ΓΡΑΜΜΙΚΗ ΑΛΓΕΒΡΑ

Διάλεξη 9

Ορθοκανονικοποίηση
Ορθογώνιο συμπλήρωμα
Ορθογώνιες προβολές

Ε. Στεφανόπουλος & Ε. Γαλλόπουλος

Τμήμα Μηχανικών Η/Υ & Πληροφορικής, Πανεπιστήμιο Πατρών

28 Νοεμβρίου 2022

1. 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 γνώση μιας τέτοιας βάσης είναι γνωστή δεν συμβαίνει το ίδιο για άλλους χώρους. Για παράδειγμα στο χώρο των πολυωνύμων $\mathbb{P}_3[-1,1]$ 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 $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ παράγεται με συστηματικό τρόπο ένα ορθοκανονικό σύνολο $S' = \{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ 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 $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ ένα σύνολο γραμμικά ανεξαρτήτων διανυσμάτων του X.

- Θέτουμε

$$\mathbf{w}_1 = \mathbf{u}_1 / \|\mathbf{u}_1\|.$$

Είναι προφανές ότι $\mathrm{span}\{\mathbf{u}_1\} = \mathrm{span}\{\mathbf{w}_1\}$.

- Θέτουμε

$$\mathbf{v}_2 = \mathbf{u}_2 - \langle \mathbf{u}_2, \mathbf{w}_1\rangle \mathbf{w}_1$$

αφαιρούμε δηλαδή από το $\mathbf{u}_2$ την $\mathbf{w}_1$-συνιστώσα του. Δείχνουμε ότι $\mathrm{span}\{\mathbf{u}_1, \mathbf{u}_2\} = \mathrm{span}\{\mathbf{w}_1, \mathbf{v}_2\}$, ισοδύναμα αν $\alpha, \beta \in \mathbb{R}$, τότε υπάρχουν $a, b \in \mathbb{R}$ ώστε $\alpha\mathbf{u}_1 + \beta\mathbf{u}_2 = a\mathbf{w}_1 + b\mathbf{v}_2$. Από τον ορισμό των $\mathbf{w}_1$ και $\mathbf{v}_2$ υπολογίζουμε

$$\begin{aligned}\alpha\mathbf{u}_1 + \beta\mathbf{u}_2 &= a\mathbf{w}_1 + b\mathbf{v}_2 \\ &= \frac{a\mathbf{u}_1}{\|\mathbf{u}_1\|} + b\Big(\mathbf{u}_2 - \langle \mathbf{u}_2, \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}\rangle \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|}\Big)\end{aligned}$$

οπότε

$$\Big(\alpha - \frac{a}{\|\mathbf{u}_1\|} + \frac{b\langle \mathbf{u}_2, \mathbf{u}_1\rangle}{\|\mathbf{u}_1\|^2}\Big)\mathbf{u}_1 + (\beta - b)\mathbf{u}_2 = \mathbf{0}.$$

Επειδή τα $\mathbf{u}_1$, $\mathbf{u}_2$ είναι γραμμικά ανεξάρτητα βρίσκουμε

$$a = \alpha\|\mathbf{u}_1\| + \beta\langle\mathbf{u}_2,\mathbf{u}_1\rangle/\|\mathbf{u_1}\| \qquad \text{και} \qquad b = \beta.$$

Επίσης

$$\begin{aligned}\langle\mathbf{w}_1,\mathbf{v}_2\rangle &= \langle\mathbf{w}_1,\mathbf{u}_2 - \langle\mathbf{u}_2,\mathbf{w}_1\rangle\mathbf{w}_1\rangle \\ &= \langle\mathbf{w}_1,\mathbf{u}_2\rangle - \langle\mathbf{u}_2,\mathbf{w}_1\rangle\langle\mathbf{w}_1,\mathbf{w}_1\rangle \\ &= 0\end{aligned}$$

αφού $\langle\mathbf{w}_1,\mathbf{w}_1\rangle = 1$. Θέτοντας

$$\mathbf{w}_2 = \mathbf{v}_2/\|\mathbf{v}_2\|$$

το $\{\mathbf{w}_1,\mathbf{w}_2\}$ είναι ορθοκανονικό και $\operatorname{span}\{\mathbf{u}_1,\mathbf{u}_2\} = \operatorname{span}\{\mathbf{w}_1,\mathbf{w}_2\}$.

• Έστω ότι για $k < n$ έχει κατασκευαστεί ένα ορθοκανονικό σύνολο $\{\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_k\}$ τέτοιο ώστε $\operatorname{span}\{\mathbf{u}_1,\mathbf{u}_2,\ldots,\mathbf{u}_k\} = \operatorname{span}\{\mathbf{w}_1,\mathbf{w}_2,\ldots,\mathbf{w}_k\}$. Θέτουμε

$$\mathbf{v}_{k+1} = \mathbf{u}_{k+1} - \langle\mathbf{u}_{k+1},\mathbf{w}_1\rangle\mathbf{w}_1 - \langle\mathbf{u}_{k+1},\mathbf{w}_2\rangle\mathbf{w}_2 - \cdots - \langle\mathbf{u}_{k+1},\mathbf{w}_k\rangle\mathbf{w}_k$$

Για $j < k+1$ έχουμε

$$\begin{aligned}\langle \mathbf{w}_j, \mathbf{v}_{k+1}\rangle &= \langle \mathbf{w}_j, \mathbf{u}_{k+1}\rangle - \langle \mathbf{u}_{k+1}, \mathbf{w}_1\rangle\langle \mathbf{w}_j, \mathbf{w}_1\rangle - \cdots - \langle \mathbf{u}_{k+1}, \mathbf{w}_k\rangle\langle \mathbf{w}_j, \mathbf{w}_k\rangle \\ &= \langle \mathbf{w}_j, \mathbf{u}_{k+1}\rangle - \langle \mathbf{u}_{k+1}, \mathbf{w}_j\rangle\langle \mathbf{w}_j, \mathbf{w}_j\rangle \\ &= 0\end{aligned}$$

