνυμο βαθμού $n-2$. Με την ίδια λογική έχουμε ότι

$$\begin{aligned}\det A'_{11} &= (a_{22}-\lambda)\det A''_{22} + \sum_{j=3}^{n}(-1)^{2+j}\det A''_{2j} \\ &= (a_{22}-\lambda)\det A''_{22} + q_{n-3}(\lambda),\end{aligned}$$

όπου το A''_{2j} προκύπτει από το A'_{11} με το γνωστό τρόπο. Επομένως

$$\begin{aligned}\det(A-\lambda I) &= (a_{11}-\lambda)[(a_{22}-\lambda)\det A''_{22} + q_{n-3}(\lambda)] + q_{n-2}(\lambda) \\ &= (a_{11}-\lambda)(a_{22}-\lambda)\det A''_{22} + (a_{11}-\lambda)q_{n-3} + q_{n-2}(\lambda) \\ &= (a_{11}-\lambda)(a_{22}-\lambda)\det A''_{22} + \tilde{q}_{n-2}(\lambda)\end{aligned}$$

όπου $\tilde{q}_{n-2}$ είναι ένα πολυώνυμο βαθμού $n-2$. Από αυτό το σημείο δεν είναι δύσκολο να δούμε πώς προκύπτει η (9.3). □

Σημείωση 9.1. Ένας ευκολώτερος τρόπος απόδειξης προκύπτει από τον τύπο του Leibniz, βλέπε Θεώρημα 8.10. Γράφοντας για το $n \times n$ μητρώο A το $A - \lambda I$ ως

$$A - \lambda I = \left(a_{ij} - \delta_{ij}\lambda\right),$$

όπου $\delta_{ij} = 1$ αν $i = j$ και $\delta_{ij} = 0$ αν $i \neq j$, το δέλτα του Kronecker, τότε

$$\det(A - \lambda I) = \sum_{\sigma \in S_n} \pi(\sigma)(a_{1\sigma(1)} - \delta_{1\sigma(1)}\lambda)(a_{2\sigma(2)} - \delta_{2\sigma(2)}\lambda)\cdots(a_{n\sigma(n)} - \delta_{n\sigma(n)}\lambda),$$

όπου S_n είναι η ομάδα των μεταθέσεων του $\{1, 2, \dots, n\}$ και $\pi(\sigma)$ είναι το πρόσημο της μετάθεσης σ.

Παρατήρηση 9.3. Από τις ιδιότητες της ορίζουσας και την (9.3) έπεται ότι

$$\begin{aligned} \det(\lambda I - A) &= (-1)^n \det(A - \lambda I) \\ &= (-1)^n[(-1)^n\lambda^n + \cdots] \\ &= (-1)^{2n}\lambda^n + \cdots \\ &= \lambda^n + \cdots \end{aligned}$$

Ορισμός 9.2. Εάν A είναι ένα $n \times n$ μητρώο ορίζουμε ως **χαρακτηριστικό πολυώνυμο** (characteristic polynomial) του A το πολυώνυμο

$$p_A(\lambda) = \det(\lambda I - A). \tag{9.4}$$

Πολλές φορές αντί του $p_A(\lambda)$ γράφουμε απλά $p(\lambda)$.

Η επιλογή να ορίζεται ως χαρακτηριστικό πολυώνυμο η έκφραση $\det(\lambda I - A)$ και όχι η $\det(A - \lambda I)$ γίνεται ώστε ο συντελεστής του μεγιστοβάθμιου όρου του πολυωνύμου να είναι το 1.

▼ **Πόρισμα 9.1.** *Αν A είναι ένα τετραγωνικό μητρώο οι ιδιοτιμές του A είναι οι ρίζες του χαρακτηριστικού πολυωνύμου, ισοδύναμα οι ρίζες της χαρακτηριστικής εξίσωσης*

$$p_A(\lambda) = \det(\lambda I - A) = 0. \tag{9.5}$$

Παράδειγμα 9.2. Να βρεθούν οι ιδιοτιμές και τα αντίστοιχα ιδιοδιανύσματα του μητρώου

$$A = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 4 \\ 0 & 0 & 4 \end{pmatrix}. \tag{9.6}$$

Διαμορφώνοντας την χαρακτηριστική εξίσωση υπολογίζουμε το χαρακτηριστικό πολυώνυμο

$$p_A(\lambda) = \det\begin{pmatrix} \lambda - 3 & 0 & -1 \\ 0 & \lambda - 3 & -4 \\ 0 & 0 & \lambda - 4 \end{pmatrix} = (\lambda - 3)\begin{vmatrix} \lambda - 3 & -4 \\ 0 & \lambda - 4 \end{vmatrix} = (\lambda - 3)^2(\lambda - 4), \tag{9.7}$$

κατά συνέπεια οι ιδιοτιμές του A είναι $\lambda_1 = \lambda_2 = 3$ και $\lambda_3 = 4$. Στη συνέχεια βρίσκουμε τα ιδιοδιανύσματα του A. Για $\lambda = 3$ έχουμε

$$(3I - A)\boldsymbol{x} = \mathbf{0} \Leftrightarrow \begin{pmatrix} 0 & 0 & -1 \\ 0 & 0 & -4 \\ 0 & 0 & -1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} -x_3 \\ -4x_3 \\ -x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \{x_3 = 0\} \Leftrightarrow \boldsymbol{x} = \begin{pmatrix} r \\ s \\ 0 \end{pmatrix}$$

όπου $r, s \in \mathbb{R}$, ενώ για $\lambda = 4$ βρίσκουμε

$$(4I - A)\boldsymbol{x} = \mathbf{0} \Leftrightarrow \begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -4 \\ 0 & 0 & 0 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

$$\Leftrightarrow \begin{pmatrix} x_1 - x_3 \\ x_2 - 4x_3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \begin{Bmatrix} x_1 = x_3 \\ x_2 = 4x_3 \end{Bmatrix} \Leftrightarrow \boldsymbol{x} = \begin{pmatrix} t \\ 4t \\ t \end{pmatrix},$$

όπου $t \in \mathbb{R}$. Τα ιδιοδιανύσματα που αντιστοιχούν στην ιδιοτιμή $\lambda = 3$ είναι της μορφής

$$\boldsymbol{x} = r\begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} + s\begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \tag{9.8}$$

ισοδύναμα ο μηδενόχωρος του $3I - A$ έχει διάσταση δύο (γιατί;). Έτσι μπορούμε να πάρουμε ως ιδιοδιανύσματα, που αντιστοιχούν στις $\lambda_1 = 3$, $\lambda_2 = 3$ και $\lambda_3 = 4$, τα

$$\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{x}_3 = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix}, \tag{9.9}$$

επιλέγοντας, αντίστοιχα, ($r = 1$ και $s = 0$), ($r = 0$ και $s = 1$), και ($t = 1$). Παρατηρούμε ότι τα ιδιοδιανύσματα $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ είναι γραμμικά ανεξάρτητα, κατά συνέπεια αποτελούν μια βάση για τον $\mathbb{R}^3$.

Σημείωση 9.2. Σε κάθε $n \times n$ μητρώο A αντιστοιχεί μοναδικό χαρακτηριστικό πολυώνυμο, αυτό που ορίζεται ως $p(\lambda) = \det(\lambda I - A)$. Αν τώρα δοθεί πολυώνυμο βαθμού n αναρωτιόμαστε αν υπάρχει $n \times n$ μητρώο τέτοιο ώστε το δοσμένο πολυώνυμο να είναι το χαρακτηριστικό πολυώνυμο του μητρώου. Η απάντηση είναι ναι αρκεί ο συντελεστής του μεγιστοβάθμιου όρου να είναι ένα. Έτσι αν

$$p(x) = x^n + a_{n-1}x^{n-1} + \cdots + a_1 x + a_0$$

αποδεικνύεται ότι το $n \times n$ μητρώο

$$A = \begin{pmatrix} 0 & 0 & \ldots & 0 & -a_0 \\ 1 & 0 & \ldots & 0 & -a_1 \\ 0 & 1 & \ldots & 0 & -a_2 \\ \vdots & \vdots & \ldots & \vdots & \\ 0 & 0 & \ldots & 1 & -a_{n-1} \end{pmatrix} \tag{9.10}$$

έχει ως χαρακτηριστικό πολυώνυμο το p. Δείχνουμε την περίπτωση $n = 2$ και αφήνουμε τη γενική περίπτωση σαν άσκηση. Πράγματι αν

$$A = \begin{pmatrix} 0 & -a_0 \\ 1 & -a_1 \end{pmatrix},$$

τότε

$$\det(\lambda I - A) = \det \begin{pmatrix} \lambda & a_0 \\ -1 & \lambda + a_1 \end{pmatrix} = \lambda(\lambda + a_1) + a_0 = \lambda^2 + a_1\lambda + a_0.$$

Το μητρώο στην (9.10) λέγεται συνοδευτικό μητρώο.

Ιδιόχωροι

Θυμίζουμε ότι ο μηδενόχωρος null A ενός μητρώου A είναι διανυσματικός υπόχωρος.

Ορισμός 9.3. Εάν A είναι ένα τετραγωνικό μητρώο και λ είναι μια ιδιοτιμή του A, ο μηδενόχωρος null$(\lambda I - A)$ λέγεται **ιδιόχωρος** (eigenspace) του A που αντιστοιχεί στην ιδιοτιμή λ. Η διάσταση του ιδιόχωρου null$(\lambda I - A)$ λέγεται **γεωμετρική πολλαπλότητα** (geometric multiplicity) της ιδιοτιμής λ.

Παρατήρηση 9.4. Για ένα πολυώνυμο βαθμού n με πραγματικούς ή μιγαδικούς συντελεστές

$$p(x) = a_n x^n + a_{n-1} x^{n-1} + \cdots + a_1 x + a_0$$

ως συνέπεια του Θεμελιώδους θεωρήματος της Άλγεβρας προκύπτει ότι υπάρχουν πραγματικοί ή μιγαδικοί αριθμοί $r_1, r_2, \ldots, r_n$, οι ρίζες, όχι αναγκαστικά ξένες μεταξύ τους ώστε

$$p(x) = a(x - r_1)(x - r_2) \cdots (x - r_n)$$

ή αριθμοί $\rho_1, \rho_2, \ldots, \rho_m$ με $m \le n$, ξένοι μεταξύ τους και θετικοί ακέραιοι $k_1, k_2, \ldots, k_m$ ώστε

$$p(x) = a(x - \rho_1)^{k_1}(x - \rho_2)^{k_2} \cdots (x - \rho_m)^{k_m}, \qquad k_1 + k_2 + \cdots + k_m = n.$$

Ο θετικός ακέραιος k_j λέγεται **αλγεβρική πολλαπλότητα** (algebraic multiplicity) της ρίζας ρ_j, $j = 1, 2, \ldots, m$. Έτσι στο Παράδειγμα 9.1 η αλγεβρική πολλαπλότητα της ιδιοτιμής $\lambda = 1$ είναι ένα και είναι ίση με τη γεωμετρική πολλαπλότητα. Το ίδιο συμβαίνει και με την ιδιοτιμή $\lambda = 2$. Στο δε Παράδειγμα 9.2 η ιδιοτιμή $\lambda = 3$ έχει αλγεβρική και γεωμετρική πολλαπλότητα ίση με δύο, ενώ η ιδιοτιμή $\lambda = 4$ έχει αλγεβρική και γεωμετρική πολλαπλότητα ίση με ένα.

Παράδειγμα 9.3. Να βρεθούν οι ιδιοτιμές και οι αντίστοιχοι ιδιόχωροι του μητρώου

$$A = \begin{pmatrix} 3 & 1 & 0 \\ 0 & 3 & 4 \\ 0 & 0 & 4 \end{pmatrix}. \tag{9.11}$$

Το χαρακτηριστικό πολυώνυμο του A είναι

$$p_A(\lambda) = \det \begin{pmatrix} \lambda - 3 & -1 & 0 \\ 0 & \lambda - 3 & -4 \\ 0 & 0 & \lambda - 4 \end{pmatrix} = (\lambda - 3) \begin{vmatrix} \lambda - 3 & -4 \\ 0 & \lambda - 4 \end{vmatrix} = (\lambda - 3)^2(\lambda - 4),$$

κατά συνέπεια οι ιδιοτιμές του A είναι $\lambda_1 = 3$ με αλγεβρική πολλαπλότητα δύο και $\lambda_2 = 4$ με αλγεβρική πολλαπλότητα ένα. Στη συνέχεια βρίσκουμε τα ιδιοδιανύσματα του A. Για $\lambda = 3$ έχουμε

$$(3I - A)\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \begin{pmatrix} 0 & -1 & 0 \\ 0 & 0 & -4 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} -x_2 \\ -4x_3 \\ -x_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \boldsymbol{x} = \begin{pmatrix} s \\ 0 \\ 0 \end{pmatrix}$$

όπου $s \in \mathbb{R}$, ενώ για $\lambda = 4$ βρίσκουμε

$$(4I - A)\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & -4 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} x_1 - x_2 \\ x_2 - 4x_3 \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \boldsymbol{x} = \begin{pmatrix} 4t \\ 4t \\ t \end{pmatrix},$$

όπου $t \in \mathbb{R}$. Έτσι ο ιδιόχωρος για την ιδιοτιμή $\lambda = 3$ είναι ο

$$\mathcal{E}_1 = \operatorname{span}\left\{ \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \right\},$$

ισοδύναμα η γεωμετρική πολλαπλότητα της ιδιοτιμής $\lambda = 3$ είναι ένα, και ο ιδιόχωρος για την $\lambda = 4$ είναι ο

$$\mathcal{E}_2 = \operatorname{span}\left\{ \begin{pmatrix} 4 \\ 4 \\ 1 \end{pmatrix} \right\} \tag{9.12}$$

κατά συνέπεια η γεωμετρική πολλαπλότητα της ιδιοτιμής $\lambda_2 = 4$ είναι επίσης ένα.

9.2 Το χαρακτηριστικό πολυώνυμο

Για το μητρώο

$$A = \begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}$$

υπολογίζουμε το χαρακτηριστικό πολυώνυμο

$$\begin{aligned} p_A(\lambda) &= \det(\lambda I - A) \\ &= \det \begin{pmatrix} \lambda - a_{11} & -a_{12} \\ -a_{21} & \lambda - a_{22} \end{pmatrix} \\ &= \lambda^2 - (a_{11} + a_{22})\lambda + a_{11}a_{22} - a_{12}a_{21} \end{aligned} \tag{9.13}$$

και παρατηρούμε τα εξής:

1. Ο συντελεστής του μεγιστοβάθμιου-δευτεροβάθμιου όρου λ^2 είναι η μονάδα, όπως έχουμε ήδη συζητήσει.

2. Ο συντελεστής του πρωτοβάθμιου όρου λ είναι το άθροισμα των στοιχείων της κυρίας διαγωνίου του A επί -1, ισοδύναμα το $-$ίχνος του A, $-\operatorname{tr} A$.

3. Ο σταθερός όρος είναι η ορίζουσα του A, $\det A$, αποτέλεσμα το οποίο προκύπτει, εξάλλου, και από τον ορισμό του χαρακτηριστικού πολυωνύμου

$$p_A(0) = \det(O - A) = \det(-A) = (-1)^2 \det A = \det A. \tag{9.14}$$

Επομένως μπορούμε να γράψουμε

$$p_A(\lambda) = \lambda^2 - (\operatorname{tr} A)\lambda + \det A \tag{9.15}$$

Επιπλέον αν λ_1 και λ_2 είναι οι ιδιοτιμές του A, από το Πόρισμα 9.1 έπεται ότι

$$p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2) = \lambda^2 - (\lambda_1 + \lambda_2)\lambda + \lambda_1\lambda_2 \tag{9.16}$$

Συμπέρασμα. Συγκρίνοντας τις (9.15) και (9.16) βλέπουμε ότι

$$\operatorname{tr} A = \lambda_1 + \lambda_2, \qquad \& \qquad \det A = \lambda_1\lambda_2.$$

Τα αποτελέσματα αυτά γενικεύονται για μητρώα μεγαλύτερου μεγέθους. Θα χρειαστούμε το παρακάτω αποτέλεσμα.

