ΓΡΑΜΜΙΚΗ ΑΛΓΕΒΡΑ

Ε. Γαλλόπουλος & Ε. Στεφανόπουλος

CEID

Διάλεξη 10

15 Μαΐου 2018

Έννοιες, λέξεις κλειδιά, υπενθύμιση I

Υπενθύμιση της 10ης δ.

- Ιδιοτιμές και ιδιοδιανύσματα: Τι, γιατί, ορισμοί.
- Χαρακτηριστικό πολυώνυμο.
- Θεώρημα Cayley-Hamilton και εφαρμογές.
- Συνοδευτικό μητρώο πολυωνύμου.
- Δεξιά και αριστερά ιδιοδιανύσματα.
- Κανονικοποιημένα ιδιοδιανύσματα.
- Υπολογισμός ιδιοτιμών και ιδιοδιανυσμάτων.
- Ελλειμματικές ιδιοτιμές και μητρώα.
- Διαγωνιοποίηση και διαγωνιοποιήσιμα μητρώα.
- Φασματικό ανάπτυγμα μητρώου.
- Ιδιοτιμές και ιδιοδιανύσματα δυνάμεων μητρώου και πολυωνύμων μητρώου.
- Απόδειξη του Θ Cayley-Hamilton.
- Υπολογισμός ιδιοδιανυσμάτων και ιδιοτιμών: η μέθοδος δύναμης.

Έννοιες, λέξεις κλειδιά, υπενθύμιση II

Σήμερα θα συζητήσουμε (όροι 11ης δ.):

- Ιδιοζεύγη γινομένου μητρώων
- Η περίπτωση των μιγαδικών μητρώων
- Ιδιόχωροι
- Ιδιότητες ιδιοτιμών βάσει του είδους του μητρώου
- Ορθογωνιότητα διανυσμάτων
- Συμμετρικά θετικά ορισμένα μητρώα
- Κανονικές μορφές και κανονική μορφή Jordan.
- Παραγοντοποίηση Schur.
- Παραγοντοποίηση ιδιαζουσών τιμών (SVD), ιδιάζουσες τιμές, ιδιάζοντα διανύσματα.
- Προσέγγιση μητρώων μέσω του SVD.
- Ψευδοαντίστροφος μητρώου.

Υπό συζήτηση ενότητες

1 Εισαγωγή στα Διανύσματα 1
- 1.1 Διανύσματα και Γραμμικοί Συνδυασμοί . . . 2
- 1.2 Μήκη και Εσωτερικά Γινόμενα . . . 13

2 Επίλυση Γραμμικών Εξισώσεων 27
- 2.1 Διανύσματα και Γραμμικές Εξισώσεις . . . 27
- 2.2 Η Έννοια της Απαλοιφής . . . 44
- 2.3 Απαλοιφή Χρησιμοποιώντας Πίνακες . . . 58
- 2.4 Κανόνες για τις Πράξεις Πινάκων . . . 71
- 2.5 Αντίστροφοι Πίνακες . . . 89
- 2.6 Απαλοιφή = Παραγοντοποίηση: $A = LU$ . . . 105
- 2.7 Ανάστροφοι και Μεταθέσεις . . . 122

3 Διανυσματικοί Χώροι και Υπόχωροι 141
- 3.1 Χώροι Διανυσμάτων . . . 141
- 3.2 Ο Μηδενόχωρος του A: Επίλυση της $Ax = 0$ . . . 156
- 3.3 Η Τάξη και η Μορφή Ανηγμένων Γραμμών . . . 171
- 3.4 Η Πλήρης Λύση της $Ax = b$ . . . 184
- 3.5 Ανεξαρτησία, Βάση και Διάσταση . . . 199
- 3.6 Διαστάσεις των Τεσσάρων Υποχώρων . . . 219