αφού $\langle \mathbf{w}_j, \mathbf{w}_j\rangle = 1$. Θέτοντας $\mathbf{w}_{k+1} = \mathbf{v}_{k+1}/\|\mathbf{v}_{k+1}\|$ το $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k, \mathbf{w}_{k+1}\}$ παράγει ότι και το $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_k, \mathbf{u}_{k+1}\}$, αφού

$$\begin{aligned}\mathbf{u}_{k+1} &= \langle \mathbf{u}_{k+1}, \mathbf{w}_1\rangle\mathbf{w}_1 + \langle \mathbf{u}_{k+1}, \mathbf{w}_2\rangle\mathbf{w}_2 + \cdots + \langle \mathbf{u}_{k+1}, \mathbf{w}_k\rangle\mathbf{w}_k + \|\mathbf{v}_{k+1}\|\mathbf{w}_{k+1} \\ &= c_1\mathbf{w}_1 + c_2\mathbf{w}_2 + \cdots + c_k\mathbf{w}_k + c_{k+1}\mathbf{w}_{k+1},\end{aligned}$$

κατά συνέπεια, από την υπόθεση της επαγωγής, έπεται ότι

$$\operatorname{span}\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_{k+1}\} = \operatorname{span}\{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_{k+1}\}.$$

Αποδείξαμε λοιπόν, με επαγωγή, το

Θεώρημα (9.1 Ορθοκανονικοποίηση Gram-Schmidt)

Αν σε χώρο X με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$ και επαγόμενη νόρμα $\|\cdot\| = \sqrt{\langle\cdot,\cdot\rangle}$ το $S = \{\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_n\}$ είναι ένα σύνολο γραμμικά ανεξαρτήτων διανυσμάτων, τότε η ακολουθία των διανυσμάτων που ορίζονται με τη σχέση

$$\begin{aligned} \mathbf{w}_1 &= \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} \\ \mathbf{w}_k &= \frac{\mathbf{u}_k - \langle\mathbf{u}_k, \mathbf{w}_1\rangle\mathbf{w}_1 - \cdots - \langle\mathbf{u}_k, \mathbf{w}_{k-1}\rangle\mathbf{w}_{k-1}}{\|\mathbf{u}_k - \langle\mathbf{u}_k, \mathbf{w}_1\rangle\mathbf{w}_1 - \cdots - \langle\mathbf{u}_k, \mathbf{w}_{k-1}\rangle\mathbf{w}_{k-1}\|}, \end{aligned} \tag{1}$$

$k = 2, \ldots, n$ 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 $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ είναι μια ορθοκανονική βάση για τον X.

Παράδειγμα (9.1)

Μια βάση για τον $\mathbb{R}^3$ αποτελούν τα διανύσματα

$$\mathbf{u}_1 = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}, \qquad \mathbf{u}_2 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix}, \qquad \mathbf{u}_3 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix},$$

τα οποία όμως δεν είναι ανά δύο ορθογώνια, ως προς το σύνηθες εσωτερικό γινόμενο. Ακολουθώντας τη διαδικασία Gram-Schmidt ορίζουμε

$$\mathbf{w}_1 = \frac{\mathbf{u}_1}{\|\mathbf{u}_1\|} = \frac{1}{\sqrt{3}} \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix},$$

και

$$\mathbf{v}_2 = \mathbf{u}_2 - \langle \mathbf{u}_2, \mathbf{w}_1 \rangle \mathbf{w}_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} - \frac{1}{3} 2 \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} = \frac{1}{3} \begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}, \qquad \|\mathbf{v}_2\| = \frac{\sqrt{2}}{\sqrt{3}},$$

τότε

Παράδειγμα (συνέχεια)

$$\mathbf{w}_2 = \frac{\mathbf{v}_2}{\|\mathbf{v}_2\|} = \frac{1}{\sqrt{6}}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix}.$$

Όμοια

$$\mathbf{v}_3 = \mathbf{u}_3 - \langle \mathbf{u}_3, \mathbf{w}_1 \rangle \mathbf{w}_1 - \langle \mathbf{u}_3, \mathbf{w}_2 \rangle \mathbf{w}_2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} - \frac{1}{3}\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - \frac{1}{6}\begin{pmatrix} 1 \\ 1 \\ -2 \end{pmatrix} = \frac{1}{2}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix},$$

με

$$\|\mathbf{v}_3\| = \frac{1}{\sqrt{2}},$$

οπότε

$$\mathbf{w}_3 = \frac{\mathbf{v}_3}{\|\mathbf{v}_3\|} = \frac{1}{\sqrt{2}}\begin{pmatrix} 1 \\ -1 \\ 0 \end{pmatrix}.$$

Έτσι αρχίζοντας με τα $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ η διαδικασία Gram-Schmidt

Παράδειγμα (συνέχεια)

παράγει την ορθοκανονική βάση

$$\mathbf{w}_1 = \frac{1}{\sqrt{3}}\begin{pmatrix}1\\1\\1\end{pmatrix}, \qquad \mathbf{w}_2 = \frac{1}{\sqrt{6}}\begin{pmatrix}1\\1\\-2\end{pmatrix}, \qquad \mathbf{w}_3 = \frac{1}{\sqrt{2}}\begin{pmatrix}1\\-1\\0\end{pmatrix}.$$

Παρατήρηση (9.1)

Παρατηρούμε ότι η ορθοκανονική βάση $\{\mathbf{w}_1, \mathbf{w}_2, \mathbf{w}_3\}$ του Παραδείγματος 9.1 που προέκυψε από την διαδικασία ορθοκανονικοποίησης της $\{\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3\}$ δεν είναι τόσο εύχρηστη όσο η κανονική βάση $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$. Για παράδειγμα ποιά είναι η αναπαράσταση του $\begin{pmatrix}2 & -1 & 3\end{pmatrix}^{\mathrm{T}}$ σε αυτή τη βάση; Αλλάζοντας την αρίθμηση στα $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3$ βλέπουμε ότι προκύπτουν $3!$ ορθοκανονικές βάσεις για το $\mathbb{R}^3$.