Θεώρημα 9.2. *Έστω A ένα $n \times n$ μητρώο και έστω*

$$p_A(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$$

το χαρακτηριστικό πολυώνυμο του A. Εάν $\lambda_1, \lambda_2, \ldots, \lambda_n$ είναι οι, όχι αναγκαστικά ξένες μεταξύ τους, ιδιοτιμές του A, τότε για τους συντελεστές του χαρακτηριστικού πολυωνύμου ισχύει

$$c_{n-1} = -\operatorname{tr} A = -(\lambda_1 + \lambda_2 + \cdots + \lambda_n). \tag{9.17}$$

$$c_0 = (-1)^n \det A = (-1)^n \lambda_1\lambda_2\cdots\lambda_n. \tag{9.18}$$

Απόδειξη. Η ισότητα μεταξύ των μελών στα άκρα των (9.17) και (9.18) προκύπτει από τις σχέσεις μεταξύ ριζών και συντελεστών (σχέσεις του Vieta) του χαρακτηριστικού πολυωνύμου αφού

$$p_A(\lambda) = (\lambda - \lambda_1)(\lambda - \lambda_2)\cdots(\lambda - \lambda_n). \tag{9.19}$$

Ο όρος $c_{n-1}\lambda^{n-1}$ του χαρακτηριστικού πολυωνύμου, σύμφωνα με το Λήμμα 9.1, περιέχεται στο ανάπτυγμα του $(\lambda - a_{11})(\lambda - a_{22})\ldots(\lambda - a_{nn})$, έτσι αναπτύσοντας[1] βρίσκουμε

$$(\lambda - a_{11})(\lambda - a_{22})\ldots(\lambda - a_{nn}) = \lambda^n - (a_{11} + a_{22} + \cdots + a_{nn})\lambda^{n-1} + \text{“πολυώνυμο βαθμού } n-2\text{”}$$

ενώ από την (9.19) υπολογίζουμε ανάλογα

$$p_A(\lambda) = \lambda^n - (\lambda_1 + \lambda_2 + \cdots + \lambda_n)\lambda^{n-1} + \text{“πολυώνυμο βαθμού } n-2\text{”},$$

κατά συνέπεια

$$c_{n-1} = -(\lambda_1 + \lambda_2 + \cdots + \lambda_n) = -(\lambda_1 + \lambda_2 + \cdots + \lambda_n)$$

που είναι η (9.17). Από τον ορισμό του χαρακτηριστικού πολυωνύμου και την (9.19) έπεται ότι

$$p_A(0) = \det(-A) = (-\lambda_1)(-\lambda_2)\cdots(-\lambda_n) \Rightarrow$$
$$c_0 = (-1)^n \det A = (-1)^n \lambda_1\lambda_2\cdots\lambda_n$$

που είναι η (9.18). □

[1] χρησιμοποιώντας κάθε φορά το λ από τις $n-1$ παρενθέσεις και το a_{kk}, για $k = 1, 2, \ldots, n$, από την παρένθεση που περισεύει

Θεώρημα 9.3. *Το $n \times n$ μητρώο A είναι αντιστρέψιμο αν και μόνο αν το μηδέν δεν είναι ιδιοτιμή του.*

Απόδειξη. Αν το A είναι αντιστρέψιμο, τότε $\det A \neq 0$, κατά συνέπεια $p_A(0) = (-1)^n \det A \neq 0$, κατά συνέπεια το 0 δεν είναι ιδιοτιμή του A. Αν τώρα το μηδέν δεν είναι ιδιοτιμή του A, τότε $p_A(0) \neq 0$, αλλά $p_A(0) = (-1)^n \det A$, επομένως $\det A \neq 0$, κατά συνέπεια το A είναι αντιστρέψιμο. □

Το θεώρημα Cayley-Hamilton

Εάν A είναι ένα $n \times n$ μητρώο και $p_A(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$ το χαρακτηριστικό πολυώνυμο του A, μας ενδιαφέρει το ιδιαίτερο μητρώο

$$p_A(A) = A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I.$$

Ας υπολογίσουμε αυτό το μητρώο για ένα τυπικό 2×2 A. Στη περίπτωση αυτή το χαρακτηριστικό πολυώνυμο είναι το $p_A(\lambda) = \lambda^n - (\operatorname{tr} A)\lambda + \det A$, επομένως

$$\begin{aligned} p_A(A) &= \begin{pmatrix} a & b \\ c & d \end{pmatrix}^2 - (a+d)\begin{pmatrix} a & b \\ c & d \end{pmatrix} + (ad-bc)\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} a^2+bc & ab+bd \\ ac+cd & bc+d^2 \end{pmatrix} - \begin{pmatrix} a^2+ad & ab+bd \\ ac+cd & ad+d^2 \end{pmatrix} + \begin{pmatrix} ad-bc & 0 \\ 0 & ad-bc \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}. \end{aligned}$$

Το αποτέλεσμα δεν είναι τυχαίο, και μάλιστα ισχύει για κάθε τετραγωνικό μητρώο οποιουδήποτε μεγέθους και αυτό εκφράζει το διάσημο θεώρημα των Cayley και Hamilton.

Θεώρημα 9.4 (Cayley-Hamilton). *Κάθε τετραγωνικό μητρώο μηδενίζει το χαρακτηριστικό του πολυώνυμο, δηλαδή αν A είναι ένα τετραγωνικό μητρώο και $p_A(\lambda)$ είναι το χαρακτηριστικό του πολυώνυμο, τότε το μητρώο $p_A(A)$ είναι το μηδενικό.*

Απόδειξη. Δίνουμε την αποδειξη στην ειδική περίπτωση κατά την οποία το $n \times n$ μητρώο A έχει n το πλήθος γραμμικά ανεξάρτητα ιδιοδιανύσματα. Αν $\lambda_1, \lambda_2, \ldots, \lambda_n$ είναι οι όχι κατ' ανάγκη ξένες μεταξύ τους ιδιοτιμές του A και $\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n$ τα αντίστοιχα γραμμικά ανεξάρτητα ιδιοδιανύσματα, τότε

$$p_A(A) = (A - \lambda_1 I)(A - \lambda_2 I) \cdots (A - \lambda_n I). \tag{9.20}$$

Οι πρωτοβάθμιοι παράγοντες στην (9.20) αντιμετατίθενται. Πράγματι ως αποτέλεσμα των ιδιοτήτων των δυνάμεων μητρώου βλέπουμε ότι

$$(A - \lambda_i I)(A - \lambda_j I) = A^2 - (\lambda_i + \lambda_j)A + \lambda_i\lambda_j I = (A - \lambda_j I)(A - \lambda_i I).$$

Έτσι από ην (9.20) βρίσκουμε

$$p_A(A)\boldsymbol{x}_k = \left[\prod_{i\neq k}(A - \lambda_i I)\right](A - \lambda_k I)\boldsymbol{x}_k = \prod_{i\neq k}(A - \lambda_i I)\mathbf{0} = \mathbf{0} \tag{9.21}$$

για $k = 1, 2, \ldots, n$, ισοδύναμα

$$\operatorname{null} A = \{\boldsymbol{x}_1, \boldsymbol{x}_2, \ldots, \boldsymbol{x}_n\} = \mathbb{R}^n,$$

αφού τα ιδιοσιανύσματα $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$ αποτελούν μαι βάση για τον $\mathbb{R}^n$, κατά συνέπεια $A = O$, το μηδενικό μητρώο (γιατί;).

Για την απόδειξη στη γενική περίπτωση παραπέμπουμε στο [5] ή στο [7]. □

Μια από τις συνέπειες του θεωρήματος είναι ότι δυνάμεις ενός μητρώου μπορούν να εκφρασθούν ως γραμμικοί συνδυασμοί διαδοχικών δυνάμεων του ίδιου μητρώου.

▼ **Πόρισμα 9.2.** *Εάν A είναι ένα $n \times n$ μητρώο και $p_A(\lambda) = \lambda^n + c_{n-1}\lambda^{n-1} + \cdots + c_1\lambda + c_0$ είναι το χαρακτηριστικό πολυώνυμο του A, τότε*

(1) $A^n = -c_{n-1}A^{n-1} - \cdots - c_1 A - c_0 I$.

(2) $A^{k+n} = -c_{n-1}A^{k+n-1} - \cdots - c_1 A^{k+1} - c_0 A^k$, *για κάθε φυσικό αριθμό k.*

(3) *Εάν το A είναι αντιστρέψιμο, τότε*

$$A^{-1} = -\frac{1}{c_0}(A^{n-1} + c_{n-1}A^{n-2} + \cdots + c_1 I). \tag{9.22}$$

Απόδειξη. Η βασική σχέση είναι το αποτέλεσμα του Θεωρήματος Cayley-Hamilton

$$A^n + c_{n-1}A^{n-1} + \cdots + c_1 A + c_0 I = O, \tag{9.23}$$

από την οποία έπονται αμέσως οι (1) και (2). Αν το A είναι αντιστρέψιμο, τότε $c_0 = \det A \neq 0$, οπότε πολλαπλασιάζοντας αρχικά την (9.23) με A^{-1} και στη συνέχεια λύνοντας ως προς A^{-1} παίρνουμε την (9.22). □

Παρατήρηση 9.5. Εάν A είναι ένα $n \times n$ μητρώο και $\boldsymbol{x}$ είναι ένα μη μηδενικό διάνυσμα στον $\mathbb{R}^n$, τότε τα $n+1$ διανύσματα

$$A^n\boldsymbol{x}, \quad A^{n-1}\boldsymbol{x}, \quad \dots, \quad A\boldsymbol{x}, \quad \boldsymbol{x}$$

είναι γραμμικά εξαρτημένα, κατά συνέπεια υπάρχουν σταθερές $b_0, b_1, \dots, b_n$, όχι όλες ίσες με μηδέν, ώστε

$$\begin{gathered} b_n A^n\boldsymbol{x} + b_{n-1}A^{n-1}\boldsymbol{x} + \cdots + b_1 A\boldsymbol{x} + b_0\boldsymbol{x} = \mathbf{0} \\ (b_n A^n + b_{n-1}A^{n-1} + \cdots + b_1 A + b_0 I)\boldsymbol{x} = \mathbf{0}. \end{gathered}$$

Το Θεώρημα Cayley-Hamilton μας λέει ότι υπάρχουν σταθερές $b_0, b_1, \dots, b_n$ οι οποίες είναι οι ίδιες για όλα τα διανύσματα $\boldsymbol{x}$ του $\mathbb{R}^n$.

9.3 Διαγωνοποιήσιμα μητρώα

Ας υποθέσουμε ότι το 3×3 μητρώο A έχει τρία γραμμικά ανεξάρτητα ιδιοδυανύσματα $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ τα οποία αντιστοιχούν στις ιδιοτιμές $\lambda_1, \lambda_2, \lambda_3$, τότε οι επιμέρους σχέσεις $A\boldsymbol{x}_k = \lambda_k\boldsymbol{x}_k$, με $k = 1, 2, 3$, συμπτύσονται στην

$$\begin{pmatrix} A\boldsymbol{x}_1 & A\boldsymbol{x}_2 & A\boldsymbol{x}_3 \end{pmatrix} = \begin{pmatrix} \lambda_1\boldsymbol{x}_1 & \lambda_2\boldsymbol{x}_2 & \lambda_3\boldsymbol{x}_3 \end{pmatrix} \Leftrightarrow A\begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{x}_1 & \boldsymbol{x}_2 & \boldsymbol{x}_3 \end{pmatrix}\begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{pmatrix}$$

έτσι γράφοντας Λ για το διαγώνιο μητρώο με τις ιδιοτιμές $\lambda_1, \lambda_2, \lambda_3$ στη διαγώνιο και X για το μητρώο με στήλες τα αντίστοιχα ιδιοδιανύσματα $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3$ έχουμε

$$AX = X\Lambda \Rightarrow A = X\Lambda X^{-1}, \tag{9.24}$$

αφού οι στήλες του X είναι γραμμικά ανεξάρτητες. Ένα από τα πλεονεκτήματα της δυνατότητας γραφής του μητρώου όπως στην (9.24) είναι ο υπολογισμός δυνάμεων του A. Πράγματι

$$A^2 = (X\Lambda X^{-1})^2 = (X\Lambda X^{-1})(X\Lambda X^{-1}) = X\Lambda^2 X^{-1}$$

και επαγωγικά

$$A^n = X\Lambda^n X^{-1} = X\begin{pmatrix} \lambda_1^n & 0 & 0 \\ 0 & \lambda_2^n & 0 \\ 0 & 0 & \lambda_3^n \end{pmatrix} X^{-1},$$

κατά συνέπεια ο υπολογισμός δυνάμεων του A προκύπτει με απλή παρατήρηση. Τα αποτελέσματα αυτά γενικεύονται σε μητρώα οιουδήποτε μεγέθους.

Ορισμός 9.4. Ένα τετραγωνικό μητρώο A λέγεται **διαγωνοποιήσιμο** (diagonalizable) εάν υπάρχουν διαγώνιο μητρώο D και αντιστρέψιμο μητρώο M, ιδίου μεγέθους με το A, ώστε

$$A = MDM^{-1}. \tag{9.25}$$

Παρατήρηση 9.6. Αν ένα τετραγωνικό μητρώο είναι διαγωνοποιήσιμο, τότε από την (9.25) έπεται ότι

$$AM = MD$$

όπου το D είναι ένα διαγώνιο μητρώο. Αν $d_1, d_2, \dots, d_n$ είναι τα στοιχεία της διαγωνίου του D και $\boldsymbol{\xi}_1, \boldsymbol{\xi}_2, \dots \boldsymbol{\xi}_n$ είναι οι στήλες του M, τότε

$$A\boldsymbol{\xi}_k = d_k\boldsymbol{\xi}_k, \qquad k = 1, 2, \dots, n,$$

κατά συνέπεια η ιδιότητα της διαγωνοποιησιμότητας μητρώου σχετίζεται άμεσα με τις ιδιοτιμές και τα ιδιοδιανύσματα του μητρώου.

Στην αρχή της ενότητας, ουσιαστικά, δείξαμε το αποτέλεσμα που ακολουθεί.

Θεώρημα 9.5. *Εάν το $n \times n$ μητρώο A έχει n γραμμικά ανεξάρτητα ιδιοδιανύσματα, τότε είναι διαγωνοποιήσιμο. Συγκεκριμένα αν Λ είναι το διαγώνιο μητρώο με τις ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_n$ στην κύρια διαγώνιο και X είναι το μητρώο με στήλες τα αντίστοιχα γραμμικά ανεξάρτητα ιδιοδιανύσματα $\boldsymbol{x}_1, \boldsymbol{x}_2, \dots, \boldsymbol{x}_n$, τότε*

$$A = X\Lambda X^{-1}. \tag{9.26}$$

Ισχύει και το αντίστροφο, δηλαδή αν το $n \times n$ μητρώο A είναι διαγωνοποιήσιμο, τότε έχει n γραμμικά ανεξάρτητα ιδιοδιανύσματα.

Απόδειξη. Αν το μητρώο A έχει n γραμμικά ανεξάρτητα ιδιοδιανύσματα, τότε η (9.26) είναι η γενίκευση του αποτελέσματος στην αρχή της παραγράφου. Αν το A είναι διαγωνοποιήσιμο και $A = MDM^{-1}$ με D διαγώνιο, τότε $AM = MD$, κατά συνέπεια οι γραμμικά ανεξάρτητες στήλες του M είναι ιδιοδιανύσματα του A με αντίστοιχες ιδιοτιμές τα αντίστοιχα διαγώνια στοιχεία του D. □

Παράδειγμα 9.4. Στο Παράδειγμα 9.2 δείξαμε ότι το μητρώο

$$A = \begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 4 \\ 0 & 0 & 4 \end{pmatrix}$$

έχει ως χαρακτηριστικό πολυώνυμο το $p_A(\lambda) = (\lambda - 3)^2(\lambda - 4)$, κατά συνέπεια οι ιδιοτιμές του είναι $\lambda_1 = \lambda_2 = 3$ και $\lambda_3 = 4$, και ως αντίστοιχα ιδιοδιανύσματα βρήκαμε τα

$$\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \qquad \boldsymbol{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{x}_3 = \begin{pmatrix} 1 \\ 4 \\ 1 \end{pmatrix},$$

Έτσι

$$\begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 4 \\ 0 & 0 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}$$

κατά συνέπεια, δεδομένου ότι τα ιδιοδιανύσματα είναι γραμμικά ανεξάρτητα

$$\begin{aligned} \begin{pmatrix} 3 & 0 & 1 \\ 0 & 3 & 4 \\ 0 & 0 & 4 \end{pmatrix} &= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}^{-1} \\ &= \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 4 \\ 0 & 0 & 1 \end{pmatrix}\begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 4 \end{pmatrix}\begin{pmatrix} 1 & 0 & -1 \\ 0 & 1 & -4 \\ 0 & 0 & 1 \end{pmatrix} \end{aligned}$$

επαληθεύοντας το Θεώρημα 9.5.

Το φυσιολογικό ερώτημα που εγείρεται είναι πότε τα ιδιοδιανύσματα ενός μητρώου είναι γραμμικά ανεξάρτητα.

Θεώρημα 9.6. *Ιδιοδιανύσματα ενός μητρώου A τα οποία αντιστοιχούν σε ξένες μεταξύ τους ιδιοτιμές είναι γραμμικά ανεξάρτητα.*

Απόδειξη. Αν $\lambda \neq \mu$ είναι ιδιοτιμές του μητρώου A με $\boldsymbol{x}, \boldsymbol{y}$ να είναι τα αντίστοιχα ιδιοδιανύσματα, θεωρούμε την εξίσωση

$$c_1\boldsymbol{x} + c_2\boldsymbol{y} = \boldsymbol{0}. \tag{9.27}$$

Θα αποδείξουμε ότι $c_1 = c_2 = 0$. Πολλαπλασιάζοντας και τα δύο μέλη της (9.27) με το μητρώο A παίρνουμε

$$A(c_1\boldsymbol{x} + c_2\boldsymbol{y}) = A\boldsymbol{0} \Rightarrow c_1A\boldsymbol{x} + c_2A\boldsymbol{y} = \boldsymbol{0}$$

επομένως

$$c_1\lambda\boldsymbol{x} + c_2\mu\boldsymbol{y} = \boldsymbol{0}. \tag{9.28}$$

Αφού $\lambda \neq \mu$ μία τουλάχιστον από τις ιδιοτιμές είναι διάφορη του μηδενός. Υποθέτοντας ότι $\lambda \neq 0$ πολλαπλασιάζουμε την (9.27) με λ και στη συνέχεια συγκρίνοντας το αποτέλεσμα με την (9.28) παίρνουμε

$$c_2(\lambda - \mu)\boldsymbol{y} = \boldsymbol{0} \Rightarrow c_2 = 0,$$

δεδομένου ότι $\lambda \neq \mu$ και $\boldsymbol{y} \neq \mathbf{0}$ ως ιδιοδιάνυσμα. Το αποτέλεσμα αυτό οδηγεί μέσω της (9.27) ότι $c_1 = 0$ επίσης. Στο ίδιο συμπέρασμα καταλήγουμε αν υποθέσουμε ότι $\mu \neq 0$. Η απόδειξη είναι πλήρης. $\square$

Παράδειγμα 9.5. Εξετάστε αν το μητρώο A είναι διαγωνοποιήσιμο

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}.$$

Το χαρακτηριστικό πολυώνυμο του A είναι

$$\begin{aligned}
\det(\lambda I - A) &= \begin{vmatrix} \lambda - 4 & -2 & -2 \\ -2 & \lambda - 4 & -2 \\ -2 & -2 & \lambda - 4 \end{vmatrix} \\
&= (\lambda - 4)\begin{vmatrix} \lambda - 4 & -2 \\ -2 & \lambda - 4 \end{vmatrix} + 2\begin{vmatrix} -2 & -2 \\ -2 & \lambda - 4 \end{vmatrix} - 2\begin{vmatrix} -2 & \lambda - 4 \\ -2 & -2 \end{vmatrix} \\
&= (\lambda - 4)[(\lambda - 4)^2 - 4] + 4[-2(\lambda - 4) - 4] \\
&= (\lambda - 4)(\lambda - 6)(\lambda - 2) - 8(\lambda - 2) = \cdots = (\lambda - 2)^2(\lambda - 8),
\end{aligned}$$

συνεπώς οι ιδιοτιμές είναι οι $\lambda_1 = \lambda_2 = 2$, αλγεβρική πολλαπλότητα ίση με 2, και $\lambda_3 = 8$. Στη συνέχεια βρίσκουμε τους αντίστοιχους ιδιόχωρους. Για $\lambda = 2$ επιλύουμε το σύστημα

$$\begin{pmatrix} -2 & -2 & -2 \\ -2 & -2 & -2 \\ -2 & -2 & -2 \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \xi_1 + \xi_2 + \xi_3 = 0,$$