Άσκηση 9.1

Αν $\mathbf{u}_1$, $\mathbf{u}_2$, $\mathbf{u}_3$ είναι όπως στο Παράδειγμα 9.1, θέτουμε $\mathbf{b}_1 = \mathbf{u}_3$, $\mathbf{b}_2 = \mathbf{u}_2$, $\mathbf{b}_3 = \mathbf{u}_1$. Να βρεθεί η ορθοκανονική βάση του $\mathbb{R}^3$ που παράγει η διαδικασία Gram-Schmidt από την βάση $\{\mathbf{b}_1, \mathbf{b}_2, \mathbf{b}_3\}$.

2. Ορθογώνιο συμπλήρωμα

Ορισμός (9.1)

Αν S είναι ένα υποσύνολο ενός διανυσματικού χώρου X με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$ το σύνολο όλων των διανυσμάτων του X τα οποία είναι ορθογώνια σε κάθε διάνυσμα του S λέγεται **ορθογώνιο συμπλήρωμα** (orthogonal complement) του S και συμβολίζεται με $S^{\perp}$, έτσι

$$S^{\perp} = \{\mathbf{x} \in X : \langle \mathbf{x}, \mathbf{s} \rangle = 0 \text{ για κάθε } \mathbf{s} \in S\}.$$

Αν το S είναι υποσύνολο του X, $\mathbf{x}, \mathbf{y} \in S^{\perp}$, $\lambda, \mu \in \mathbb{R}$ και $\mathbf{s} \in S$, τότε

$$\begin{aligned} \langle \lambda\mathbf{x} + \mu\mathbf{y}, \mathbf{s} \rangle &= \langle \lambda\mathbf{x}, \mathbf{s} \rangle + \langle \mu\mathbf{y}, \mathbf{s} \rangle \\ &= \lambda\langle \mathbf{x}, \mathbf{s} \rangle + \mu\langle \mathbf{y}, \mathbf{s} \rangle = 0 \end{aligned}$$

από τον ορισμό του $S^{\perp}$, κατά συνέπεια $\lambda\mathbf{x} + \mu\mathbf{y} \in S^{\perp}$. Δείξαμε λοιπόν ότι

Θεώρημα (9.2)

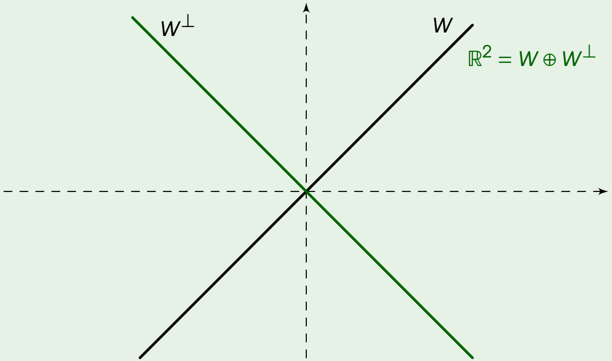
Αν το S είναι υποσύνολο του διανυσματικού χώρου X με εσωτερικό γινόμενο, το ορθογώνιο συμπλήρωμα $S^{\perp}$ του S είναι υπόχωρος του X.

Παράδειγμα (9.2)

Η ευθεία με εξίσωση $y = x$ είναι ο υπόχωρος του $\mathbb{R}^2$

$$W = \left\{ \begin{pmatrix} x \\ x \end{pmatrix} : x \in \mathbb{R} \right\}.$$

Να βρεθεί ο υπόχωρος $W^\perp$.



Παράδειγμα (συνέχεια)

Αν $\begin{pmatrix} y & z \end{pmatrix}^{\mathrm{T}} \in W^{\perp}$, τότε

$$\begin{pmatrix} x \\ x \end{pmatrix} \cdot \begin{pmatrix} y \\ z \end{pmatrix} = xy + xz = x(y+z) = 0$$

για κάθε $x \in \mathbb{R}$, επομένως $z = -y$. 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 $y = -x$.

´Ασκηση 9.2

Να βρεθεί το ορθογώνιο συμπλήρωμα του υποσυνόλου του $\mathbb{R}^2$

$$S = \left\{ \begin{pmatrix} x \\ x+1 \end{pmatrix} : x \in \mathbb{R} \right\}.$$

Παρατήρηση (9.2)

Σε σχέση με το Παράδειγμα 9.2 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 $\mathbf{x} \in \mathbb{R}^2$ γράφεται σαν το άθροισμα $\mathbf{x} = \mathbf{w} + \mathbf{w}'$, όπου $\mathbf{w} \in W$ και $\mathbf{w}' \in W^{\perp}$. Το αποτέλεσμα αυτό γενικεύεται.

Θεώρημα (9.3)

Αν το W είναι υπόχωρος του διανυσματικού χώρου X πεπερασμένης διάστασης, με εσωτερικό γινόμενο, τότε $X = W \oplus W^{\perp}$, δηλαδή ο X είναι το ευθύ άθροισμα των W και $W^{\perp}$.

Πόρισμα (9.1)

Αν το W είναι υπόχωρος του διανυσματικού χώρου X διάστασης n, με εσωτερικό γινόμενο, τότε $\dim W^{\perp} = n - \dim W$.

Πρόταση (9.1)

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

$$X = W \oplus W^{\perp} \qquad \text{και} \qquad X = W^{\perp} \oplus (W^{\perp})^{\perp},$$

ως συνέπεια του Θεωρήματος 9.4.

Απόδειξη.

Δείχνουμε πρώτα ότι $(W^{\perp})^{\perp} \subseteq W$. Αν $\mathbf{x} \in (W^{\perp})^{\perp}$, τότε ως στοιχείο του χώρου X γράφεται ως $\mathbf{x} = \mathbf{w} + \mathbf{z}$, όπου $\mathbf{w} \in W$ και $\mathbf{z} \in W^{\perp}$, βλέπε Θεώρημα 9.4. Έτσι υπολογίζουμε

$$\mathbf{0} = \langle \mathbf{x}, \mathbf{z} \rangle = \langle \mathbf{w} + \mathbf{z}, \mathbf{z} \rangle = \langle \mathbf{w}, \mathbf{z} \rangle + \langle \mathbf{z}, \mathbf{z} \rangle = \|\mathbf{z}\|^2$$

αφού $\mathbf{w} \perp \mathbf{z}$, συνεπώς, από τον ορισμό της νόρμας, παίρνουμε $\mathbf{z} = \mathbf{0}$, ισοδύναμα $\mathbf{x} = \mathbf{w} \in W$, δηλαδή $(W^{\perp})^{\perp} \subseteq W$. Απομένει να δείξουμε ότι $W \subseteq (W^{\perp})^{\perp}$. Πράγματι αν $\mathbf{x} \in W$, τότε $\langle \mathbf{x}, \mathbf{y} \rangle = 0$ για κάθε $\mathbf{y} \in W^{\perp}$, επομένως $\mathbf{x} \in (W^{\perp})^{\perp}$, που είναι το ζητούμενο. □

Ερώτημα: Αν X είναι διανυσματικός χώρος με εσωτερικό γινόμενο και V, W είναι υπόχωροι του X τέτοιοι ώστε

$$X = V \oplus V^{\perp}, \qquad \text{και} \qquad X = V \oplus W,$$

είναι αλήθεια ότι $W = V^{\perp}$;

Παράδειγμα (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 $L_2 = L_3$.

3. Το Θεμελιώδες Θεώρημα, 2ο μέρος

Στο Θεμελιώδες Θεώρημα αποδείξαμε ότι εάν A είναι ένα $n \times m$ μητρώο, τότε

$$\mathbb{R}^m = \operatorname{null} A \oplus \operatorname{range} A^{\mathrm{T}}, \qquad \text{και} \qquad \mathbb{R}^n = \operatorname{null} A^{\mathrm{T}} \oplus \operatorname{range} A,$$

ενώ από το Θεώρημα 9 έπεται ότι

$$\mathbb{R}^m = \operatorname{null} A \oplus (\operatorname{null} A)^{\perp}, \qquad \text{και} \qquad \mathbb{R}^n = (\operatorname{range} A)^{\perp} \oplus \operatorname{range} A.$$

Στη συνέχεια αποδεικνύουμε το επιπλέον αποτέλεσμα

Θεώρημα (9.4 Θεμελιώδες Θεώρημα, μέρος 2)

Εάν A είναι ένα $n \times m$ μητρώο, τότε

1. $(\operatorname{null} A)^{\perp} = \operatorname{range} A^{\mathrm{T}}$*, δηλαδή το ορθογώνιο συμπλήρωμα του* $\operatorname{null} A$ *είναι ο* $\operatorname{range} A^{\mathrm{T}}$.
2. $(\operatorname{range} A)^{\perp} = \operatorname{null} A^{\mathrm{T}}$*, δηλαδή το ορθογώνιο συμπλήρωμα του* $\operatorname{range} A$ *είναι ο* $\operatorname{null} A^{\mathrm{T}}$.

Απόδειξη.

(1) Αν $\mathbf{x} \in (\operatorname{null} A)^{\perp}$, επειδή $\mathbb{R}^m = \operatorname{null} A \oplus \operatorname{range} A^{\mathrm{T}}$, έπεται ότι $\mathbf{x} \in \operatorname{range} A^{\mathrm{T}}$, οπότε $(\operatorname{null} A)^{\perp} \subseteq \operatorname{range} A^{\mathrm{T}}$. Δείχνουμε ότι $\operatorname{range} A^{\mathrm{T}} \subseteq (\operatorname{null} A)^{\perp}$. Αν $\mathbf{x} \in \operatorname{range} A^{\mathrm{T}}$, με $\mathbf{x} \neq \mathbf{0}$, τότε $\mathbf{x} = A^{\mathrm{T}}\mathbf{y}$, για κάποιο $\mathbf{y} \in \mathbb{R}^n$, και

$$0 < \|\mathbf{x}\|^2 = \langle \mathbf{x}, \mathbf{x} \rangle = \langle A^{\mathrm{T}}\mathbf{y}, \mathbf{x} \rangle = (A^{\mathrm{T}}\mathbf{y})^{\mathrm{T}}\mathbf{x} = \mathbf{y}^{\mathrm{T}}A\mathbf{x} = \langle \mathbf{y}, A\mathbf{x} \rangle$$

επομένως $A\mathbf{x} \neq \mathbf{0}$, συνεπώς $\mathbf{x} \notin \operatorname{null} A$, άρα $\mathbf{x} \in (\operatorname{null} A)^{\perp}$, γεγονός που συμπληρώνει την απόδειξη του (1). Για το (2) παρατηρούμε ότι με A^{T} στη θέση του A από το (1) έχουμε $(\operatorname{null} A^{\mathrm{T}})^{\perp} = \operatorname{range} A$, έτσι

$$((\operatorname{null} A^{\mathrm{T}})^{\perp})^{\perp} = (\operatorname{range} A)^{\perp} \Leftrightarrow \operatorname{null} A^{\mathrm{T}} = (\operatorname{range} A)^{\perp}.$$

□

Σημειώνουμε ότι εάν A είναι ένα $n \times m$ μητρώο, $\mathbf{x} \in \mathbb{R}^m$ και $\mathbf{y} \in \mathbb{R}^n$, τότε

$$\langle A\mathbf{x}, \mathbf{y} \rangle = \langle \mathbf{x}, A^{\mathrm{T}}\mathbf{y} \rangle,$$

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 $\mathbb{R}^m$.