επομένως ως ιδιοδιάνυσμα παίρνουμε το

$$\boldsymbol{x} = \begin{pmatrix} -\xi_2 - \xi_3 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} -\xi_2 \\ \xi_2 \\ 0 \end{pmatrix} + \begin{pmatrix} -\xi_3 \\ 0 \\ \xi_3 \end{pmatrix} = \xi_2 \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} + \xi_3 \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

Τα δύο διανύσματα που συνθέτουν το $\boldsymbol{x}$ είναι γραμμικά ανεξάρτητα, κατά συνέπεια μια βάση για τον ιδιόχωρο που αντιστοιχεί στην ιδιοτιμή $\lambda = 2$ αποτελείται από αυτά τα ιδιοδιανύσματα

$$\boldsymbol{u}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{u}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

Για $\lambda = 8$ επιλύουμε το σύστημα

$$\begin{pmatrix} 4 & -2 & -2 \\ -2 & 4 & -2 \\ -2 & -2 & 4 \end{pmatrix}\begin{pmatrix} \xi_1 \\ \xi_2 \\ \xi_3 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix} \Leftrightarrow \xi_1 = \xi_2 = \xi_3,$$

παρατηρώντας ότι το άθροισμα των στοιχείων κάθε γραμμής του μητρώου είναι ίσο με το μηδέν, επομένως ως ιδιοδιάνυσμα παίρνουμε το

$$\boldsymbol{u}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

το οποίο σύμφωνα με το Θεώρημα 9.6 είναι ορθογώνιο στα $\boldsymbol{u}_1$ και $\boldsymbol{u}_2$. Έτσι θα έχουμε

$$A = \begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3\end{pmatrix} \operatorname{diag}\begin{pmatrix}2 & 2 & 8\end{pmatrix}\begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3\end{pmatrix}^{-1}$$
$$A = \begin{pmatrix}-1 & -1 & 1\\ 1 & 0 & 1\\ 0 & 1 & 1\end{pmatrix}\begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 8\end{pmatrix}\begin{pmatrix}-1 & -1 & 1\\ 1 & 0 & 1\\ 0 & 1 & 1\end{pmatrix}^{-1}.$$

Συνδυάζοντας το αποτέλεσμα του Θεωρήματος 9.5 με εκείνο του Θεωρήματος 9.6 έχουμε το

Θεώρημα 9.7. *Το $n \times n$ μητρώο A είναι διαγωνοποιήσιμο αν και μόνο αν η αλγεβρική πολλαπλότητα κάθε ιδιοτιμής είναι ίση με τη γεωμετρική πολλαπλότητα της ιδιοτιμής. Ειδικά αν η αλγεβρική πολλαπλότητα κάθε ιδιοτιμής είναι ίση με ένα, δηλαδή οι ιδιοτιμές είναι ξένες μεταξύ τους, τότε το μητρώο είναι διαγωνοποιήσιμο.*

Απόδειξη. Αφήνεται σαν άσκηση. □

Παράδειγμα 9.6. Εξετάξτε αν το μητρώο

$$A = \begin{pmatrix}2 & -1\\ 1 & 4\end{pmatrix}$$

είναι διαγωνοποιήσιμο.
Το χαρακτηριστικό πολυώνυμο του A είναι

$$p_A(\lambda) = \begin{vmatrix}\lambda - 2 & 1\\ -1 & \lambda - 4\end{vmatrix} = (\lambda - 2)(\lambda - 4) + 1 = (\lambda - 3)^2,$$

κατά συνέπεια το A έχει μία ιδιοτιμή με αλγεβρική πολλαπλότητα δύο. Ψάχνοντας για ιδιοδιανύσματα του A επιλύουμε το σύστημα

$$(3I - A)\boldsymbol{x} = \boldsymbol{0} \Leftrightarrow \begin{pmatrix}1 & 1\\ -1 & -1\end{pmatrix} = \begin{pmatrix}x_1\\ x_2\end{pmatrix} = \begin{pmatrix}0\\ 0\end{pmatrix} \Leftrightarrow x_1 = -x_2,$$

έτσι ο ιδιόχωρος που αντιστοιχεί στην ιδιοτιμή $\lambda = 1$ είναι

$$\mathcal{E} = \operatorname{span}\left\{\begin{pmatrix}1\\ -1\end{pmatrix}\right\},$$

έχει δηλαδή διάσταση ένα, ισοδύναμα η γεωμετρική πολλαπλότητα της ιδιοτιμής είναι ίση με ένα, διαφορετική από την αλγεβρική πολλαπλότητα. Συνεπώς το A δεν διαγωνοποιείται.

9.4 Ορθογώνια διαγωνοποίηση

Ας υποθέσουμε ότι ένα τετραγωνικό μητρώο A είναι διαγωνοποιήσιμο και το μητρώο που υλοποιεί την διαγωνοποίηση είναι ορθογώνιο ισχύει δηλαδή ότι

$$A = PDP^{-1} = PDP^{\mathrm{T}},$$

όπου το D είναι διαγώνιο. Στόχος μας είναι να προσδιορίσουμε αυτού του τύπου τα μητρώα.

Ορισμός 9.5. Ένα τετραγωνικό μητρώο A θα λέγεται **ορθογώνια διαγωνοποιήσιμο** (orthogonally diagonalizable) εάν υπάρχουν διαγώνιο μητρώο D και ορθογώνιο μητώο Q, ώστε

$$A = QDQ^{\mathrm{T}}. \tag{9.29}$$

Θεώρημα 9.8. *Εάν A είναι ένα ορθογώνια διαγωνοποιήσιμο μητρώο, τότε το A είναι συμμετρικό.*

Απόδειξη. Αν $A = QDQ^{\mathrm{T}}$, με Q ορθοώνιο και D διαγώνιο αναστρέφοντας

$$A^{\mathrm{T}} = (QDQ^{\mathrm{T}})^{\mathrm{T}} = (Q^{\mathrm{T}})^{\mathrm{T}}D^{\mathrm{T}}Q^{\mathrm{T}} = QDQ^{\mathrm{T}} = A.$$

αποδεικνύεται το ζητούμενο. □

Στην παράγραφο αυτή αποδεικνύουμε ότι αν το A είναι πραγματικό τότε ισχύει και το αντίστροφο, δηλαδή ένα συμμετρικό, πραγματικό μητρώο είναι ορθογώνια διαγωνοποιήσιμο.

Ένα πολυώνυμο με πραγματικούς συντελεστές κάλιστα μπορεί να έχει μιγαδικές ρίζες, κατά συνέπεια ένα πραγματικό μητρώο μπορεί να έχει μιγαδικές ιδιοτιμές, επομένως και μιγαδικά ιδιοδιανύσματα. Στην περίπτωση όμως όπου το μητρώο είναι συμμετρικό κάτι τέτοιο δεν συμβαίνει. Πριν αποδείξουμε το σχετικό αποτέλεσμα ας δούμε μια χαρακτηριστική ιδότητα των συμμετρικών μητρώων με πραγματικά στοιχεία. Θυμίζουμε ότι ένας αριθμός a είναι πραγματικός αν και μόνο αν είναι ίσος με τον συζυγή του $\overline{a}$. Αντιμετωπίζοντας μιγαδικές ποσότητες και προκειμένου η ευκλείδια νόρμα να παράγεται από εσωτερικό γινόμενο στον $\mathbb{F}^n$, όπου $\mathbb{F} = \mathbb{R}$, ή $\mathbb{F} = \mathbb{C}$ χρησιμοποιούμε το εσωτερικό γινόμενο

$$\langle \boldsymbol{u}, \boldsymbol{v} \rangle = \boldsymbol{u}^{\mathrm{T}}\overline{\boldsymbol{v}} = \sum_{i=1}^{n} u_i\overline{v_i},$$

βλέπε Ορισμό 5.2 και Άσκηση 5.4. Αν $A = \left(a_{ij}\right)$ είναι ένα συμμετρικό, πραγματικό μητρώο, τότε

$$\overline{A} = \left(\overline{a_{ij}}\right) = \left(a_{ij}\right) = A$$

και

$$\langle A\boldsymbol{u}, \boldsymbol{v} \rangle = (A\boldsymbol{u})^{\mathrm{T}}\overline{\boldsymbol{v}} = \boldsymbol{u}^{\mathrm{T}}A^{\mathrm{T}}\overline{\boldsymbol{v}} = \boldsymbol{u}^{\mathrm{T}}A\overline{\boldsymbol{v}} = \boldsymbol{u}^{\mathrm{T}}\overline{A}\overline{\boldsymbol{v}} = \boldsymbol{u}^{\mathrm{T}}\overline{A\boldsymbol{v}} = \langle \boldsymbol{u}, A\boldsymbol{v} \rangle \tag{9.30}$$

Θεώρημα 9.9. *Εάν το A είναι ένα συμμετρικό, πραγματικό μητρώο, τότε όλες οι ιδιοτιμές του είναι πραγματικοί αριθμοί.*

Απόδειξη. Έστω λ μια ιδιοτιμή του A και έστω $\boldsymbol{x}$ το αντίστοιχο ιδιοδιάνυσμα. Αποδεικνύουμε ότι $\overline{\lambda} = \lambda$. Έτσι υπολογίζουμε, αφενός

$$\langle A\boldsymbol{x}, \boldsymbol{x}\rangle = \langle \lambda\boldsymbol{x}, \boldsymbol{x}\rangle = \lambda\langle \boldsymbol{x}, \boldsymbol{x}\rangle = \lambda\|\boldsymbol{x}\|^2 \tag{9.31}$$

και αφετέρου από την (9.30)

$$\langle A\boldsymbol{x}, \boldsymbol{x}\rangle = \langle \boldsymbol{x}, A\boldsymbol{x}\rangle = \langle \boldsymbol{x}, \lambda\boldsymbol{x}\rangle = \overline{\lambda}\langle \boldsymbol{x}, \boldsymbol{x}\rangle = \overline{\lambda}\|\boldsymbol{x}\|^2. \tag{9.32}$$

Από τις (9.31), (9.32) έπεται ότι $\lambda\|\boldsymbol{x}\|^2 = \overline{\lambda}\|\boldsymbol{x}\|^2$, συνεπώς $\lambda = \overline{\lambda}$, αφού $\|\boldsymbol{x}\| > 0$, που είναι ό,τι θέλαμε να αποδείξουμε. □

Για τα ιδιοδιανύσματα συμμετρικού, πραγματικού μητρώου έχουμε

Θεώρημα 9.10. *Εάν το A είναι ένα συμμετρικό, πραγματικό μητρώο, τότε σε κάθε ιδιοτιμή του αντιστοιχεί ένα πραγματικό ιδιοδιάνυσμα.*

Απόδειξη. Έστω λ μια ιδιοτιμή του A με $\boldsymbol{x}$ ένα αντίστοιχο ιδιοδιάνυσμα. Έστω $\boldsymbol{x} = \boldsymbol{\xi} + i\boldsymbol{\zeta}$ με $\boldsymbol{\xi}$ και $\boldsymbol{\zeta}$ πραγματικά διανύσματα, τότε επειδή $\lambda \in \mathbb{R}$ βρίσκουμε

$$\begin{aligned} A\boldsymbol{x} = \lambda\boldsymbol{x} &\Rightarrow A(\boldsymbol{\xi} + i\boldsymbol{\zeta}) = \lambda(\boldsymbol{\xi} + i\boldsymbol{\zeta}) \\ &\Rightarrow A\boldsymbol{\xi} + iA\boldsymbol{\zeta} = \lambda\boldsymbol{\xi} + i\lambda\boldsymbol{\zeta} \end{aligned}$$

απ΄ όπου εξισώνοντας τα πραγματικά και τα φανταστικά μέρη παίρνουμε

$$A\boldsymbol{\xi} = \lambda\boldsymbol{\xi} \qquad \text{και} \qquad A\boldsymbol{\zeta} = \lambda\boldsymbol{\zeta}.$$

Επειδή $\boldsymbol{x} \neq \boldsymbol{0}$ τουλάχιστον ένα από τα $\boldsymbol{\xi}$ και $\boldsymbol{\zeta}$ είναι διάφορο του μηδενικού διανύσματος, το οποίο τελικά είναι το ζητούμενο πραγματικό ιδιοδιάνυσμα. □

Θεώρημα 9.11. *Εάν το A είναι ένα συμμετρικό, πραγματικό μητρώο, τότε ιδιοδιανύσματα που αντιστοιχούν σε ξένες μεταξύ τους ιδιοτιμές του A είναι ορθογώνια μεταξύ τους.*

Απόδειξη. Εάν $\lambda \neq \mu$ είναι ιδιοτιμές του A με αντίστοιχα ιδιοδιανύσματα $\boldsymbol{x}$ και $\boldsymbol{y}$, τότε, όπως στην απόδειξη του Θεωρήματος 9.9, έχουμε

$$\begin{aligned} \langle A\boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, A\boldsymbol{y}\rangle &\Rightarrow \langle \lambda\boldsymbol{x}, \boldsymbol{y}\rangle = \langle \boldsymbol{x}, \mu\boldsymbol{y}\rangle \\ &\Rightarrow \lambda\langle \boldsymbol{x}, \boldsymbol{y}\rangle = \overline{\mu}\langle \boldsymbol{x}, \boldsymbol{y}\rangle \\ &\Rightarrow (\lambda - \mu)\langle \boldsymbol{x}, \boldsymbol{y}\rangle = 0 \end{aligned}$$

αφού $\overline{\mu} = \mu$. Επειδή $\lambda \neq \mu$ από την τελευταία σχέση έπεται ότι $\langle \boldsymbol{x}, \boldsymbol{y}\rangle = 0$, ισοδύναμα τα $\boldsymbol{x}$, $\boldsymbol{y}$ είναι ορθογώνια μεταξύ τους. □

Παρατήρηση 9.7. Το αποτέλεσμα του Θεωρήματος 9.10 μπορεί ισοδύναμα να διατυπωθεί ως: Εάν $\lambda_i \neq \lambda_j$ είναι ιδιοτιμές ενός συμμετρικού, πραγματικού μητρώου A, τότε οι αντίστοιχοι ιδιόχωροι είναι ορθογώνιοι μεταξύ τους, ισοδύναμα

$$\mathrm{null}(\lambda_i I - A) \perp \mathrm{null}(\lambda_j I - A), \qquad \lambda_i \neq \lambda_j. \tag{9.33}$$

Θεώρημα 9.12. *Έστω A ένα συμμετρικό, πραγματικό μητρώο, και έστω $\boldsymbol{x}$ ένα ιδιοδιάνυσμα του A. Εάν $\boldsymbol{v}$ είναι ένα διάνυσμα ορθογώνιο στο $\boldsymbol{x}$, τότε και το $A\boldsymbol{v}$ είναι ορθογώνιο στο $\boldsymbol{x}$.*

Απόδειξη. Εάν λ είναι η ιδιοτιμή του A που αντιστοιχεί στο ιδιοδιάνυσμα $\boldsymbol{x}$, από την (9.30) παίρνουμε

$$\langle A\boldsymbol{v}, \boldsymbol{x}\rangle = \langle \boldsymbol{v}, A\boldsymbol{x}\rangle = \langle \boldsymbol{v}, \lambda\boldsymbol{x}\rangle = \lambda\langle \boldsymbol{v}, \boldsymbol{x}\rangle = 0$$

γεγονός που αποδεικνύει τον ισχυρισμό μας. □

Θεώρημα 9.13. *Εάν το A είναι ένα συμμετρικό, πραγματικό μητρώο, τότε το A είναι ορθογώνια διαγωνοποιήσιμο.*

Απόδειξη. Έστω ότι το A είναι $n \times n$ και έστω ότι k είναι το πλήθος των διακριτών ιδιοτιμών του A. Διακρίνουμε τις περιπτώσεις.

Περίπτωση 1. $k = n$, τότε από το Θεώρημα 9.11 έπεται ότι υπάρχουν n ορθογώνια μεταξύ τους ιδιοδιανύσματα, άρα γραμμικά ανεξάρτητα, τα οποία απο το Θεώρημα 9.10 είναι πραγματικά. Συνεπώς το μητρώο Q με στήλες τα κανονοικοποιημένα ιδιοδιανύσματα είναι ορθογώνιο και το ζητούμενο έπεται από το Θεώρημα 9.5.