Παράρτημα: 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 **;;**, τα πολυώνυμα της φυσιολογικής βάσης $\mathscr{B} = \{1, x, x^2, x^3, \dots\}$ δεν είναι ορθογώνια μεταξύ τους η διαδικασία οθοκανονικοποίησης Gram-Schmidt παράγει μια ακολουθία από ορθογώνια πολυώνυμα. Στη πράξη και σε εφαρμογές στα Μαθηματικά και στη Φυσική υπάρχουν διάφορες ακολουθίες ορθογωνίων πολυωνύμων, τα οποία προκύπτουν από την $\mathscr{B}$ με τη διαδικασία ορθογωνοποίησης, 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 δίνουμε μερικά χακτηριστικά παραδείγματα βάσεων-συστημάτων ορθογωνίων πολυωνύμων.

Πολυώνυμα Legendre. Παράγονται από την $\mathscr{B}$ 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. $P_2(x) = (1/2)(3x^2 - 1)$
4. $P_3(x) = (1/2)(5x^3 - 3x)$
5. $P_4(x) = (1/8)(35x^4 - 30x^2 + 3)$

Αποδεικνύεται ότι τα πολυώνυμα Legendre δίνονται από την αναδρομική σχέση

$$P_0(x) = 1, \qquad P_1(x) = x$$
$$(n+1)P_{n+1}(x) = (2n+1)xP_n(x) - nP_{n-1}(x), \quad n = 1, 2, 3, \ldots$$

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 $\mathscr{B}$ 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

$$T_0(x) = 1, \qquad T_1(x) = x$$
$$T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x), \quad n = 1, 2, 3, \ldots$$

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 $\mathscr{B}$ 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 $U_n(x)$, $n = 0, 1, 2, \ldots$. 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. $U_4(x) = 16x^4 - 12x^2 + 1$

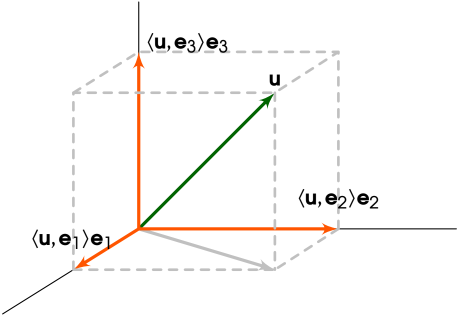
Αποδεικνύεται ότι τα πολυώνυμα Chebyshev δεύτερου είδους δίνονται από την αναδρομική σχέση

$$U_0(x) = 1, \qquad U_1(x) = 2x$$
$$U_{n+1}(x) = 2xU_n(x) - U_{n-1}(x), \quad n = 1, 2, 3, \ldots$$

Κάποιες χαρακτηριστικές ιδιότητες

$$\int_{-1}^{1} U_n(x) U_m(x) \sqrt{1-x^2}\,\mathrm{d}x = \frac{\pi}{2}\delta_{nm}$$
$$U_n(-x) = (-1)^n U_n(x), \qquad \int_{-1}^{1} U_n(x)\,\mathrm{d}x = \ldots, \quad n > 1.$$

4. Προβολές



Γράφοντας το διάνυσμα $\mathbf{u} \in \mathbb{R}^3$ ως

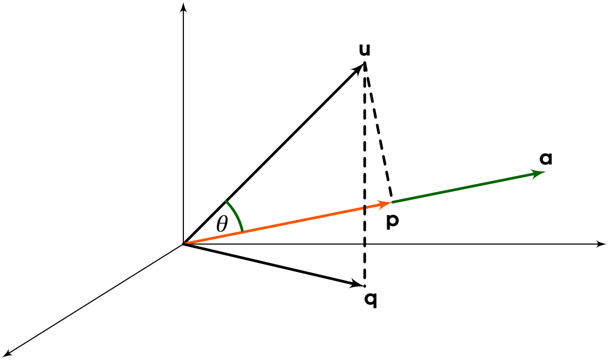
$$\mathbf{u} = \langle \mathbf{u}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle \mathbf{u}, \mathbf{e}_2 \rangle \mathbf{e}_2 + \langle \mathbf{u}, \mathbf{e}_3 \rangle \mathbf{e}_3,$$

ονομάζουμε τα διανύσματα

$$\mathbf{p}_1 = \langle \mathbf{u}, \mathbf{e}_1 \rangle \mathbf{e}_1, \qquad \mathbf{p}_2 = \langle \mathbf{u}, \mathbf{e}_2 \rangle \mathbf{e}_2, \qquad \mathbf{p}_3 = \langle \mathbf{u}, \mathbf{e}_3 \rangle \mathbf{e}_3$$

προβολές (projections) του $\mathbf{u}$, αντίστοιχα, στον x-άξονα, y-άξονα, z-άξονα, ή στους υποχώρους του $\mathbb{R}^3$, $\text{span}\{\mathbf{e}_1\}$, $\text{span}\{\mathbf{e}_2\}$, και $\text{span}\{\mathbf{e}_3\}$ αντίστοιχα.

Αν **a** είναι ένα σταθερό μη μηδενικό διάνυσμα στον $\mathbb{R}^3$ και **u** είναι ένα τυχαίο διάνυσμα στον $\mathbb{R}^3$ μας ενδιαφέρει να βρούμε την προβολή του **u** επί της ευθείας η οποία περιέχει το **a**.



Αν θ είναι η γωνία μεταξύ των **a** και **u** και **p** είναι η προβολή του **u** επί της ευθείας δια του **a**, τότε

$$\|\mathbf{p}\| = \|\mathbf{u}\| \cos\theta$$

κατά συνέπεια

$$\mathbf{p} = \|\mathbf{u}\| \cos\theta \frac{\mathbf{a}}{\|\mathbf{a}\|} = \frac{\|\mathbf{a}\| \|\mathbf{u}\| \cos\theta}{\|\mathbf{a}\|^2} \mathbf{a}$$

επομένως από τον ορισμό του συνήθους εσωτερικού γινομένου προκύπτει ότι

$$\mathbf{p} = \frac{\mathbf{a} \cdot \mathbf{u}}{\mathbf{a} \cdot \mathbf{a}} \mathbf{a} = \frac{\mathbf{a}^{\mathrm{T}} \mathbf{u}}{\mathbf{a}^{\mathrm{T}} \mathbf{a}} \mathbf{a} \tag{2}$$

Στο Σχήμα βλέπουμε, επίσης, την προβολή $\mathbf{q}$ του διανύσματος $\mathbf{u}$ επί του xy-επιπέδου, δηλαδή επί του υπόχωρου $\operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2\}$ του $\mathbb{R}^3$, η οποία είναι

$$\mathbf{q} = \mathbf{p}_1 + \mathbf{p}_2 = \frac{\mathbf{e}_1 \cdot \mathbf{u}}{\mathbf{e}_1 \cdot \mathbf{e}_1} \mathbf{e}_1 + \frac{\mathbf{e}_2 \cdot \mathbf{u}}{\mathbf{e}_2 \cdot \mathbf{e}_2} \mathbf{e}_2$$

ισοδύναμα, αφού $\mathbf{e}_1 \cdot \mathbf{e}_1 = \mathbf{e}_2 \cdot \mathbf{e}_2 = 1$,

$$\mathbf{q} = \langle \mathbf{u}, \mathbf{e}_1 \rangle \mathbf{e}_1 + \langle \mathbf{u}, \mathbf{e}_2 \rangle \mathbf{e}_2$$

όπου $\langle \cdot, \cdot \rangle$ είναι το σύνηθες εσωτερικό γινόμενο στον $\mathbb{R}^3$.

Ορισμός

Έστω X ένας διανυσματικός χώρος με εσωτερικό γινόμενο $\langle\cdot,\cdot\rangle$. Εάν W είναι ένας πεπερασμένης διάστασης υπόχωρος του X και $\{\mathbf{w}_1, \mathbf{w}_2, \ldots, \mathbf{w}_n\}$ είναι μια ορθοκανονική βάση του W και $\mathbf{u}$ είναι ένα διάνυσμα του X ορίζουμε την **ορθογώνια προβολή του u επί του** W το διάνυσμα

$$\mathrm{proj}_W \mathbf{u} := \langle \mathbf{u}, \mathbf{w}_1\rangle \mathbf{w}_1 + \langle \mathbf{u}, \mathbf{w}_2\rangle \mathbf{w}_2 + \cdots + \langle \mathbf{u}, \mathbf{w}_n\rangle \mathbf{w}_n. \tag{3}$$

Το διάνυσμα

$$\mathbf{u} - \mathrm{proj}_W \mathbf{u} = \mathbf{u} - \langle \mathbf{u}, \mathbf{w}_1\rangle \mathbf{w}_1 - \langle \mathbf{u}, \mathbf{w}_2\rangle \mathbf{w}_2 - \cdots - \langle \mathbf{u}, \mathbf{w}_n\rangle \mathbf{w}_n \tag{4}$$

θα το λέμε **ορθογώνια στο** W **συνιστώσα του u**.

Ο ορισμός του $\mathbf{u} - \mathrm{proj}_W \mathbf{u}$ ως ορθογώνια στο W συνιστώσα του $\mathbf{u}$ διακαιολογείται από το γεγονός ότι γράφοντας

$$\mathbf{u} = \mathrm{proj}_W \mathbf{u} + \mathbf{u} - \mathrm{proj}_W \mathbf{u}$$

ισχύει

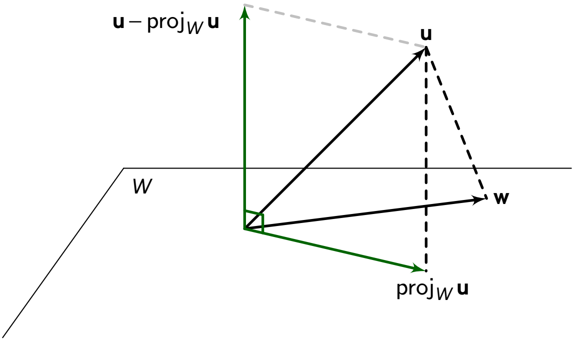
$$\langle \mathrm{proj}_W \mathbf{u}, \mathbf{u} - \mathrm{proj}_W \mathbf{u}\rangle = 0 \Rightarrow \mathbf{u} - \mathrm{proj}_W \mathbf{u} \in W^{\perp}.$$

Θεώρημα (Βέλτιστη προσέγγιση)

Εάν W είναι ένας πεπερασμένης διάστασης υπόχωρος ενός διανυσματικού χώρου X με εσωτερικό γινόμενο και επαγόμενη νόρμα $\|\cdot\|$, και εάν $\mathbf{u} \in X$, τότε για κάθε $\mathbf{w} \in W$ με $\mathbf{w} \neq \mathrm{proj}_W \mathbf{u}$ ισχύει

$$\|\mathbf{u} - \mathrm{proj}_W \mathbf{u}\| < \|\mathbf{u} - \mathbf{w}\|$$

δηλαδή το διάνυσμα $\mathrm{proj}_W \mathbf{u}$ είναι η βέλτιστη προσέγγιση του $\mathbf{u}$ από το W.



Απόδειξη.