Περίπτωση 2. $k < n$. Δίνουμε την απόδειξη με επαγωγή στη διάσταση n. Εάν το A είναι 1×1, $n = 1$, τότε το αποτέλεσμα ισχύει, αφού $(a) = (1)(a)(1)$. Υποθέτουμε ότι (υπόθεση της επαγωγής) κάθε $(n-1)\times(n-1)$, με $n > 1$, συμμετρικό πραγματικό μητρώο είναι ορθογώνια διαγωνοποιήσιμο. Θα αποδείξουμε ότι το αυτό ισχύει για κάθε $n\times n$ συμμετρικό, πραγματικό μητρώο A. Εάν $\lambda_1, \ldots, \lambda_k$, με $1 \le k < n$ είναι οι διακριτές ιδιοτιμές του A, και $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$ τα αντίστοιχα ορθογώνια μεταξύ τους (Θεώρημα 9.11) πραγματικά (Θεώρημα 9.10) ιδιοδιανύσματα, κανονικοποιημένα, τότε εμπλουτίζοντας προκύπτει μια ορθοκανονική βάση $\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k, \boldsymbol{v}_{k+1}, \ldots \boldsymbol{v}_n\}$ του $\mathbb{R}^n$. Αν $E = \operatorname{span}\{\boldsymbol{x}_1, \ldots, \boldsymbol{x}_k\}$, τότε $E^{\perp} = \operatorname{span}\{\boldsymbol{v}_{k+1}, \ldots \boldsymbol{v}_n\}$. Το μητρώο

$$P = \begin{pmatrix} \boldsymbol{x}_1 & \ldots & \boldsymbol{x}_k & \boldsymbol{v}_{k+1} & \ldots & \boldsymbol{v}_n \end{pmatrix}$$

είναι ορθογώνιο, επομένως το $P^{\mathrm{T}}AP$ είναι συμμετρικό. Επειδή

$$AP = \begin{pmatrix} \lambda_1\boldsymbol{x}_1 & \ldots & \lambda_k\boldsymbol{x}_k & A\boldsymbol{v}_{k+1} & \ldots & A\boldsymbol{v}_n \end{pmatrix}$$

και $A\boldsymbol{v}_j \in E^{\perp}$, $j = m+1, \ldots, n$, ειδικά κανένα από τα $A\boldsymbol{v}_j$ δεν εκφράζεται μέσω των $\boldsymbol{x}_i$, όπως εξασφαλίζει το Θεώρημα 9.12, συμπεραίνουμε ότι

$$P^{\mathrm{T}}AP = \begin{pmatrix} \Lambda & O_{k\times(n-k)} \\ O_{(n-k)\times k} & B \end{pmatrix} = \begin{pmatrix} \lambda_1 & & & \\ & \ddots & & \\ & & \lambda_k & \\ & & & B \end{pmatrix}$$

όπου το B είναι ένα $(n-m)\times(n-m)$ αναγκαστικά συμμετρικό (γιατί;) πραγματικό μητρώο. Κατά συνέπεια από την υπόθεση της επαγωγής υπάρχει ορθογώνιο μητρώο U και διαγώνιο μητρώο D_1, ώστε $B = U^{\mathrm{T}}D_1U$. Αν

$$R = \begin{pmatrix} I_k & O_{k\times(n-k)} \\ O_{(n-k)\times k} & U^{\mathrm{T}} \end{pmatrix}$$

τότε το R είναι ορθογώνιο και

$$R^{\mathrm{T}} = \begin{pmatrix} I_k & O_{k\times(n-k)} \\ O_{(n-k)\times k} & U \end{pmatrix}$$

Ορίζουμε $Q = PR$, τότε το Q είναι ορθογώνιο

$$Q^{\mathrm{T}} = R^{\mathrm{T}}P^{\mathrm{T}} = R^{-1}P^{-1} = (PR)^{-1} = Q^{-1}$$

και

$$\begin{aligned} Q^{\mathrm{T}}AQ &= R^{\mathrm{T}}P^{\mathrm{T}}APR \\ &= R^{\mathrm{T}}\begin{pmatrix} \Lambda & O_{k\times(n-k)} \\ O_{(n-k)\times k} & B \end{pmatrix}R \\ &= \begin{pmatrix} I_k & O_{k\times(n-k)} \\ O_{(n-k)\times k} & U \end{pmatrix}\begin{pmatrix} \Lambda & O_{k\times(n-k)} \\ O_{(n-k)\times k} & B \end{pmatrix}\begin{pmatrix} I_k & O_{k\times(n-k)} \\ O_{(n-k)\times k} & U^{\mathrm{T}} \end{pmatrix} \\ &= \begin{pmatrix} \Lambda & O_{k\times(n-k)} \\ O_{(n-k)\times k} & D_1 \end{pmatrix} \\ &= D \end{aligned}$$

ισοδύναμα $A = QDQ^{\mathrm{T}}$, ή το μητρώο A είναι ορθογώνια διαγωνοποιήσιμο. Κατά συνέπεια το συμπέρασμα είναι αληθές για κάθε φυσικό n. Η απόδειξη είναι πλήρης. □

Παράδειγμα 9.7. Το μητρώο

$$A = \begin{pmatrix} 4 & 2 & 2 \\ 2 & 4 & 2 \\ 2 & 2 & 4 \end{pmatrix}$$

είναι συμμετρικό, επομένως υπάρχει ορθογώνιο μητρώο Q που το διαγωνοποιεί. Στο Παράδειγμα 9.5 δείξαμε ότι οι ιδιοτιμές του A είναι οι $\lambda_1 = \lambda_2 = 2$, και $\lambda_3 = 8$. Είδαμε ότι μια βάση για τον ιδιόχωρο που αντιστοιχεί στην ιδιοτιμή $\lambda = 2$ αποτελείται από τα ιδιοδιανύσματα

$$\boldsymbol{u}_1 = \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{u}_2 = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}$$

ενώ ο ιδιόχωρος που αντιστοιχεί στην ιδιοτιμή $\lambda = 8$ παράγεται από το

$$\boldsymbol{v}_3 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}.$$

Ενώ το $\boldsymbol{u}_3$ είναι ορθογώνιο στα $\boldsymbol{u}_1$ και $\boldsymbol{u}_2$ τα δύο τελευταία δεν είναι ορθογώνια μεταξύ τους. Με τη διαδικασία ορθοκανονικοποίησης βρίσκουμε

$$\boldsymbol{v}_1 = \frac{1}{\|\boldsymbol{u}_1\|}\boldsymbol{u}_1 = \frac{1}{\sqrt{2}}\begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix}$$

και

$$\boldsymbol{u}_2 \cdot \boldsymbol{v}_1 = \frac{1}{\sqrt{2}}, \qquad \boldsymbol{w} = \boldsymbol{u}_2 - (\boldsymbol{u}_2 \cdot \boldsymbol{v}_1)\boldsymbol{v}_1 = \frac{1}{2}\begin{pmatrix} -1 \\ -1 \\ 2 \end{pmatrix}, \qquad \|\boldsymbol{w}\| = \sqrt{\frac{3}{2}}$$

οπότε

$$\boldsymbol{v}_2 = \frac{1}{\|\boldsymbol{w}\|}\boldsymbol{w} = \frac{1}{\sqrt{6}}\begin{pmatrix}-1\\-1\\2\end{pmatrix}.$$

Αν

$$\boldsymbol{v}_3 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}$$

η βάση $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ είναι ορθοκανονική, κατά συνέπεια το ζητούμενο μητρώο είναι το $Q = \begin{pmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \boldsymbol{v}_3\end{pmatrix}$. Έτσι όπως εύκολα επαληθεεύεται

$$A = \begin{pmatrix}-1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3}\\ 1/\sqrt{2} & -1/\sqrt{6} & 1/\sqrt{3}\\ 0 & 2/\sqrt{6} & 1/\sqrt{3}\end{pmatrix}\begin{pmatrix}2 & 0 & 0\\ 0 & 2 & 0\\ 0 & 0 & 8\end{pmatrix}\begin{pmatrix}-1/\sqrt{2} & 1/\sqrt{2} & 0\\ -1/\sqrt{6} & -1/\sqrt{6} & 2/\sqrt{6}\\ 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3}\end{pmatrix}.$$

Συνδυάζοντας τα αποτελέσματα των Θεωρημάτων 9.8 και 9.13 έχουμε το σημαντικό

Θεώρημα 9.14 (το Φασματικό Θεώρημα). *Ένα πραγματικό $n \times n$ μητρώο είναι ορθογώνια διαγωνοποιήσιμο αν και μόνο αν είναι συμμετρικό.*

Παρατήρηση 9.8. Εάν το A είναι ένα $n \times n$ συμμετρικό, πραγματικό μητρώο, τότε υπάρχουν ορθογώνιο μητρώο Q και διαγώνιο μητρώο D ώστε $A = QDQ^{\mathrm{T}}$, ισοδύναμα $AQ = QD$, ή, όπως ήδη επισημάναμε στην Παρατήρηση 9.6

$$A\begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n\end{pmatrix} = \begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n\end{pmatrix}\begin{pmatrix}d_1 & 0 & \cdots & 0\\ 0 & d_2 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & d_n\end{pmatrix}$$

$$\begin{pmatrix}A\boldsymbol{u}_1 & A\boldsymbol{u}_2 & \cdots & A\boldsymbol{u}_n\end{pmatrix} = \begin{pmatrix}d_1\boldsymbol{u}_1 & d_2\boldsymbol{u}_2 & \cdots & d_n\boldsymbol{u}_n\end{pmatrix}$$

κατά συνέπεια

$$A\boldsymbol{u}_j = d_j\boldsymbol{u}_j, \qquad j = 1, 2, \dots, n. \tag{9.34}$$

Συνεπώς κάθε πραγματικό συμμετρικό $n \times n$ μητρώο έχει n ορθογώνια μεταξύ τους ιδιοδιανύσματα, ισοδύναμα παράγει μια ορθοκανονική βάση στον $\mathbb{R}^n$. Παρόμοια από την αρχική σχέση $A = QDQ^{\mathrm{T}}$ παίρνουμε

$$A = \begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n\end{pmatrix}\begin{pmatrix}d_1 & 0 & \cdots & 0\\ 0 & d_2 & \dots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & d_n\end{pmatrix}\begin{pmatrix}\boldsymbol{u}_1^{\mathrm{T}}\\ \boldsymbol{u}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{u}_n^{\mathrm{T}}\end{pmatrix} = \begin{pmatrix}d_1\boldsymbol{u}_1 & d_2\boldsymbol{u}_2 & \cdots & d_n\boldsymbol{u}_n\end{pmatrix}\begin{pmatrix}\boldsymbol{u}_1^{\mathrm{T}}\\ \boldsymbol{u}_2^{\mathrm{T}}\\ \vdots\\ \boldsymbol{u}_n^{\mathrm{T}}\end{pmatrix}$$

ισοδύναμα

$$A = d_1\boldsymbol{u}_1\boldsymbol{u}_1^{\mathrm{T}} + d_2\boldsymbol{u}_2\boldsymbol{u}_2^{\mathrm{T}} + \cdots + d_n\boldsymbol{u}_n\boldsymbol{u}_n^{\mathrm{T}} \tag{9.35}$$

Κάθε υπόχωρος $\mathrm{span}\{\boldsymbol{u}_j\}$ είναι μια ευθεία δια του $\mathbf{0}$, ένας άξονας, και κάθε μητρώο $\boldsymbol{u}_j\boldsymbol{u}_j^{\mathrm{T}}$ είναι το μητρώο προβολής του $\mathbb{R}^n$ επί του υπόχωρου $\mathrm{span}\{\boldsymbol{u}_j\}$. Συνεπώς η (9.35) εκφράζει το μητρώο

A ως γραμμικό συνδυασμό προβολών επί n το πλήθος ορθογώνιων αξόνων στον $\mathbb{R}^n$ με βάρη τις ιδιοτιμές του A. Έτσι αν $\boldsymbol{x} \in \mathbb{R}^n$ βρίσκουμε

$$\begin{aligned} A\boldsymbol{x} &= d_1\boldsymbol{u}_1\boldsymbol{u}_1^{\mathrm{T}}\boldsymbol{x} + d_2\boldsymbol{u}_2\boldsymbol{u}_2^{\mathrm{T}}\boldsymbol{x} + \cdots + d_n\boldsymbol{u}_n\boldsymbol{u}_n^{\mathrm{T}}\boldsymbol{x} \\ &= d_1 \operatorname{proj}_{\boldsymbol{u}_1} \boldsymbol{x} + d_2 \operatorname{proj}_{\boldsymbol{u}_2} \boldsymbol{x} + \cdots + d_n \operatorname{proj}_{\boldsymbol{u}_n} \boldsymbol{x} \end{aligned}$$

μέσω της (7.14) αφού $\boldsymbol{u}_j^{\mathrm{T}}\boldsymbol{u}_j = 1$. Η σχέση (9.35) λέγεται **φασματική διάσπαση** (spectral decomposition) του A και αποκαλύπτει τη δράση του A στον γραμμικό μετασχηματισμό $\boldsymbol{x} \mapsto A\boldsymbol{x}$.

9.5 Όμοια μητρώα

Ορισμός 9.6. Εάν A και B είναι τετραγωνικά μητρώα θα λέμε ότι το A είναι **όμοιο** (similar) με το B εάν υπάρχει αντιστρέψιμο μητρώο M ώστε $A = MBM^{-1}$. Αν το A είναι όμοιο με το B γράφουμε $A \sim B$.

Υπό αυτή την έννοια, της ομοιότητας, μπορούμε να πούμε ότι ένα τετραγωνικό μητρώο είναι διαγωνοποιήσιμο αν είναι όμοιο με ένα διαγώνιο μητρώο.

Παρατήρηση 9.9. Ας δούμε λίγο πιο προσεκτικά τη σχέση της ομοιότητας.

1. Επειδή $A = IAI^{-1}$ κάθε τετραχωνικό μητρώο είναι όμοιο με τον εαυτό του.

2. Παρατηρούμε ότι αν το μητρώο A είναι όμοιο με το B, τότε $A = MBM^{-1}$, επομένως
$$B = M^{-1}AM = NAN^{-1},$$
όπου $N = M^{-1}$, κατά συνέπεια το B είναι όμοιο με το A.

3. Έστω ότι το μητρώο A είναι όμοιο με το B, και έστω ότι το B είναι όμοιο με το C, τότε για αντιστρέψιμα μητρώα M και N ισχύουν οι σχέσεις $A = MBM^{-1}$ και $B = NCN^{-1}$, κατά συνέπεια
$$A = MBM^{-1} = M(NCN^{-1})M^{-1} = (MN)C(MN)^{-1},$$
δηλαδή το A είναι όμοιο με το C. Συνεπώς η σχέση ομοιότητας, "$\sim$" είναι αυτοπαθής, συμμετρική και μεταβατική.

Θεώρημα 9.15. *Εάν τα μητρώα A και B είναι όμοια, τότε έχουν τις ίδιες ιδιοτιμές.*

Απόδειξη. Έστω ότι $A = MBM^{-1}$, τότε για κάθε σταθερά λ έχουμε

$$\begin{aligned} \lambda I - A &= \lambda I - MBM^{-1} \\ &= M\lambda IM^{-1} - MBM^{-1} \\ &= M(\lambda I - B)M^{-1} \end{aligned}$$

κατά συνέπεια

$$\begin{aligned} \det(\lambda I - A) &= \det(M(\lambda I - B)M^{-1}) \\ &= (\det M)\det(\lambda I - B)(\det M)^{-1} \\ &= \det(\lambda I - B) \end{aligned}$$

ισοδύναμα $p_A(\lambda) = p_B(\lambda)$, επομένως τα μητρώα έχουν τις ίδιες ιδιοτιμές. □

9.6 Ασκήσεις

1. Για κάθε ένα από τα μητρώα

(α′) $\begin{pmatrix} 1 & i \\ i & -2 \end{pmatrix}$ (β′) $\begin{pmatrix} 1 & i \\ -i & 1 \end{pmatrix}$ (γ′) $\begin{pmatrix} i & 1 \\ 1 & -i \end{pmatrix}$ (δ′) $\begin{pmatrix} 1 & 2 \\ 2 & -2 \end{pmatrix}$

να βρεθούν οι ιδιοτιμές και τα αντίστοιχα ιδιοδιανύσματα.

2. Θεωρήστε το ορθογώνιο μητρώο

$$Q = \begin{pmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{pmatrix}.$$

(α′) Περιγράψτε τη δράση του Q επί ενός μοναδιαίου διανύσματος $\boldsymbol{u} = \begin{pmatrix} \cos\phi & \sin\phi \end{pmatrix}^{\mathrm{T}}$, περιγράψτε δηλαδή γεωμετρικά το μετασχηματισμό $\boldsymbol{u} \to Q\boldsymbol{u}$.

(β′) Δείξτε ότι το Q έχει πραγματικές ιδιοτιμές και αντίστοιχα ιδιοδιανύσματα στο $\mathbb{R}^2$.

3. Ένα 3×3 μητρώο A έχει ιδιοτιμές $\lambda_1 = 1$, $\lambda_2 = 2$ και $\lambda_3 = 2$ με αντίστοιχα ιδιοδιανύσματα τα

$$\boldsymbol{x}_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \quad \boldsymbol{x}_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \quad \boldsymbol{x}_3 = \begin{pmatrix} -1 \\ 1 \\ 1 \end{pmatrix}.$$

(α′) Είναι το μητρώο A αντιστρέψιμο; Δικαιολογήστε την απάντησή σας.

(β′) Είναι το μητρώο A διαγωνοποιήσιμο; Δικαιολογήστε την απάντησή σας.

(γ′) Να βρεθεί το μητρώο A.

4. Έστω A ένα 4×4 μητρώο με χαρακτηριστικό πολυώνυμο

$$p(\lambda) = (\lambda - 1)(\lambda + 2)(\lambda - i)(\lambda + 2i).$$

Δείξτε ότι το μητρώο είναι διαγωνοποιήσιμο στο σώμα των μιγαδικών αριθμών.

5. Εάν ζ είναι ένας μιγαδικός αριθμός διάφορος του μηδενός, δείξτε ότι το μητρώο

$$A = \begin{pmatrix} 1 & \zeta \\ 0 & 1 \end{pmatrix},$$

δεν είναι διαγωνοποιήσιμο.

Κεφάλαιο 10

Γραμμικοί μετασχηματισμοί

10.1 Παραδείγματα Γραμμικών Μετασχηματισμών

Θυμίζουμε ότι αν $X(\mathbb{K})$ και $Y(\mathbb{K})$ είναι δύο διανυσματικοί χώροι μια συνάρτηση $L : X \to Y$, ή $L : X \to \mathbb{K}$ λέγεται **γραμμικός μετασχηματισμός** [1] εάν

$$L(\boldsymbol{u} + \boldsymbol{v}) = L(\boldsymbol{u}) + L(\boldsymbol{v}), \qquad \text{και} \qquad L(\lambda \boldsymbol{u}) = \lambda L(\boldsymbol{u}),$$

ή, ισοδύναμα,

$$L(\lambda \boldsymbol{u} + \mu \boldsymbol{v}) = \lambda L(\boldsymbol{u}) + \mu L(\boldsymbol{v}), \tag{10.1}$$

για όλα τα διανύσματα $\boldsymbol{u}, \boldsymbol{v} \in X$ και για όλες τις σταθερές $\lambda, \mu \in \mathbb{K}$.

Παράδειγμα 10.1. Θεωρούμε την απεικόνιση $T : \mathbb{R}^3 \to \mathbb{R}^2$ όπου

$$T(\boldsymbol{u}) = T\left(\begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix}\right) = \begin{pmatrix} 2u_1 - u_2 + u_3 \\ u_2 + 3u_3 \end{pmatrix}.$$

Ο T είναι γραμμικός μετασχηματισμός.
Πράγματι αν $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^3$, τότε

$$\begin{aligned} T(\boldsymbol{x} + \boldsymbol{y}) = T\left(\begin{pmatrix} x_1 + y_1 \\ x_2 + y_2 \\ x_3 + y_3 \end{pmatrix}\right) &= \begin{pmatrix} 2(x_1 + y_1) - (x_2 + y_2) + (x_3 + y_3) \\ (x_2 + y_2) + 3(x_3 + y_3) \end{pmatrix} \\ &= \begin{pmatrix} 2x_1 - x_2 + x_3 \\ x_2 + 3x_3 \end{pmatrix} + \begin{pmatrix} 2y_1 - y_2 + y_3 \\ y_2 + 3y_3 \end{pmatrix} \\ &= T(\boldsymbol{x}) + T(\boldsymbol{y}) \end{aligned}$$

και αν $\lambda \in \mathbb{R}$, τότε

$$\begin{aligned} T(\lambda \boldsymbol{x}) = T\left(\begin{pmatrix} \lambda x_1 \\ \lambda x_2 \\ \lambda x_3 \end{pmatrix}\right) &= \begin{pmatrix} 2\lambda x_1 - \lambda x_2 + \lambda x_3 \\ \lambda x_2 + 3\lambda x_3 \end{pmatrix} \\ &= \lambda \begin{pmatrix} 2x_1 - x_2 + x_3 \\ x_2 + 3x_3 \end{pmatrix} \\ &= \lambda T(\boldsymbol{x}) \end{aligned}$$

[1] Ειδικά ένας γραμμικός μετασχηματισμός $L : X(\mathbb{K}) \to \mathbb{K}$ λέγεται **γραμμικό συναρτησιακό** (linear functional).

κατά συνέπεια η (10.1) ικανοποιείται, επομένως ο T είναι γραμμικός μετασχηματισμός.

Παράδειγμα 10.2. Δείχνουμε ότι η απεικόνιση $T : \mathbb{R}^2 \to \mathbb{M}^{2,2}$, όπου

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x & x+y \\ x-y & y \end{pmatrix},$$

είναι γραμμικός μετασχηματισμός.
Πράγματι

$$\begin{aligned} T\left(\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} u \\ v \end{pmatrix}\right) &= T\left(\begin{pmatrix} x+u \\ y+v \end{pmatrix}\right) = \begin{pmatrix} x+u & x+u+y+v \\ x+u-y-v & y+v \end{pmatrix} \\ &= \begin{pmatrix} x & x+y \\ x-y & y \end{pmatrix} + \begin{pmatrix} u & u+v \\ u-v & v \end{pmatrix} \\ &= T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) + T\left(\begin{pmatrix} u \\ v \end{pmatrix}\right) \end{aligned}$$

και αν $\lambda \in \mathbb{R}$, τότε

$$\begin{aligned} T\left(\lambda \begin{pmatrix} x \\ y \end{pmatrix}\right) &= T\left(\begin{pmatrix} \lambda x \\ \lambda y \end{pmatrix}\right) = \begin{pmatrix} \lambda x & \lambda x + \lambda y \\ \lambda x - \lambda y & \lambda y \end{pmatrix} \\ &= \lambda \begin{pmatrix} x & x+y \\ x-y & y \end{pmatrix} \\ &= \lambda T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) \end{aligned}$$

κατά συνέπεια η (10.1) ικανοποιείται, επομένως ο T είναι γραμμικός μετασχηματισμός.

Παράδειγμα 10.3. Εάν A είναι ένα $n\times m$ πραγματικό μητρώο, η απεικόνιση $T : \mathbb{R}^m \to \mathbb{R}^n$ που ορίζεται ως

$$T(\boldsymbol{x}) = A\boldsymbol{x}$$

είναι ένας γραμμικός μετασχηματισμός.
Πράγματι αν $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^m$ και $\lambda, \mu \in \mathbb{R}$ ως άμεση συνέπεια της άλγεβρας μητρώων έχουμε

$$\begin{aligned} T(\lambda\boldsymbol{x} + \mu\boldsymbol{y}) &= A(\lambda\boldsymbol{x} + \mu\boldsymbol{y}) \\ &= \lambda A\boldsymbol{x} + \mu A\boldsymbol{y} \\ &= \lambda T(\boldsymbol{x}) + \mu T(\boldsymbol{y}). \end{aligned}$$

Παράδειγμα 10.4 (Περιστροφή και Ανάκλαση). Ειδικά αν

$$A = \begin{pmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{pmatrix}, \qquad \text{ή} \qquad A = \begin{pmatrix} \cos 2\theta & \sin 2\theta \\ \sin 2\theta & -\cos 2\theta \end{pmatrix}$$

είναι αντίστοιχα το μητρώο περιστροφής κατά γωνία θ, βλέπε Παράδειγμα 5.14, ή το μητρώο ανάκλασης στο επίπεδο ως προς διάνυσμα που σχηματίζει γωνία θ με τον οριζόντιο άξονα (ισοδύναμα με το διάνυσμα $\boldsymbol{e}_1$ της κανονικής βάσης), βλέπε Παράδειγμα 5.15, τότε ο μετασχηματισμός $T(\boldsymbol{x}) = A\boldsymbol{x}$ περιστροφής ή ανάκλασης αντίστοιχα είναι γραμμικός.

Παράδειγμα 10.5. Αν $p(x) = a_0 + a_1x + a_2x^2$ είναι ένα πολυώνυμο βαθμού το πολύ δύο, θεωρούμε την απεικόνιση $S : \mathbb{P}_2 \to \mathbb{R}$ με $S(p) = a_0$. Ο μετασχηματισμός S είναι γραμμικός.
Παρατηρούμε ότι μπορούμε να γράψουμε $S(p) = p(0)$. Τότε αν $p, q \in \mathbb{P}_2$ και $\lambda \in \mathbb{R}$ έχουμε

$$\begin{aligned} S(p+q) &= (p+q)(0) = p(0) + q(0) = S(p) + S(q) \\ S(\lambda p) &= (\lambda p)(0) = \lambda p(0) = \lambda S(p) \end{aligned}$$

που είναι ό,τι θέλουμε να αποδείξουμε.

Παράδειγμα 10.6. Αν p είναι ένα πολυώνυμο ορίζουμε ένα νέο πολυώνυμο P με τη σχέση

$$P(x) = xp(x).$$

Η σχέση αυτή ορίζει ένα μετασχηματισμό $L : \mathbb{P}_n \to \mathbb{P}_{n+1}$ μέσω της $L(p) = P$, ισοδύναμα

$$P(x) = L(p)(x) = xp(x).$$

Έτσι αν $p(x) = a_0 + a_1x + a_2x^2$ τότε

$$L(p)(x) = a_0x + a_1x^2 + a_2x^3.$$

Ο μετασχηματισμός $L : \mathbb{P}_n \to \mathbb{P}_{n+1}$ είναι γραμμικός.
Πράγματι αν $p, q \in \mathbb{P}_n$ και $\lambda \in \mathbb{R}$ έχουμε

$$\begin{aligned} L(p+q)(x) &= x(p+q)(x) = x[p(x) + q(x)] = xp(x) + xq(x) = L(p)(x) + L(q)(x) \\ L(\lambda p)(x) &= x(\lambda p)(x) = \lambda xp(x) = \lambda L(p)(x) \end{aligned}$$

που είναι το ζητούμενο.

► **Άσκηση 10.1.** Αν f είναι μια συνεχής συνάρτηση στο διάστημα $[a, b]$ ορίζουμε

$$J(f) = \int_a^b f(x)\,\mathrm{d}x.$$

Αποδείξτε ότι ο $J : C[a, b] \to \mathbb{R}$ είναι ένας γραμμικός μετασχηματισμός.

► **Άσκηση 10.2.** Αν $X(\mathbb{K})$, όπου $\mathbb{K} = \mathbb{R}$, ή $\mathbb{K} = \mathbb{C}$, είναι ένας διανυσματικός χώρος με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$ και $\boldsymbol{a}$ ένα σταθερό διάνυσμα του X ορίζουμε την συνάρτηση $F : X \to \mathbb{K}$ με

$$F(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{a} \rangle, \qquad \boldsymbol{x} \in X,$$

Αποδείξτε ότι ο F είναι ένας γραμμικός μετασχηματισμός.

10.2 Βασικές έννοιες και χαρακτηριστικές ιδιότητες

Μια πρώτη χαρακτηριστική ιδιότητα των γραμμικών μετασχηματισμών είναι ότι απεικονίζουν το μηδενικό διάνυσμα στο μηδενικό διάνυσμα. Πράγματι αν $L : X \to Y$ είναι ένας γραμμικός μετασχηματισμός, τότε

$$L(\mathbf{0}) = L(\mathbf{0} + \mathbf{0}) = L(\mathbf{0}) + L(\mathbf{0}) \Rightarrow L(\mathbf{0}) = \mathbf{0}.$$

Έτσι ο μετασχηματισμός $M : \mathbb{R}^2 \to \mathbb{R}^2$ με

$$M\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} x + y \\ y + 1 \end{pmatrix}$$

δεν είναι γραμμικός αφού $M(\mathbf{0}) \neq \mathbf{0}$.

Ορισμός 10.1. Αν X και Y είναι διανυσματικοί χώροι και $L : X \to Y$ είναι ένας γραμμικός μετασχηματισμός ορίζουμε

(1) Την **εικόνα** (image) του L, $\operatorname{image} L$, να είναι η εικόνα του X μέσω του L

$$\operatorname{image} L = \{L(\boldsymbol{x}) : \boldsymbol{x} \in X\} = L(X).$$

(2) Τον **πυρήνα** (kernel) του L, $\operatorname{kernel} L$, να είναι το σύνολο των διανυσμάτων του X των οποίων η εικόνα μέσω του L είναι το $\mathbf{0} \in Y$, δηλαδή

$$\operatorname{kernel} L = \{x : L(\boldsymbol{x}) = \mathbf{0}\} = L^{-1}(\{\mathbf{0}\}).$$

Από το σχόλιο στην αρχή της παραγράφου έπεται ότι ο πυρήνας κάθε γραμμικού μετασχηματισμού L είναι μη κενό σύνολο αφού $\mathbf{0} \in \operatorname{kernel} L$.

Θεώρημα 10.1. *Εάν $L : X \to Y$ είναι ένας γραμμικός μετασχηματισμός, τότε η εικόνα* $\operatorname{image} L$ *είναι διανυσματικός υπόχωρος του Y, ενώ ο πυρήνας* $\operatorname{kernel} L$ *είναι διανυσματικός υπόχωρος του X.*

Σημειώνουμε ότι αν ο μετασχηματισμός L δίνεται μέσω ενός μητρώου A, δηλαδή $L(\boldsymbol{x}) = A\boldsymbol{x}$, τότε $\operatorname{image} L = \operatorname{range} A = C(A)$ ο χώρος στηλών του A και $\operatorname{kernel} L = \operatorname{null} A$, κατά συνέπεια τα αποτελέσματα του Θεωρήματος, στην περίπτωση αυτή αλλά όχι μόνο όπως θα δούμε παρακάτω, είναι ήδη γνωστά.

Απόδειξη. Η απόδειξη είναι άμεση συνέπεια της γραμμικότητας και αφήνεται σαν άσκηση. □

Παράδειγμα 10.7. Έστω $L : \mathbb{P}_2 \to \mathbb{P}_3$ ο μετασχηματισμός με $L(p)(x) = xp(x)$, βλέπε Παράδειγμα 10.6, τότε

1. $q \in \operatorname{image} L$ αν και μόνο αν υπάρχει $p \in \mathbb{P}_2$ ώστε $q(x) = xp(x)$, ισοδύναμα

$$q(x) = x(a_0 + a_1 x + a_2 x^2) = a_0 x + a_1 x^2 + a_2 x^3$$

κατά συνέπεια

$$\operatorname{image} L = \{q \in \mathbb{P}_3 : q(0) = 0\}.$$

Για παράδειγμα $x + x^2 \in \operatorname{image} L$ αφού $x + x^2 = x(1 + x)$, ενώ $1 - x^2 \notin \operatorname{image} L$.

2. $q(x) = a_0 + a_1 x + a_2 x^2 \in \operatorname{kernel} L$ αν και μόνο αν $xq(x) = x(a_0 + a_1 x + a_2 x^2) = 0$ για όλες τις τιμές του x, ισοδύναμα $a_0 + a_1 x + a_2 x^2 = 0$ για όλες τις τιμές του x, ισοδύναμα $a_0 = a_1 = a_2 = 0$ (γιατί;), κατά συνέπεια $q = 0$

$$\operatorname{kernel} L = \{0\}.$$

Ορισμός 10.2. Έστω $L : X \to Y$ ένας γραμμικός μετασχηματισμός. Εάν η εικόνα του L έχει πεπερασμένη διάσταση ορίζουμε την **τάξη** του L, $\operatorname{rank} L$ να είναι η διάσταση της εικόνας του L, δηλαδή $\operatorname{rank} L = \dim(\operatorname{image} L)$.

Θεώρημα 10.2. *Εάν $L : X \to Y$ είναι ένας γραμμικός μετασχηματισμός και $\dim X = n$, τότε*

$$\operatorname{rank} L + \dim(\operatorname{kernel} L) = n. \tag{10.2}$$

Απόδειξη. Ο πυρήνας $\operatorname{kernel} L$ είναι υπόχωρος του X, κατά συνέπεια $\dim(\operatorname{kernel} L) \le n$.
(i) Αν $\dim(\operatorname{kernel} L) = n$, τότε $\operatorname{kernel} L = X$ και $\operatorname{image} L = \{\mathbf{0}\}$, οπότε το αποτέλεσμα ισχύει.
(ii) Έστω $1 \le \dim(\operatorname{kernel} L) < n$. Αν $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_k\}$ είναι μια βάση του $\operatorname{kernel} L$, τότε υπάρχει επέκταση $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_k, \boldsymbol{w}_{k+1}, \dots, \boldsymbol{w}_n\}$ η οποία είναι μια βάση για τον X, βλέπε Θεώρημα 4.8. Δείχνουμε ότι η συλλογή $\{L(\boldsymbol{w}_{k+1}), \dots, L(\boldsymbol{w}_n)\}$ είναι μια βάση για τον υπόχωρο $\operatorname{image} L$. Πράγματι αν $\boldsymbol{x} \in X$, τότε

$$\boldsymbol{x} = c_1 \boldsymbol{w}_1 + \cdots + c_k \boldsymbol{w}_k + c_{k+1} \boldsymbol{w}_{k+1} + \cdots + c_n \boldsymbol{w}_n$$

επομένως

$$\begin{aligned} L(\boldsymbol{x}) &= c_1 L(\boldsymbol{w}_1) + \cdots + c_k L(\boldsymbol{w}_k) + c_{k+1} L(\boldsymbol{w}_{k+1}) + \cdots + c_n L(\boldsymbol{w}_n) \\ &= c_{k+1} L(\boldsymbol{w}_{k+1}) + \cdots + c_n L(\boldsymbol{w}_n) \end{aligned}$$

αφού $L(\boldsymbol{w}_j) = \mathbf{0}$, για $j = 1, \dots, k$, κατά συνέπεια

$$\operatorname{image} L = \operatorname{span}\{L(\boldsymbol{w}_{k+1}), \dots, L(\boldsymbol{w}_n)\}.$$

Στη συνέχεια δείχνουμε ότι τα διανύσματα $L(\boldsymbol{w}_{k+1}), \dots, L(\boldsymbol{w}_n)$ είναι γραμμικά ανεξάρτητα. Υποθέτοντας ότι

$$c_{k+1} L(\boldsymbol{w}_{k+1}) + \cdots + c_n L(\boldsymbol{w}_n) = \mathbf{0}$$

έπεται ότι

$$L(c_{k+1} \boldsymbol{w}_{k+1} + \cdots + c_n \boldsymbol{w}_n) = \mathbf{0}$$

ισοδύναμα

$$c_{k+1} \boldsymbol{w}_{k+1} + \cdots + c_n \boldsymbol{w}_n \in \operatorname{kernel} L.$$

Επειδή όμως τα $\boldsymbol{w}_{k+1}, \cdots, \boldsymbol{w}_n$ δεν περιέχονται στον $\ker L$ έπεται ότι

$$c_{k+1} \boldsymbol{w}_{k+1} + \cdots + c_n \boldsymbol{w}_n = 0 \Leftrightarrow c_{k+1} = \cdots = c_n = 0,$$

από τη γραμμική ανεξαρτησία, συνεπώς τα $L(\boldsymbol{w}_{k+1}), \dots, L(\boldsymbol{w}_n)$ είναι γραμμικά ανεξάρτητα, άρα αποτελούν μια βάση για το $\operatorname{image} L$, γεγονός που έχει σαν συνέπεια την (10.2).
(iii) Τέλος, έστω $\dim(\operatorname{kernel} L) = 0$, τότε $\operatorname{kernel} L = \{\mathbf{0}\}$ και αν $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ είναι μια βάση του X, όπως στην περίπτωση (ii) η συλλογή $\{L(\boldsymbol{w}_1), \dots, L(\boldsymbol{w}_n)\}$ είναι μια βάση για το $\operatorname{image} L$, επομένως $\operatorname{rank} L = n$, οπότε η (10.2) ισχύει. □

Παρατήρηση 10.1. Αν ο γραμμικός μετασχηματισμός $L : X \to Y$ είναι ένα-προς-ένα, τότε υπάρχει ο αντίστροφος μετασχηματισμός $L^{-1} : \operatorname{image} L \to X$. Ο αντίστροφος μετασχηματισμός είναι επίσης γραμμικός. Πράγματι αν $\boldsymbol{u}, \boldsymbol{v} \in \operatorname{image} L$ και $\lambda, \mu \in \mathbb{R}$, τότε υπάρχουν μοναδικά διανύσματα $\boldsymbol{x}, \boldsymbol{y} \in X$, με $L(\boldsymbol{x}) = \boldsymbol{u}$ και $L(\boldsymbol{y}) = \boldsymbol{v}$ και

$$\begin{aligned} L^{-1}(\lambda \boldsymbol{u} + \mu \boldsymbol{v}) &= L^{-1}(\lambda L(\boldsymbol{x}) + \mu L(\boldsymbol{y})) \\ &= L^{-1}(L(\lambda \boldsymbol{x} + \mu \boldsymbol{y})) \\ &= \lambda \boldsymbol{x} + \mu \boldsymbol{y} \\ &= \lambda L^{-1}(\boldsymbol{u}) + \mu L^{-1}(\boldsymbol{u}) \end{aligned}$$

που είναι το ζητούμενο.

Θεώρημα 10.3. *Ο γραμμικός μετασχηματισμός $L : X \to Y$ έχει αντίστροφο αν και μόνο αν* $\operatorname{kernel} L = \{\boldsymbol{0}\}$.