Γράφοντας

$$\mathbf{u}-\mathbf{w}=(\operatorname{proj}_W\mathbf{u}-\mathbf{w})+(\mathbf{u}-\operatorname{proj}_W\mathbf{u})$$

επειδή $\operatorname{proj}_W\mathbf{u}-\mathbf{w}$ είναι σαν διαφορά διανυσμάτων του W διάνυσμα του W θα είναι ορθογώνιο στο $\mathbf{u}-\operatorname{proj}_W\mathbf{u}\in W^{\perp}$, κατά συνέπεια από το Πυθαγόρειο Θεώρημα θα έχουμε ότι

$$\|\mathbf{u}-\mathbf{w}\|^2=\|\operatorname{proj}_W\mathbf{u}-\mathbf{w}\|^2+\|\mathbf{u}-\operatorname{proj}_W\mathbf{u}\|^2.$$

Αφού $\mathbf{w}\neq\operatorname{proj}_W\mathbf{u}$ θα είναι $\|\operatorname{proj}_W\mathbf{u}-\mathbf{w}\|>0$, επομένως από την τελευταία ισότητα έπεται ότι

$$\|\mathbf{u}-\mathbf{w}\|^2>\|\mathbf{u}-\operatorname{proj}_W\mathbf{u}\|^2.$$

απ' όπου έπεται το ζητούμενο. □

5. Το μητρώο της ορθογώνιας προβολής

Η εξοικείωσή μας με τα στοιχειώδη μητρώα επιτρέπει να δούμε, αναφερόμαστε στο αρχικό παράδειγμα, όπου $\mathbf{u} = \begin{pmatrix} u_1 & u_2 & u_3 \end{pmatrix}^{\mathrm{T}}$, ότι

$$\mathbf{p}_1 = \begin{pmatrix} u_1 \\ 0 \\ 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{u}, \qquad \mathbf{p}_2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{u}, \qquad \mathbf{p}_3 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \mathbf{u}$$

ή

$$\mathbf{p}_1 = \begin{pmatrix} \mathbf{e}_1 & \mathbf{0} & \mathbf{0} \end{pmatrix} \mathbf{u} = P_1 \mathbf{u}, \qquad \mathbf{p}_2 = \begin{pmatrix} \mathbf{0} & \mathbf{e}_2 & \mathbf{0} \end{pmatrix} \mathbf{u} = P_2 \mathbf{u}, \qquad \mathbf{p}_3 = \begin{pmatrix} \mathbf{0} & \mathbf{0} & \mathbf{e}_3 \end{pmatrix} \mathbf{u} = P_3 \mathbf{u}$$

και

$$\mathbf{q} = \begin{pmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{0} \end{pmatrix} \mathbf{u} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \mathbf{u} = Q \mathbf{u}.$$

Ερώτημα: Ποιο είναι το μητρώο P της προβολής ενός διανύσματος $\mathbf{u}$ του $\mathbb{R}^m$ πρώτον επί ενός διανύσματος $\mathbf{a}$, και δεύτερον επί ενός υποχώρου W του $\mathbb{R}^m$;

1. Αναζητάμε μητρώο P ώστε $P\mathbf{u} = \mathbf{p} = \mathrm{proj}_{\mathbf{a}}\mathbf{u}$, ισοδύναμα για ποιο P είναι
$$P\mathbf{u} = \frac{\mathbf{a}^{\mathrm{T}}\mathbf{u}}{\mathbf{a}^{\mathrm{T}}\mathbf{a}}\mathbf{a}?$$
Επειδή $(\mathbf{a}^{\mathrm{T}}\mathbf{u})\mathbf{a} = \mathbf{a}(\mathbf{a}^{\mathrm{T}}\mathbf{u}) = (\mathbf{a}\mathbf{a}^{\mathrm{T}})\mathbf{u}$ (γιατί;), τελικά βρίσκουμε
$$\mathrm{proj}_{\mathbf{a}}\mathbf{u} = P\mathbf{u} = \frac{\mathbf{a}\mathbf{a}^{\mathrm{T}}}{\mathbf{a}^{\mathrm{T}}\mathbf{a}}\mathbf{u}. \tag{5}$$
2. Έστω $\mathbf{a}_1, \mathbf{a}_2 \dots, \mathbf{a}_n$ γραμμικά ανεξάρτητα διανύσματα στο $\mathbb{R}^m$ και έστω $W = \mathrm{span}\{\mathbf{a}_1, \mathbf{a}_2 \dots, \mathbf{a}_n\}$. Αν A είναι το μητρώο με στήλες τα $\mathbf{a}_j$, τότε $\mathrm{range}\, A = W$, δηλαδή ο υπόχωρος W είναι ο χώρος στηλών του A. Έστω $\mathbf{c} \in \mathbb{R}^m$ τέτοιο ώστε $\mathrm{proj}_W \mathbf{u} = A\mathbf{c}$. Επειδή το $\mathbf{u} - A\mathbf{c}$ είναι ορθογώνιο στον W έπεται ότι $\mathbf{a}_j^{\mathrm{T}}(\mathbf{u} - A\mathbf{c}) = 0$ για $j = 1, 2, \dots, n$. Κατά συνέπεια
$$A^{\mathrm{T}}(\mathbf{u} - A\mathbf{c}) = 0 \Rightarrow A^{\mathrm{T}}\mathbf{u} = A^{\mathrm{T}}A\mathbf{c}$$
Το μητρώο $A^{\mathrm{T}}A$ είναι αντιστρέψιμο, βλέπε επόμενο Θεώρημα, επομένως
$$(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\mathbf{u} = \mathbf{c} \Rightarrow A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\mathbf{u} = A\mathbf{c}$$
κατά συνέπεια
$$\mathrm{proj}_W \mathbf{u} = P\mathbf{u} = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}}\mathbf{u}. \tag{6}$$

Θεώρημα

Εάν το $m \times n$ μητρώο A έχει γραμμικά ανεξάρτητες στήλες, τότε το μητρώο $A^{\mathrm{T}}A$ είναι αντιστρέψιμο.

Απόδειξη.