Απόδειξη. Αρκεί να δείξουμε ότι ο L είναι ένα-προς-ένα αν και μόνο αν $\operatorname{kernel} L = \{\boldsymbol{0}\}$. Αν ο L είναι ένα-προς-ένα και $\boldsymbol{x} \in \operatorname{kernel} L$, τότε $L(\boldsymbol{x}) = \boldsymbol{0} = L(\boldsymbol{0})$, οπότε $\boldsymbol{x} = \boldsymbol{0}$, συνεπώς $\operatorname{kernel} L = \{\boldsymbol{0}\}$. Αν τώρα $\operatorname{kernel} L = \{\boldsymbol{0}\}$ και για τα διανύσματα $\boldsymbol{x}$ και $\boldsymbol{y}$ ισχύει $L(\boldsymbol{x}) = L(\boldsymbol{y})$, τότε

$$L(\boldsymbol{x}) - L(\boldsymbol{y}) = \boldsymbol{0} \Rightarrow L(\boldsymbol{x} - \boldsymbol{y}) = \boldsymbol{0} \Rightarrow \boldsymbol{x} - \boldsymbol{y} \in \operatorname{kernel} L \Rightarrow \boldsymbol{x} - \boldsymbol{y} = \boldsymbol{0},$$

συνεπώς $\boldsymbol{x} = \boldsymbol{y}$, ισοδύναμα ο L είναι ένα-προς-ένα. □

Θεώρημα 10.4. *Εάν $L_1 : X \to Y$ και $L_2 : Y \to Z$ είναι γραμμικοί μετασχηματισμοί, τότε*

(1) *Η σύνθεση $L_2 \circ L_1 : X \to Z$, η οποία ως συνήθως ορίζεται με τη σχέση $(L_2 \circ L_1)(\boldsymbol{x}) = L_2(L_1(\boldsymbol{x}))$, είναι γραμμικός μετασχηματισμός.*

(2) *Αν οι L_1, L_2 είναι ένα-προς-ένα, τότε $L_2 \circ L_1$ είναι ένα-προς-ένα.*

(3) *Αν οι L_1, L_2 είναι ένα-προς-ένα, τότε $(L_2 \circ L_1)^{-1} = L_1^{-1} \circ L_2^{-1}$.*

Απόδειξη. Αφήνεται σαν άσκηση. □

Ορισμός 10.3. Εάν $L : X \to Y$ είναι ένας γραμμικός μετασχηματισμός ο οποίος είναι ένα-προς-ένα και επί θα λέμε ότι ο L είναι ένας **ισομορφισμός** μεταξύ των χώρων X και Y.

Θυμίζουμε, βλέπε Παρατήρηση 4.9, ότι εάν $X(\mathbb{K})$ είναι ένας διανυσματικός χώρος πεπερασμένης διάστασης και $\mathcal{B}$ είναι μια βάση για τον X, ο μετασχηματισμός $\Phi : X \to \mathbb{K}^n$ που απεικονίζει κάθε διάνυσμα $\boldsymbol{x}$ του X στο $(\boldsymbol{x})_{\mathcal{B}}$, το διάνυσμα συντεταγμένων του $\boldsymbol{x}$ ως προς τη βάση, συγκεκριμένα αν $\mathcal{B} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{n}_n\}$ και $\boldsymbol{x} = c_1\boldsymbol{u}_1 + \cdots + c_n\boldsymbol{u}_n$, τότε

$$\Phi(\boldsymbol{x}) = (\boldsymbol{x})_{\mathcal{B}} = \begin{pmatrix} c_1 \\ \vdots \\ c_n \end{pmatrix},$$

είναι ισομορφισμός. Ο μετασχηματισμός Φ λέγεται **φυσιολογικός** ισομορφισμός.

10.3 Το μητρώο γραμμικού μετασχηματισμού

Παρατήρηση 10.2. Για τον μετασχηματισμό T στο Παράδειγμα 10.1 παρατηρούμε ότι

$$T(\boldsymbol{u}) = \begin{pmatrix} 2u_1 - u_2 + u_3 \\ u_2 + 3u_3 \end{pmatrix} = \begin{pmatrix} 2 & -1 & 1 \\ 0 & 1 & 3 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \\ u_3 \end{pmatrix} = A\boldsymbol{u}$$

όπου A είναι το 2×3 μητρώο στην τελευταία σχέση, και η γραμμικότητα του T απορρέει από τις ιδιότητητες της άλγεβρας των μητρών.

Το αποτέλεσμα αυτό γενικεύεται.

▲ **Πρόταση 10.1.** *Εάν $T : \mathbb{R}^m \to \mathbb{R}^n$ είναι ένας γραμμικός μετασχηματισμός, τότε υπάρχει $A \in \mathbb{R}^{n\times m}$ ώστε $T(\boldsymbol{x}) = A\boldsymbol{x}$ για κάθε $\boldsymbol{x} \in \mathbb{R}^m$.*

Απόδειξη. Αν $\mathcal{B}_m = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \ldots, \boldsymbol{e}_m\}$ είναι η συνήθης, κανονική βάση στον $\mathbb{R}^m$ και αν $\boldsymbol{x} \in \mathbb{R}^m$, τότε

$$\boldsymbol{x} = x_1\boldsymbol{e}_1 + x_2\boldsymbol{e}_2 + \cdots + x_m\boldsymbol{e}_m = (\boldsymbol{x})_{\mathcal{B}_m}$$

επομένως

$$\begin{aligned} T(\boldsymbol{x}) &= x_1T(\boldsymbol{e}_1) + x_2T(\boldsymbol{e}_2) + \cdots + x_mT(\boldsymbol{e}_m) \\ &= \begin{pmatrix} T(\boldsymbol{e}_1) & T(\boldsymbol{e}_2) & \cdots & T(\boldsymbol{e}_m) \end{pmatrix} \boldsymbol{x}, \end{aligned}$$

κατά συνέπεια το

$$A = \begin{pmatrix} T(\boldsymbol{e}_1) & T(\boldsymbol{e}_2) & \cdots & T(\boldsymbol{e}_m) \end{pmatrix}$$

είναι το ζητούμενο μητρώο. □

Το αποτέλεσμα της Πρότασης 10.1 σε συνδυασμό με το Παράδειγμα 10.3 μπορεί να διατυπωθεί ως: *μια απεικόνιση $T : \mathbb{R}^m \to \mathbb{R}^n$ είναι ένας γραμμικός μετασχηματισμός, αν και μόνο αν υπάρχει $n \times m$ πραγματικό μητρώο A ώστε $T(\boldsymbol{x}) = A\boldsymbol{x}$ για κάθε $\boldsymbol{x} \in \mathbb{R}^m$.*

Παράδειγμα 10.8 (Κλιμάκωση στην κατεύθυνση του $\boldsymbol{e}_1$). Για $k \geq 0$ θεωρούμε τον μετασχηματισμό $T_{k\boldsymbol{e}_1} : \mathbb{R}^2 \to \mathbb{R}^2$ με

$$T_{k\boldsymbol{e}_1}(\boldsymbol{u}) = \begin{pmatrix} ku_1 \\ u_2 \end{pmatrix}$$

έτσι έχουμε συρρίκνωση αν $0 \leq k < 1$, ή επιμήκυνση αν $k > 1$ στη κατεύθυνση του $\boldsymbol{e}_1$. Υπολογίζουμε

$$T_{k\boldsymbol{e}_1}(\boldsymbol{e}_1) = \begin{pmatrix} k \\ 0 \end{pmatrix} = k\boldsymbol{e}_1, \qquad T_{k\boldsymbol{e}_1}(\boldsymbol{e}_2) = \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \boldsymbol{e}_2 \tag{10.3}$$

και παρατηρούμε ότι

$$T_{k\boldsymbol{e}_1}(\boldsymbol{u}) = \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = A\boldsymbol{u}$$

γεγονός που επιβεβαιώνει το αποτέλεσμα της Πρότασης 10.1, ότι οι στήλες του μητρώου

$$A = \begin{pmatrix} k & 0 \\ 0 & 1 \end{pmatrix}$$

που υλοποιεί τον μετασχηματισμό $T_{k\boldsymbol{e}_1}$ είναι τα διανύσματα $T_{k\boldsymbol{e}_1}(\boldsymbol{e}_1)$ και $T_{k\boldsymbol{e}_1}(\boldsymbol{e}_2)$.

Ορισμός 10.4. Αν $T : \mathbb{R}^m \to \mathbb{R}^n$ είναι ένας γραμμικός μετασχηματισμός, το $n \times m$ μητρώο A το οποίο υλοποιεί τον μετασχηματισμό με την έννοια ότι $T(\boldsymbol{u}) = A\boldsymbol{u}$ λέγεται **μητρώο αναπαράστασης** του μετασχηματισμού.

Παράδειγμα 10.9 (Κλιμάκωση στην κατεύθυνση του $\boldsymbol{e}_2$). Αυτή υλοποιείται μέσω του μετασχηματισμού $T_{k\boldsymbol{e}_2} : \mathbb{R}^2 \to \mathbb{R}^2$ με

$$T_{k\boldsymbol{e}_2}(\boldsymbol{u}) = \begin{pmatrix} u_1 \\ ku_2 \end{pmatrix}, \qquad k \geq 0$$

έτσι, όπως στο Παράδειγμα 10.8 έχουμε συρρίκνωση αν $0 \leq k < 1$, ή επιμήκυνση αν $k > 1$ στη κατεύθυνση του $\boldsymbol{e}_2$. Αφού

$$T_{k\boldsymbol{e}_2}(\boldsymbol{e}_1) = \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \boldsymbol{e}_1, \qquad T_{k\boldsymbol{e}_2}(\boldsymbol{e}_2) = \begin{pmatrix} 0 \\ k \end{pmatrix} = k\boldsymbol{e}_2 \tag{10.4}$$

το μητρώο αναπαράστασης του $T_{k\boldsymbol{e}_2}$ είναι το

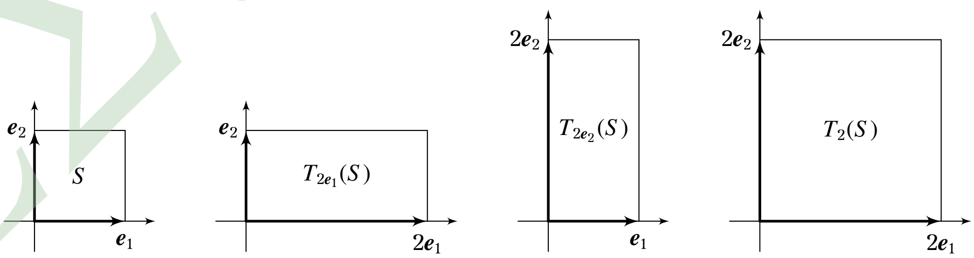
$$A = \begin{pmatrix} 1 & 0 \\ 0 & k \end{pmatrix}.$$

Παράδειγμα 10.10. Ο μετασχηματισμός **συστολής/διαστολής** στο $\mathbb{R}^2$ ορίζεται με τη σχέση

$$T_k(\boldsymbol{u}) = k\boldsymbol{u}, \qquad k \geq 0.$$

Παρατηρούμε ότι $\|T_k(\boldsymbol{u})\| = k\|\boldsymbol{u}\|$, κατά συνέπεια για $0 \leq k < 1$ το αρχικό διάνυσμα συστέλλεται ενώ για $k > 1$ το διάνυσμα διαστέλλεται. Το μητρώο αναπαράστασης του μετασχηματισμού είναι το

$$A = \begin{pmatrix} k & 0 \\ 0 & k \end{pmatrix}.$$



Σχήμα 10.1: Το S και οι εικόνες $T_{2\boldsymbol{e}_1}(S)$, $T_{2\boldsymbol{e}_2}(S)$ και $T_2(S)$ ($k = 2$).

Το Σχήμα 10.1 δείχνει την δράση των μετασχηματισμών $T_{k\boldsymbol{e}_1}$, $T_{k\boldsymbol{e}_2}$ και T_k αντίστοιχα, για $k = 2$, στο μοναδιαίο τετράγωνο S με πλευρές τα διανύσματα $\boldsymbol{e}_1$ και $\boldsymbol{e}_2$.

Παράδειγμα 10.11 (Κλιμάκωση). Ο μετασσχηματισμός $T_{k_1,k_2} = T_{k_1\boldsymbol{e}_1} \circ T_{k_2\boldsymbol{e}_2}$ ως σύνθεση γραμμικών μετασχηματισμών είναι γραμμικός μετασχηματισμός. Έτσι

$$
\begin{aligned}
T_{k_1\boldsymbol{e}_1} \circ T_{k_2\boldsymbol{e}_2}(\boldsymbol{u}) &= T_{k_1\boldsymbol{e}_1}(T_{k_2\boldsymbol{e}_2}(\boldsymbol{u})) \\
&= T_{k_1\boldsymbol{e}_1}\left(\begin{pmatrix} u_1 \\ k_2u_2 \end{pmatrix}\right) \\
&= \begin{pmatrix} k_1u_1 \\ k_2u_2 \end{pmatrix}.
\end{aligned}
$$

Το μητρώο αναπαράστασης του T_{k_1,k_2} είναι το

$$
A = \begin{pmatrix} k_1 & 0 \\ 0 & k_2 \end{pmatrix}.
$$

Αν $k_1 = k_2 = k$ ο μετασχηματισμός κλιμάκωσης είναι ο μετασχηματισμός συστολής/διαστολής.

Παράδειγμα 10.12 (Διατμήσεις). Ένας γραμμικός μετασχηματισμός $T : \mathbb{R}^2 \to \mathbb{R}^2$ με

$$
T(\boldsymbol{u}) = \begin{pmatrix} u_1 + ku_2 \\ u_2 \end{pmatrix}, \qquad k \neq 0
$$

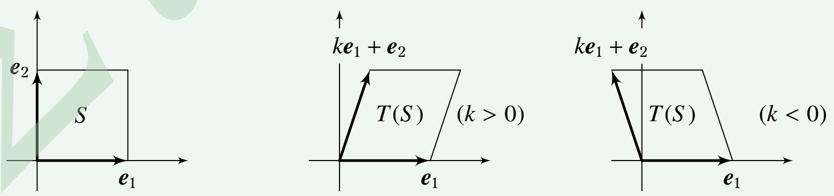
λέγεται **διάτμηση** (shear) στην κατεύθυνση του $\boldsymbol{e}_1$ κατά παράγοντα k. Παρατηρούμε ότι

$$
T(\boldsymbol{e}_1) = \boldsymbol{e}_1, \qquad T(\boldsymbol{e}_2) = k\boldsymbol{e}_1 + \boldsymbol{e}_2.
$$

Το μητρώο αναπαράστασης για τον μετασχηματισμό διάτμησης είναι το

$$
A = \begin{pmatrix} 1 & k \\ 0 & 1 \end{pmatrix}.
$$

Στο Σχήμα 10.2 αποτυπώνονται οι εικόνες του μοναδιαίου τετραγώνου με πλευρές τα διανύσματα $\boldsymbol{e}_1$ και $\boldsymbol{e}_2$ μέσω του μετασχηματισμού διάτμησης στην κατεύθυνση του $\boldsymbol{e}_1$ για $k > 0$ και $k < 0$.



Σχήμα 10.2: Η δράση επί του S της διάτμησης στην κατεύθυνση του $\boldsymbol{e}_1$ για $k = \pm 1/2$.

► **Άσκηση 10.3.** Ο ανάλογος μετασχηματισμός με

$$
T(\boldsymbol{u}) = \begin{pmatrix} u_1 \\ u_2 + ku_1 \end{pmatrix}, \qquad k \neq 0
$$

λέγεται **διάτμηση** στην κατεύθυνση του $\boldsymbol{e}_2$ κατά παράγοντα k.

(α΄) Να βρεθεί το μητρώο αναπαράστασης για τον μετασχηματισμό διάτμησης στην κατεύθυνση του $\boldsymbol{e}_2$.

(β΄) Να περιγραφεί/αποτυπωθεί η δράση του μετασχηματισμού διάτμησης στην κατεύθυνση του $\boldsymbol{e}_2$ επί του μοναδιαίου τετραγώνου S για $k > 0$ και $k < 0$.

Διατυπώνεται τώρα το ερώτημα κατά πόσον ένας γραμμικός μετασχηματισμός $L : X \to Y$ μεταξύ γενικών διανυσματικών χώρων μπορεί να παρασταθεί μέσω κάποιου σχετικού μητρώου A. Είναι σχεδόν προφανές ότι μια τέτοια αναπαράσταση δεν μπορεί να είναι της μορφής $L(\boldsymbol{x}) = A\boldsymbol{x}$ (γιατί;). Ας δούμε πως μπορούμε να αντιμετωπίσουμε αυτό το πρόβλημα με ένα παράδειγμα.

Παράδειγμα 10.13. Θεωρούμε τους διανυσματικούς χώρους των πολυωνύμων $\mathbb{P}_3$ και $\mathbb{P}_2$ και τον μετασχηματισμό παράγωγο $T : \mathbb{P}_3 \to \mathbb{P}_2$ όπου για $p \in \mathbb{P}_3$ είναι

$$T(p) = \frac{\mathrm{d}p}{\mathrm{d}x} = p',$$

ισοδύναμα

$$T(a_0 + a_1x + a_2x^2 + a_3x^3) = a_1 + 2a_2x + 3a_3x^2 \tag{10.5}$$

Συνέπεια των ιδιοτήτων της παραγώγου είναι ότι ο μετασχηματισμός T είναι γραμμικός. Σχετικά με το μητρώο αναπαράστασης του T σκεφτόμαστε, βλέπε Παρατήρηση 4.9, ότι οι χώροι $\mathbb{P}_3$ και $\mathbb{P}_2$ με κάποια έννοια ταυτίζονται με τους $\mathbb{R}^4$ και $\mathbb{R}^3$ αντίστοιχα. Οι μετασχηματισμοί $\Phi_3 : \mathbb{P}_3 \to \mathbb{R}^4$ και $\Phi_2 : \mathbb{P}_2 \to \mathbb{R}^3$ με

$$\Phi_3(a_0 + a_1x + a_2x^2 + a_3x^3) = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} \quad \text{και} \quad \Phi_2(b_0 + b_1x + b_2x^2) = \begin{pmatrix} b_0 \\ b_1 \\ b_2 \end{pmatrix}$$

ισοδύναμα $\Phi_3(p) = (p)_{\mathcal{B}_3}$ και $\Phi_2(p) = (p)_{\mathcal{B}_2}$ όπου $\mathcal{B}_k = \{1, x, \ldots, x^k\}$ είναι η συνήθης βάση για τον $\mathbb{P}_k$, είναι ισομορφισμοί. Τώρα όμως, σύμφωνα με την Πρόταση 10.1, το μητρώο που υλοποιεί τον, σχετικό με τον (10.5), μετασχηματισμό

$$(p)_{\mathcal{B}_3} = \Phi_3(p) = \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} \to \begin{pmatrix} a_1 \\ 2a_2 \\ 3a_3 \end{pmatrix} = \Phi_2(T(p)) = (T(p))_{\mathcal{B}_2}$$

είναι το

$$A = \Big((T(1)_{\mathcal{B}_2} \ \ T(x)_{\mathcal{B}_2} \ \ T(x^2)_{\mathcal{B}_2} \ \ T(x^3)_{\mathcal{B}_2}\Big) = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix}$$

ώστε

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \end{pmatrix} \begin{pmatrix} a_0 \\ a_1 \\ a_2 \\ a_3 \end{pmatrix} = \begin{pmatrix} a_1 \\ 2a_2 \\ 3a_3 \end{pmatrix}.$$

Οι διανυσματικοί χώροι που εμπλέκονται και οι μεταξύ τους μετασχηματισμοί αποτυπώνονται στο διάγραμμα

$$\begin{tikzcd} \mathbb{P}_3 \arrow[r, "T"] \arrow[d, "\Phi_3"] & \mathbb{P}_2 \arrow[d, "\Phi_2"] \\ \mathbb{R}^4 \arrow[r, "A"] & \mathbb{R}^3 \end{tikzcd} \quad \text{ή} \quad \begin{tikzcd} p \arrow[r, "T"] \arrow[d, "\Phi_3"] & T(p) \arrow[d, "\Phi_2"] \\ (p)_{\mathcal{B}_3} \arrow[r, "A"] & (T(p))_{\mathcal{B}_2} \end{tikzcd}$$

Λόγω ισομορφισμών έχουμε ότι

$$T(p) = \Phi_2^{-1}(A\Phi_3(p))$$

και υπό αυτή την έννοια λέμε ότι το μητρώο A παριστάνει τον μετασχηματισμό T.

Ορισμός 10.5. Έστω $T : X \to Y$ ένας γραμμικός μετασχηματισμός και ας υποθέσουμε ότι $\dim X = m$, $\dim Y = n$ και $\mathcal{B}_X = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \dots, \boldsymbol{u}_m\}$ είναι μια βάση για τον X, ενώ $\mathcal{B}_Y = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$ είναι μια βάση για τον Y. Αν $\Phi_X : X \to \mathbb{R}^m$ και $\Phi_Y : Y \to \mathbb{R}^n$ είναι οι ισομορφισμοί

$$\Phi_X(\boldsymbol{x}) = (\boldsymbol{x})_{\mathcal{B}_X} \qquad \text{και} \qquad \Phi_Y(\boldsymbol{y}) = (\boldsymbol{y})_{\mathcal{B}_Y}$$

και αν

$$T(\boldsymbol{u}_j) = a_{1j}\boldsymbol{v}_1 + a_{2j}\boldsymbol{v}_2 + \cdots + a_{nj}\boldsymbol{v}_n, \qquad (T(\boldsymbol{u}_j))_{\mathcal{B}_Y} = \begin{pmatrix} a_{1j} \\ a_{2j} \\ \vdots \\ a_{nj} \end{pmatrix},$$

το $n \times m$ μητρώο

$$A_{T,\mathcal{B}_X,\mathcal{B}_Y} = \Big((T(\boldsymbol{u}_1))_{\mathcal{B}_Y} \;\; (T(\boldsymbol{u}_2))_{\mathcal{B}_Y} \;\; \cdots \;\; (T(\boldsymbol{u}_n))_{\mathcal{B}_Y} \Big), \tag{10.6}$$

ικανοποιεί τη σχέση

$$A_{T,\mathcal{B}_X,\mathcal{B}_Y}(\boldsymbol{x})_{\mathcal{B}_X} = (T(\boldsymbol{x}))_{\mathcal{B}_Y} \tag{10.7}$$

και λέγεται **μητρώο αναπαράστασης** του T ως προς τις βάσεις $\mathcal{B}_X$ και $\mathcal{B}_Y$.

Παρατήρηση 10.3. Από τον ορισμό του μητρώου αναπαράστασης (10.6) βλέπουμε ότι το μητρώο εξαρτάται από τις βάσεις. Πράγματι για τον μετασχηματισμό στο Παράδειγμα 10.1

$$T(\boldsymbol{x}) = T\left(\begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix}\right) = \begin{pmatrix} 2x_1 - x_2 + x_3 \\ x_2 + 3x_3 \end{pmatrix}$$

το μητρώο αναπαράστασης ως προς τις συνήθεις βάσεις των $\mathbb{R}^3$ και $\mathbb{R}^2$ είναι

$$A = \begin{pmatrix} 2 & -1 & 1 \\ 0 & 1 & 3 \end{pmatrix}.$$

Αν πάρουμε ως βάσεις των $\mathbb{R}^3$ και $\mathbb{R}^2$ τις

$$\mathcal{B} = \left\{ \boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 2 \\ 0 \end{pmatrix}, \boldsymbol{u}_3 = \begin{pmatrix} 2 \\ 0 \\ 1 \end{pmatrix} \right\}, \qquad \mathcal{B}' = \left\{ \boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 3 \end{pmatrix} \right\}$$

υπολογίζουμε

$$T(\boldsymbol{u}_1) = \begin{pmatrix} 2 \\ 4 \end{pmatrix} = \boldsymbol{v}_1 + \boldsymbol{v}_2, \qquad T(\boldsymbol{u}_2) = \begin{pmatrix} 0 \\ 2 \end{pmatrix} = -\boldsymbol{v}_1 + \boldsymbol{v}_2, \qquad T(\boldsymbol{u}_3) = \begin{pmatrix} 5 \\ 3 \end{pmatrix} = 6\boldsymbol{v}_1 - \boldsymbol{v}_2$$

κατά συνέπεια

$$(T(\boldsymbol{u}_1))_{\mathcal{B}'} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad (T(\boldsymbol{u}_2))_{\mathcal{B}'} = \begin{pmatrix} -1 \\ 1 \end{pmatrix}, \qquad (T(\boldsymbol{u}_3))_{\mathcal{B}'} = \begin{pmatrix} 6 \\ -1 \end{pmatrix}$$

και το μητρώο αναπαράστασης του T ως προς τις βάσεις $\mathcal{B}$ και $\mathcal{B}'$ είναι το

$$A_{T,\mathcal{B},\mathcal{B}'} = \begin{pmatrix} 1 & -1 & 6 \\ 1 & 1 & -1 \end{pmatrix}$$

το οποίο είναι διαφορετικό από το A. Παρατηρήστε ότι ενώ $T(x) = A\boldsymbol{x}$ στην αναπαράσταση ως προς τις βάσεις $\mathcal{B}$ και $\mathcal{B}'$ είναι

$$T(\boldsymbol{x}) = \Phi_{\mathcal{B}'}^{-1}(A_{T,\mathcal{B},\mathcal{B}'}\Phi_{\mathcal{B}}(\boldsymbol{x})) \tag{10.8}$$

όπου $\Phi_{\mathcal{B}}(\boldsymbol{x})$ είναι το διάνυσμα συντεταγμένων του $\boldsymbol{x}$ ως προς τη βάση $\mathcal{B}$ και $\Phi_{\mathcal{B}'}(\boldsymbol{y})$ είναι το διάνυσμα συντεταγμένων του $\boldsymbol{y}$ ως προς τη βάση $\mathcal{B}'$.

► **Άσκηση 10.4.** Βρείτε τους ισομορφισμούς $\Phi_{\mathcal{B}}$ και $\Phi_{\mathcal{B}'}$ και επαληθεύστε την (10.8).

10.3.1 Το μητρώο της σύνθεσης γραμμικών μετασχηματισμών

Ο Arthur Cayley, μελετώντας τους μετασχηματισμούς της γραμμικής κίνησης, Παράδειγμα 2.15 διείδε την αλγεβρική υπόσταση των μητρώων και όρισε τον πολλαπλασιασμό μεταξύ αυτών των αντικειμένων ώστε να αντιστοιχεί στη σύνθεση των μετασχηματισμών. Κατά συνέπεια το αποτέλεσμα που ακολουθεί είναι αναμενόμενο.

Θεώρημα 10.5. *Εάν X, Y και Z είναι διανυσματικοί χώροι πεπερασμένης διάστασης, $L_1 : X \to Y$ και $L_2 : Y \to Z$ είναι γραμμικοί μετασχηματισμοί με αντίστοιχα μητρώα αναπαράστασης A_{L_1} και A_{L_2}, τότε για το μητρώο αναπαράστασης της σύνθεσης $L_2 \circ L_1 : X \to Z$ ισχύει*

$$A_{L_2 \circ L_1} = A_{L_2} A_{L_1}.$$

Απόδειξη. Αν $\mathcal{B}_X = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_k\}$, $\mathcal{B}_Y = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_m\}$ και $\mathcal{B}_Z = \{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$ είναι βάσεις των X, Y και Z, αντίστοιχα, τότε

$$A_{L_2 \circ L_1} = \Big((L_2 \circ L_1(\boldsymbol{u}_1))_{\mathcal{B}_Z} \;\; (L_2 \circ L_1(\boldsymbol{u}_2))_{\mathcal{B}_Z} \;\; \cdots \;\; (L_2 \circ L_1(\boldsymbol{u}_k))_{\mathcal{B}_Z} \Big).$$

Αλλά για κάθε $j = 1, 2, \ldots, k$ από την (10.7) έχουμε

$$(L_2 \circ L_1(\boldsymbol{u}_j))_{\mathcal{B}_Z} = (L_2(L_1(\boldsymbol{u}_j)))_{\mathcal{B}_Z} = A_{L_2}(L_1(\boldsymbol{u}_j))_{\mathcal{B}_Y} = A_{L_2}A_{L_1}(\boldsymbol{u}_j)_{\mathcal{B}_X}$$

από το οποίο έπεται το ζητούμενο. □

10.3.2 Το μητρώο αλλαγής βάσης

Έστω X ένας διανυσματικός χώρος πεπερασμένης διάστασης, έστω n. Αν $\mathcal{U} = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n\}$ και $\mathcal{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ είναι βάσεις του X, τότε για κάθε $k = 1, 2, \ldots, n$ υπάρχουν σταθερές $a_{1k}, a_{2k}, \ldots, a_{nk}$ ώστε

$$\boldsymbol{u}_k = a_{1k}\boldsymbol{v}_1 + a_{2k}\boldsymbol{v}_2 + \cdots + a_{2k}\boldsymbol{v}_n,$$

κατά συνέπεια

$$\begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n\end{pmatrix} = \begin{pmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n\end{pmatrix}\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \cdots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}$$

ή συνοπτικά $U = VA$, με την προφανή αντιστοιχία. Το μητρώο A λέγεται **μητρώο μετάβασης** από τη βάση $\mathcal{U}$ στην $\mathcal{V}$ και συμβολίζεται ως $M_{\mathcal{U}\to\mathcal{V}}$. Παρατηρούμε ότι αν $\boldsymbol{x} \in X$ και

$$\boldsymbol{x} = x_1\boldsymbol{u}_1 + x_2\boldsymbol{u}_2 + \cdots + x_n\boldsymbol{u}_n$$

γράφοντας ισοδύναμα

$$\begin{aligned}\boldsymbol{x} &= \begin{pmatrix}\boldsymbol{u}_1 & \boldsymbol{u}_2 & \cdots & \boldsymbol{u}_n\end{pmatrix}(\boldsymbol{x})_{\mathcal{U}} \\ &= \begin{pmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2 & \cdots & \boldsymbol{v}_n\end{pmatrix}A(\boldsymbol{x})_{\mathcal{U}}\end{aligned}$$

θα έχουμε

$$(\boldsymbol{x})_{\mathcal{V}} = M_{\mathcal{U}\to\mathcal{V}}(\boldsymbol{x})_{\mathcal{U}} \tag{10.9}$$

γεγονός που δικαιολογεί τον ορισμό του σχετικού μητρώου ως μητρώο μετάβασης από τη βάση $\mathcal{U}$ στη $\mathcal{V}$. Σημειώνουμε ότι το μητρώο $M_{\mathcal{U}\to\mathcal{V}}$ είναι αντιστρέψιμο αφού τα διανύσματα $\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots, \boldsymbol{u}_n$ είναι γραμμικά ανεξάρτητα, και μάλιστα από την (10.9) έπεται ότι

$$M^{-1}_{\mathcal{U}\to\mathcal{V}} = M_{\mathcal{V}\to\mathcal{U}}. \tag{10.10}$$

Παράδειγμα 10.14. Τα διανύσματα

$$\boldsymbol{v}_1 = \begin{pmatrix}2\\-1\end{pmatrix}, \qquad \boldsymbol{v}_2 = \begin{pmatrix}-1\\1\end{pmatrix},$$

είναι γραμμικά ανεξάρτητα, κατά συνέπεια αποτελούν μια βάση $\mathcal{V}$ για τον $\mathbb{R}^2$. Να βρεθεί το μητρώο μετάβασης από τη συνήθη βάση $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ στη $\mathcal{V}$.
Παρατηρούμε ότι

$$\boldsymbol{v}_1 = 2\boldsymbol{e}_1 - \boldsymbol{e}_2, \qquad \text{και} \qquad \boldsymbol{v}_2 = -\boldsymbol{e}_1 + \boldsymbol{e}_2,$$

συνεπώς

$$\begin{pmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2\end{pmatrix} = \begin{pmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2\end{pmatrix}\begin{pmatrix}2 & -1\\ -1 & 1\end{pmatrix}$$

απ' όπου πολλαπλασιάζοντας με το αντίστροφο του μητρώου των συντελεστών βρίσκουμε

$$\begin{pmatrix}\boldsymbol{e}_1 & \boldsymbol{e}_2\end{pmatrix} = \begin{pmatrix}\boldsymbol{v}_1 & \boldsymbol{v}_2\end{pmatrix}\begin{pmatrix}1 & 1\\ 1 & 2\end{pmatrix}$$

και το ζητούμενο μητρώο μετάβασης είναι

$$M_{\mathcal{B}\to\mathcal{V}} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}.$$

Παρατήρηση 10.4. Από το Παράδειγμα 10.14 εξάγονται κάποια συμπεράσματα τα οποία ισχύουν και στη γενική περίπτωση, στο $\mathbb{R}^n$ και για γενικές βάσεις.

(1) Παρατηρούμε αφενός ότι

$$M_{\mathcal{B}\to\mathcal{V}} = \begin{pmatrix} (\boldsymbol{e}_1)_{\mathcal{V}} & (\boldsymbol{e}_2)_{\mathcal{V}} \end{pmatrix}. \tag{10.11}$$

(2) Αφετέρου αν

$$\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

τότε

$$\begin{aligned} \boldsymbol{x} &= \begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{pmatrix}\begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{v}_1 & \boldsymbol{v}_2 \end{pmatrix}\begin{pmatrix} x_1 + x_2 \\ x_1 + 2x_2 \end{pmatrix} \end{aligned}$$

κατά συνέπεια

$$(\boldsymbol{x})_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \qquad \text{και} \qquad (\boldsymbol{x})_{\mathcal{V}} = \begin{pmatrix} x_1 + x_2 \\ x_1 + 2x_2 \end{pmatrix}.$$

Παρατήρηση 10.5. Αν $\mathcal{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2\}$ είναι η συνήθης βάση στο $\mathbb{R}^2$ και $\mathcal{V} = \{\boldsymbol{v}_1, \boldsymbol{v}_2\}$, όπου

$$\boldsymbol{v}_1 = \begin{pmatrix} 2 \\ -1 \end{pmatrix}, \qquad \boldsymbol{v}_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

όπως στο Παράδειγμα 10.14, και $T : \mathbb{R}^2 \to \mathbb{R}^2$ είναι ο ταυτοτικός μετασχηματισμός, $T(\boldsymbol{x}) = \boldsymbol{x}$ θέλουμε να βρούμε το μητρώο αναπαράστασης $A_{T,\mathcal{B},\mathcal{V}}$ του μετασχηματισμού ως προς τις βάσεις $\mathcal{B}$ και $\mathcal{V}$.