Το μητρώο $A^{\mathrm{T}}A$ είναι τετραγωνικό ($n \times n$), οπότε αρκεί να δείξουμε ότι $\operatorname{null} A^{\mathrm{T}}A = \{\mathbf{0}\}$. Πρώτα δείχνουμε 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

$$A\mathbf{x} = \mathbf{0} \Rightarrow A^{\mathrm{T}}A\mathbf{x} = A^{\mathrm{T}}\mathbf{0} = \mathbf{0}$$

έπεται ότι $\operatorname{null} A \subseteq \operatorname{null} A^{\mathrm{T}}A$, ενώ από την

$$A^{\mathrm{T}}A\mathbf{x} = \mathbf{0} \Rightarrow \mathbf{x}^{\mathrm{T}}A^{\mathrm{T}}A\mathbf{x} = \mathbf{x}^{\mathrm{T}}\mathbf{0} \Rightarrow (A\mathbf{x})^{\mathrm{T}}A\mathbf{x} = 0 \Rightarrow \|A\mathbf{x}\|^2 = 0$$

έπεται ότι $\operatorname{null} A^{\mathrm{T}}A \subseteq \operatorname{null} A$. Στη συνέχεια δείχνουμε ότι $\operatorname{null} A = \{\mathbf{0}\}$. Πράγματι αν $A\mathbf{x} = \mathbf{0}$ το διάνυσμα $A\mathbf{x}$ ως γραμμικός συνδυασμός των γραμμικά ανεξαρτήτων στηλών του A θα είναι το μηδενικό διάνυσμα, συνεπώς $\mathbf{x} = \mathbf{0}$. Έτσι τελικά έχουμε ότι $\operatorname{null} A^{\mathrm{T}}A = \operatorname{null} A = \{\mathbf{0}\}$ που είναι το ζητούμενο. $\square$

Παράδειγμα

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 15, 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 $\operatorname{rank}(AA^{\mathrm{T}}) \leq \operatorname{rank} A = 2$. Έτσι το μητρώο

Παράδειγμα (συνέχεια)

$$P = A(A^{\mathrm{T}}A)^{-1}A^{\mathrm{T}} = \frac{1}{6}\begin{pmatrix} 1 & 1 \\ 2 & 0 \\ -1 & 1 \end{pmatrix}\begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 & -1 \\ 1 & 0 & 1 \end{pmatrix} = \frac{1}{3}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}$$

είναι η προβολή επί του χώρου στηλών του A, ισοδύναμα για κάθε $\mathbf{u} \in \mathbb{R}^3$, έστω $\mathbf{u} = \begin{pmatrix} x & y & z \end{pmatrix}^{\mathrm{T}}$, υπάρχουν σταθερές λ και μ ώστε

$$P\mathbf{u} = A\begin{pmatrix} \lambda \\ \mu \end{pmatrix} \Leftrightarrow \frac{1}{3}\begin{pmatrix} 2 & 1 & 1 \\ 1 & 2 & -1 \\ 1 & -1 & 2 \end{pmatrix}\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \lambda\begin{pmatrix} 1 \\ 2 \\ -1 \end{pmatrix} + \mu\begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix} = \begin{pmatrix} \lambda + \mu \\ 2\lambda \\ -\lambda + \mu \end{pmatrix}$$

απ' όπου επιλύοντας βρίσκουμε

$$\lambda = \frac{1}{6}(x + 2y - z), \qquad \mu = \frac{1}{2}(x + z).$$

Παρατήρηση

Εάν P είναι μητρώο ορθογώνιας προβολής στο $\mathbb{R}^m$ επί υπόχωρου W, τότε από την (5) ή την (6), ανάλογα την περίπτωση, προκύπτει ότι το P είναι συμμετρικό, δηλαδή $P^{\mathrm{T}} = P$.

Επιπλέον ισχύει ότι $P^2 = P$. Πράγματι για κάθε διάνυσμα $\mathbf{u}$ έχουμε ότι $P\mathbf{u} \in W$, κατά συνέπεια

$$P^2\mathbf{u} = P(P\mathbf{u}) = P\mathbf{u}.$$

Το $I - P$, όπου I είναι το ταυτοτικό μητρώο, είναι επίσης μητρώο προβολής. Προβάλει κάθε διάνυσμα επί του ορθογωνίου συμπληρώματος $W^{\perp}$ του W. Πράγματι για κάθε διάνυσμα $\mathbf{u}$ είναι

$$\mathbf{u} = P\mathbf{u} + \mathbf{u} - P\mathbf{u} = P\mathbf{u} + (I - P)\mathbf{u}$$

και $P\mathbf{u} \perp (I - P)\mathbf{u}$, αφού για κάθε διάνυσμα $\mathbf{v}$, άρα και για $\mathbf{v} = \mathbf{u}$ είναι

Παρατήρηση (συνέχεια)

$$
\begin{aligned}
\langle P\mathbf{u}, (I-P)\mathbf{v}\rangle &= \langle P^2\mathbf{u}, (I-P)\mathbf{v}\rangle \\
&= \langle P\mathbf{u}, P^{\mathrm{T}}(I-P)\mathbf{v}\rangle \\
&= \langle P\mathbf{u}, P(I-P)\mathbf{v}\rangle \\
&= \langle P\mathbf{u}, (P-P^2)\mathbf{v}\rangle \\
&= \langle P\mathbf{u}, (P-P)\mathbf{v}\rangle \\
&= \langle P\mathbf{u}, P\mathbf{v}\rangle - \langle P\mathbf{u}, P\mathbf{v}\rangle,
\end{aligned}
$$

οπότε $\langle P\mathbf{u}, (I-P)\mathbf{v}\rangle = 0$. Έτσι μπορούμε να γράψουμε (γιατί;)

1. $\operatorname{range} P = W$ και $\operatorname{null} P = \operatorname{range}(I-P)$.
2. $\mathbb{R}^m = \operatorname{range} P \oplus \operatorname{null} P$ και $\operatorname{range} P \perp \operatorname{null} P$.

Σημείωση:

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 $\operatorname{range} P$ αν και μόνον αν $P^2 = P$. 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 $\lambda \neq 0$.

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.

Θεώρημα

Μια προβολή P, ένα μητρώο δηλαδή με την ιδιότητα $P^2 = P$, είναι ορθογώνια αν και μόνο αν $P^{\mathrm{T}} = P$.