Το ζητούμενο μητρώο, αφού

$$\boldsymbol{e}_1 = \boldsymbol{v}_1 + \boldsymbol{v}_2, \qquad \text{και} \qquad \boldsymbol{e}_2 = \boldsymbol{v}_1 + 2\boldsymbol{v}_2,$$

είναι το

$$\begin{pmatrix} (T(\boldsymbol{e}_1))_{\mathcal{V}} & (T(\boldsymbol{e}_2))_{\mathcal{V}} \end{pmatrix} = \begin{pmatrix} (\boldsymbol{e}_1)_{\mathcal{V}} & (\boldsymbol{e}_2)_{\mathcal{V}} \end{pmatrix} = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix},$$

κατά συνέπεια το μητρώο αναπαράστασης του ταυτοτικού μετασχηματισμού ως προς τις βάσεις $\mathcal{B}$ και $\mathcal{V}$ είναι το μητρώο μετάβασης από τη $\mathcal{B}$ στη $\mathcal{V}$

$$A_{T,\mathcal{B},\mathcal{V}} = M_{\mathcal{B}\to\mathcal{V}}.$$

Παράδειγμα 10.15. Θωρούμε το γραμμικό μετασχηματισμό $T : \mathbb{R}^3 \to \mathbb{R}^3$ με τύπο

$$T\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} x+y \\ y-z \\ x+y+z \end{pmatrix}$$

και τη βάση $\mathscr{B}' = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ του $\mathbb{R}^3$, όπου

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \qquad \boldsymbol{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 1 \end{pmatrix}.$$

1. Υπολογίζουμε το μητρώο μετάβασης $M_{\mathscr{B}\to\mathscr{B}'}$, όπου $\mathscr{B} = \{\boldsymbol{e}_1, \boldsymbol{e}_2, \boldsymbol{e}_3\}$ είναι η συνήθης βάση του $\mathbb{R}^3$. Επειδή

$$\boldsymbol{u}_1 = \boldsymbol{e}_1 + \boldsymbol{e}_2, \qquad \boldsymbol{u}_2 = \boldsymbol{e}_1 + \boldsymbol{e}_3, \qquad \boldsymbol{u}_3 = \boldsymbol{e}_2 + \boldsymbol{e}_3$$

έπεται ότι

$$M_{\mathscr{B}'\to\mathscr{B}} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}$$

κατά συνέπεια

$$M_{\mathscr{B}\to\mathscr{B}'} = \begin{pmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 \\ -1/2 & 1/2 & 1/2 \end{pmatrix}.$$

Διαφορετικά με πράξεις βρίσκουμε

$$\begin{pmatrix} \boldsymbol{e}_1 & \boldsymbol{e}_2 & \boldsymbol{e}_3 \end{pmatrix} = \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{pmatrix} \begin{pmatrix} 1/2 & 1/2 & -1/2 \\ 1/2 & -1/2 & 1/2 \\ -1/2 & 1/2 & 1/2 \end{pmatrix}$$

2. Αν

$$\boldsymbol{v} = \begin{pmatrix} x \\ y \\ z \end{pmatrix} = p\boldsymbol{u}_1 + q\boldsymbol{u}_2 + r\boldsymbol{u}_3 = \begin{pmatrix} \boldsymbol{u}_1 & \boldsymbol{u}_2 & \boldsymbol{u}_3 \end{pmatrix} \begin{pmatrix} p \\ q \\ r \end{pmatrix} \tag{10.12}$$

όπου

$$\begin{pmatrix} p \\ q \\ r \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix}, \tag{10.13}$$

τότε

$$T\boldsymbol{v} = pT\boldsymbol{u}_1 + qT\boldsymbol{u}_2 + rT\boldsymbol{u}_3 = \begin{pmatrix} T\boldsymbol{u}_1 & T\boldsymbol{u}_2 & T\boldsymbol{u}_3 \end{pmatrix} \begin{pmatrix} p \\ q \\ r \end{pmatrix}. \tag{10.14}$$

3. Υπολογίζοντας

$$T\boldsymbol{u}_1 = \begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix}, \qquad T\boldsymbol{u}_2 = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}, \qquad T\boldsymbol{u}_3 = \begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix}$$

βρίσκουμε αφενός για το μητρώο αναπαράστασης

$$A_{T,\mathcal{B}',\mathcal{B}} = \begin{pmatrix} 2 & 1 & 1 \\ 1 & -1 & 0 \\ 2 & 2 & 2 \end{pmatrix}$$

και αφετέρου

$$(T\boldsymbol{u}_1)_{\mathcal{B}'} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 2 \\ 1 \\ 2 \end{pmatrix} = \begin{pmatrix} 1/2 \\ 3/2 \\ 1/2 \end{pmatrix}$$

$$(T\boldsymbol{u}_2)_{\mathcal{B}'} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} = \begin{pmatrix} -1 \\ 2 \\ 0 \end{pmatrix}$$

$$(T\boldsymbol{u}_3)_{\mathcal{B}'} = \frac{1}{2}\begin{pmatrix} 1 & 1 & -1 \\ 1 & -1 & 1 \\ -1 & 1 & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 0 \\ 2 \end{pmatrix} = \begin{pmatrix} -1/2 \\ 3/2 \\ 1/2 \end{pmatrix}$$

κατά συνέπεια

$$A_{T,\mathcal{B}',\mathcal{B}'} = \begin{pmatrix} 1/2 & -1 & -1/2 \\ 3/2 & 2 & 3/2 \\ 1/2 & 0 & 1/2 \end{pmatrix}$$

Για το $A_{T,\mathcal{B},\mathcal{B}'}$ υπολογίζουμε

$$T\boldsymbol{e}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \boldsymbol{u}_2, \qquad T\boldsymbol{e}_2 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{2}\boldsymbol{u}_1 + \frac{1}{2}\boldsymbol{u}_2 + \frac{1}{2}\boldsymbol{u}_2, \qquad T\boldsymbol{e}_3 = \begin{pmatrix} 0 \\ -1 \\ 1 \end{pmatrix} = -\boldsymbol{u}_1 + \boldsymbol{u}_2$$

συνεπώς

$$A_{T,\mathcal{B},\mathcal{B}'} = \begin{pmatrix} 0 & 1/2 & -1 \\ 1 & 1/2 & 1 \\ 0 & 1/2 & 0 \end{pmatrix}.$$

Παράδειγμα 10.16. Τα σύνολα διανυσμάτων του $\mathbb{R}^3$

$$\mathcal{U} = \left\{\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \boldsymbol{u}_3 = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}\right\} \quad \text{και} \quad \mathcal{V} = \left\{\boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}, \boldsymbol{v}_3 = \begin{pmatrix} 0 \\ 1 \\ -1 \end{pmatrix}\right\}$$

αποτελούν βάσεις για τον $\mathbb{R}^3$. Αν $T : \mathbb{R}^3 \to \mathbb{R}^3$ είναι ένας γραμμικός μεσχηματισμός ώστε

$$T(\boldsymbol{u}_1) = \boldsymbol{v}_1 + 2\boldsymbol{v}_2, \qquad T(\boldsymbol{u}_2) = \boldsymbol{v}_2 + \boldsymbol{v}_3, \qquad T(\boldsymbol{u}_3) = \boldsymbol{v}_1 - \boldsymbol{v}_3, \tag{10.15}$$

να βρεθεί ο τύπος του μετασχηματισμού.

Από την Πρόταση 10.1 έπεται ότι ο τύπος του μετασχηματισμού προκύπτει από τη σχέση $T(\boldsymbol{x}) = A\boldsymbol{x}$, όπου A είναι το μητρώο αναπαράστασης του T ως προς τη συνήθη βάση του $\mathbb{R}^3$. Αν $M_{\mathcal{U}}$ και $M_{\mathcal{V}}$ είναι αντίστοιχα τα μητρώα μετάβασης από τις $\mathcal{U}$ και $\mathcal{V}$

στη συνήθη βάση $\mathcal{B}$, τότε δηλώντας την εξάρτηση του χώρου από τη βάση έχουμε το διάγραμμα

$$\begin{array}{ccc} \mathbb{R}^3_{\mathcal{U}} & \xrightarrow{T'} & \mathbb{R}^3_{\mathcal{V}} \\ \big\downarrow L_{\mathcal{U}} & & \big\downarrow L_{\mathcal{V}} \\ \mathbb{R}^3_{\mathcal{B}} & \xrightarrow{T} & \mathbb{R}^3_{\mathcal{B}} \end{array}$$

όπου $L_{\mathcal{U}}$ και $L_{\mathcal{V}}$ είναι οι γραμμικοί μετασχηματισμοί με μητρώα $M_{\mathcal{U}}$ και $M_{\mathcal{V}}$ αντίστοιχα και T' ο μετασχηματισμός όπως δίνεται στην (10.15) με μητρώο

$$A' = \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix}.$$

Έτσι σύμφωνα με το διάγραμμα θα έχουμε

$$T(\boldsymbol{x}) = (L_{\mathcal{V}} \circ T' \circ L_{\mathcal{U}}^{-1})(\boldsymbol{x}),$$

ή ισοδύναμα μέσω μητρώων

$$T(\boldsymbol{x}) = A\boldsymbol{x} = (M_{\mathcal{V}} A' M_{\mathcal{U}}^{-1})\boldsymbol{x}.$$

Επειδή

$$\boldsymbol{v}_1 = \boldsymbol{e}_1 - \boldsymbol{e}_3, \qquad \boldsymbol{v}_2 = \boldsymbol{e}_1 + \boldsymbol{e}_3, \qquad \boldsymbol{v}_3 = \boldsymbol{e}_2 - \boldsymbol{e}_3$$
$$\boldsymbol{u}_1 = \boldsymbol{e}_1 + \boldsymbol{e}_3, \quad \boldsymbol{u}_2 = \boldsymbol{e}_1 + \boldsymbol{e}_2, \quad \boldsymbol{u}_3 = \boldsymbol{e}_2 \;\Rightarrow\; \boldsymbol{e}_1 = \boldsymbol{u}_2 - \boldsymbol{u}_3, \quad \boldsymbol{e}_2 = \boldsymbol{u}_3, \quad \boldsymbol{e}_3 = \boldsymbol{u}_1 - \boldsymbol{u}_2 + \boldsymbol{u}_3,$$

έπεται ότι για $\boldsymbol{x} = \begin{pmatrix} x & y & z \end{pmatrix}^{\mathrm{T}}$ είναι

$$\begin{aligned} T(\boldsymbol{x}) &= \begin{pmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 1 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 2 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} \\ &= \begin{pmatrix} 3 & 1 & 1 \\ 0 & 1 & -1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} z \\ x - z \\ -x + y + z \end{pmatrix} \\ &= \begin{pmatrix} y + 3z \\ 2x - y - 2z \\ z \end{pmatrix}. \end{aligned}$$

10.4 Ασκήσεις

1. Τα διανύσματα

$$\boldsymbol{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \boldsymbol{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \boldsymbol{u}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}$$

αποτελούν μια βάση για το $\mathbb{R}^3$. Αν για ένα γραμμικό μετασχηματισμό $T : \mathbb{R}^3 \to \mathbb{R}^2$ ισχύει

$$T(\boldsymbol{u}_1) = \begin{pmatrix} 2 \\ 3 \end{pmatrix}, \qquad T(\boldsymbol{u}_2) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad T(\boldsymbol{u}_3) = \begin{pmatrix} 3 \\ 2 \end{pmatrix},$$

να βρεθεί ο τύπος του μετασχηματισμού.

2. Έστω $T : \mathbb{R}^3 \to \mathbb{R}^3$ να είναι η ορθογώνια προβολή του $\mathbb{R}^3$ επί του επιπέδου W με εξίσωση $x + y + z = 0$. Να βρεθεί ο τύπος του T.

3. **Θεώρημα αναπαράστασης.** Έστω $T : \mathbb{R}^n \to \mathbb{R}$ ένας γραμμικός μετασχηματισμός. Δείξτε ότι υπάρχει διάνυσμα $\boldsymbol{\xi} \in \mathbb{R}^n$ ώστε

$$T(\boldsymbol{x}) = \langle \boldsymbol{x}, \boldsymbol{\xi} \rangle$$

για όλα τα $\boldsymbol{x} \in \mathbb{R}^n$, όπου $\langle \cdot, \cdot \rangle$ είναι το σύνηθες εσωτερικό γινόμενο.
Υπόδειξη. Αν $\mathcal{B} = \{\boldsymbol{u}_1, \ldots, \boldsymbol{u}_n\}$ είναι μια βάση στο $\mathbb{R}^n$ και $\boldsymbol{x} \in \mathbb{R}^n$, τότε $\boldsymbol{x} = x_1\boldsymbol{u}_1 + \cdots + x_n\boldsymbol{u}_n$.

4. Έστω $X = C[0,1]$ ο χώρος των συνεχών συναρτήσεων στο διάστημα $[0,1]$. Ορίζουμε τον μετασχηματισμό $I : X \to X$, όπου $I(f) = F$ με

$$F(x) = \int_0^x f(t)\,\mathrm{d}t, \qquad 0 \le x \le 1.$$

(α΄) Δείξτε ότι ο I είναι γραμμικός μετασχηματισμός.

(β΄) Περιγράψτε την εικόνα $\operatorname{image} I$, του I.

(γ΄) Δείξτε ότι ο I είναι ένα-προς-ένα. **Υπόδειξη:** $I(f)(0) = 0$ για κάθε $f \in X$.

(δ΄) Να βρεθεί ο αντίστροφος μετασχηματισμός I^{-1}.

5. Έστω $T : \mathbb{R}^4 \to \mathbb{R}^3$ με

$$T(\boldsymbol{x}) = \begin{pmatrix} 4 & 1 & -2 & -3 \\ 2 & 1 & 1 & -4 \\ 6 & 0 & -9 & 9 \end{pmatrix} \boldsymbol{x}.$$

(α΄) Εξετάστε αν τα διανύσματα

$$\begin{pmatrix} 1 \\ 3 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 0 \\ 6 \end{pmatrix}, \quad \begin{pmatrix} 2 \\ 4 \\ 1 \end{pmatrix}$$

περιέχονται στην εικόνα $\operatorname{image} T$.

(β΄) Εξετάστε αν τα διανύσματα

$$\begin{pmatrix} 0 \\ -4 \\ 1 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 3 \\ -8 \\ 2 \\ 0 \end{pmatrix}, \quad \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$$

περιέχονται στον πυρήνα $\operatorname{kernel} T$.

6. Έστω $T : \mathbb{R}^2 \to \mathbb{R}^3$ με

$$T\left(\begin{pmatrix} x \\ y \end{pmatrix}\right) = \begin{pmatrix} y \\ -5x + 13y \\ -7x + 16y \end{pmatrix}.$$

Να βρεθεί το μητρώο του T ως προς τις βάσεις

$$\mathcal{B} = \left\{ \boldsymbol{u}_1 = \begin{pmatrix} 3 \\ 1 \end{pmatrix}, \boldsymbol{u}_2 = \begin{pmatrix} 5 \\ 2 \end{pmatrix} \right\} \qquad \text{και} \qquad \mathcal{B}' = \left\{ \boldsymbol{v}_1 = \begin{pmatrix} 1 \\ 0 \\ -1 \end{pmatrix}, \boldsymbol{v}_2 = \begin{pmatrix} -1 \\ 2 \\ 2 \end{pmatrix}, \boldsymbol{v}_3 = \begin{pmatrix} 0 \\ 1 \\ 2 \end{pmatrix} \right\}.$$

7. Έστω $\boldsymbol{u} = \begin{pmatrix} a & b & c \end{pmatrix}^{\mathrm{T}}$ ένα μοναδιαίο διάνυσμα στο $\mathbb{R}^3$. Το μητρώο $P = (I - \boldsymbol{u}\boldsymbol{u}^{\mathrm{T}})$ είναι το μητρώο προβολής επί του επιπέδου με εξίσωση $ax + by + cz = 0$, βλέπε Παράδειγμα 7.3. Θεωρούμε τον γραμμικό μετασχηματισμό $T : \mathbb{R}^3 \to \mathbb{R}^3$ που ορίζεται ως $T(\boldsymbol{x}) = P\boldsymbol{x}$.

   (α´) Να βρεθεί ο τύπος του T.

   (β´) Να βρεθεί η εικόνα $\operatorname{image} T$ του T.

   (γ´) Να βρεθεί ο πυρήνας $\operatorname{kernel} T$ του μετασχηματισμού.

8. Θεωρούμε τους διανυσματικούς χώρους των πολυωνύμων $\mathbb{P}_2$ και $\mathbb{P}_3$ και τον μετασχηματισμό ολοκλήρωμα $J : \mathbb{P}_2 \to \mathbb{P}_3$ όπου για $p \in \mathbb{P}_3$ είναι $J(p) = P$ με

$$P(x) = \int_0^x p(t)\,\mathrm{d}t.$$

   (α´) Να βρεθεί η εικόνα $\operatorname{image} J$ του J.

   (β´) Εξετάστε αν ο μετασχηματισμός ένα-προς-ένα.

   (γ´) Να βρεθεί το μητρώο που παριστάνει τον J.

9. Θεωρούμε τους διανυσματικούς χώρους των πολυωνύμων $\mathbb{P}_2$ και $\mathbb{P}_3$ και τον μετασχηματισμό $T : \mathbb{P}_2 \to \mathbb{P}_2$ όπου για $p \in \mathbb{P}_3$ είναι $T(p) = P$ με

$$P(x) = xp(x).$$

   (α´) Εξετάστε αν ο μετασχηματισμός ένα-προς-ένα.

   (β´) Να βρεθεί το μητρώο που παριστάνει τον T.

10. Αν X και Y είναι διανυσματικοί χώροι με βάσεις $\mathcal{B}_X = \{\boldsymbol{u}_1, \boldsymbol{u}_2, \boldsymbol{u}_3\}$ και $\mathcal{B}_Y = \{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3\}$ αντίστοιχα, και αν $T : X \to Y$ είναι ένας γραμμικός μετασχηματισμός, να βρεθεί το μητρώο αναπαράστασης του T ως προς αυτές τις βάσεις σε κάθε μια από τις περιπτώσεις.

   (α´) $T(\boldsymbol{u}_i) = \boldsymbol{v}_i$, $i = 1, 2, 3$.

   (β´) $T(\boldsymbol{u}_1) = \boldsymbol{v}_3$, $T(\boldsymbol{u}_2) = \boldsymbol{v}_2$, $T(\boldsymbol{u}_3) = \boldsymbol{v}_1$.

   (γ´) $T(\boldsymbol{u}_1) = \boldsymbol{v}_1$, $T(\boldsymbol{u}_2) = \boldsymbol{v}_1 + \boldsymbol{v}_2$, $T(\boldsymbol{u}_3) = \boldsymbol{v}_1 + \boldsymbol{v}_2 + \boldsymbol{v}_3$.

   (δ´) $T(\boldsymbol{u}_1) = \boldsymbol{v}_1 + \boldsymbol{v}_2 + \boldsymbol{v}_3$, $T(\boldsymbol{u}_2) = \boldsymbol{v}_1 + \boldsymbol{v}_2$, $T(\boldsymbol{u}_3) = \boldsymbol{v}_3$.

Βιβλιογραφία

[1] Ε. Γαλλόπουλος: *Γραμμική Άλγεβρα Διαλέξεις*. Αναρτημένες στη πλατφόρμα e-class του Πανεπιστημίου Πατρών, Πάτρα, 2018.

[2] H. Anton: *Elementary Linear Algebra*, 11th edition, Wiley, 2013.

[3] S. Friedberg, A. Insel, L. Spence: *Linear Algebra*, 4th edition, Prentice Hall, New Jersey, 2003.

[4] P. R. Halmos: *Finite-Dimensional Vector Spaces*, Corr. 2nd printing, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1993.

[5] K. Hoffman, R. Kunze: *Linear Algebra*, 2th edition, Prentice Hall, New Jersey, 1971.

[6] K. Jänich: *Linear Algebra*, Undergraduate Texts in Mathematics, Springer-Verlag, New York, 1994.

[7] S. Lang: *Linear Algebra*, 3rd edition, Undergraduate Texts in Mathematics, Springer, New York, 1978.

[8] C. D. Meyer: *Matrix Analysis and Applied Linear Algebra*, SIAM 2000.

[9] G. Strang: *Εισαγωγή στη Γραμμική Άλγεβρα*, Μετάφραση, Εκδόσεις Πανεπιστημίου Πατρών, Πάτρα, 2006.

[10] G. Strang: *Γραμμική Άλγεβρα και εφαρμογές*, Μετάφραση, Πανεπιστημιακές Εκδόσεις Κρήτης, Ηράκλειο, 2015 (ημερομηνία τρέχουσας έκδοσης).

[11] J. J. Sylvester: *Additions to the articles "On a new class of theorems" and "On Pascal's Theorem"*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, Series 3, Volume 37, 1850, pages 363-370.

[12] J. J. Sylvester: *On the relation between the minor determinants of linearly equivalent quadratic functions*, The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science, Series 4, Volume 1, 1851, pages 295-305.

[13] F. Zhang: *Linear Algebra, Chalenging Problems for Students*, 2nd edition, The Johns Hopkins University Press, Baltimore, 2009.

Ευρετήριο όρων