4 Ορθογωνιότητα 233
- 4.1 Ορθογωνιότητα των Τεσσάρων Υποχώρων . . . 233
- 4.2 Προβολές . . . 246
- 4.3 Προσεγγίσεις Ελάχιστων Τετραγώνων . . . 261
- 4.4 Ορθογώνιες Βάσεις και **Gram – Schmidt** . . . 277

5 Ορίζουσες 295
- 5.1 Οι Ιδιότητες των Οριζουσών . . . 295
- 5.2 Μεταθέσεις και Αλγεβρικά Συμπληρώματα . . . 309
- 5.3 Κανόνας **Cramer**, Αντίστροφοι και Όγκοι . . . 327

6 Ιδιοτιμές και Ιδιοδιανύσματα 347
- 6.1 Εισαγωγή στις Ιδιοτιμές . . . 347
- 6.2 Διαγωνοποιώντας έναν Πίνακα . . . 365
- 6.3 Εφαρμογές στις Διαφορικές Εξισώσεις . . . 383
- 6.4 Συμμετρικοί Πίνακες . . . 401
- 6.5 Θετικά Ορισμένοι Πίνακες . . . 416
- 6.6 Όμοιοι Πίνακες . . . 432
- 6.7 Ανάλυση Ιδιαζουσών Τιμών (**SVD**) . . . 443

7 Γραμμικοί Μετασχηματισμοί 457
- 7.1 Η Έννοια του Γραμμικού Μετασχηματισμού . . . 457
- 7.2 Ο Πίνακας ενός Γραμμικού Μετασχηματισμού . . . 468
- 7.3 Αλλαγή Βάσης . . . 485
- 7.4 Η Διαγωνοποίηση και ο Ψευδοαντίστροφος . . . 494

8 Εφαρμογές 507
- 8.1 Πίνακες στη Μηχανική . . . 507
- 8.2 Γραφήματα και Δίκτυα . . . 521
- 8.3 Πίνακες **Markov** και Οικονομικά Μοντέλα . . . 535
- 8.4 Γραμμικός Προγραμματισμός . . . 545
- 8.5 Σειρές **Fourier**: Γραμμική Άλγεβρα για Συναρτήσεις . . . 553
- 8.6 Γραφικά με Ηλεκτρονικό Υπολογιστή . . . 561

9 Αριθμη[illegible] γραμμι[illegible]έβρα 569
- 9.1 [illegible] **Gauss** σ[illegible]ράξη . . . 569
- [illegible]όρμες κ[illegible]είκτες Κ[illegible]τασης . . . 581
- [illegible]ναληπτικές [illegible]θοδοι [illegible] τη Γραμμική Άλγεβρα . . . 589

10 Μιγαδικά Διανύσματα και Πίνακες 603
- 10.1 Μιγαδικοί Αριθμοί . . . 603
- 10.2 Ερμιτιανοί και Μοναδιαίοι Πίνακες . . . 614
- 10.3 Ο Ταχύς Μετασχηματισμός **Fourier** . . . 625

Λύσεις σε Επιλεγμένες Ασκήσεις 635

Ένα Τελικό Διαγώνισμα 689

Παραγοντοποιήσεις Πινάκων 693

Ερωτήσεις Ανασκόπησης επί των Εννοιών 697

Ιδιοτιμές γινομένου μητρώων (Strang, 6.2) I

<u>ΠΡΟΣΟΧΗ</u> ισχύουν μόνον όταν έχουμε ένα εμπλεκόμενο μητρώο και δυνάμεις του.

- $\Lambda(A+B) \neq \Lambda(A)+\Lambda(B)$
- υπάρχουν $\lambda \in \Lambda(AB)$ τέτοια ώστε $\lambda \neq \lambda_i(A)\lambda_j(B)$ για κανένα i,j

Το πρώτο δεν εκπλήσσει, εξάλλου $\det(A)+\det(B) \neq \det(A+B)$.

Μεταθετικότητα και ιδιοδιανύσματα

Ισχύει ότι $AB = BA$ αν και μόνον αν υπάρχει ίδιο X που διαγωνιοποιεί αμφότερα τα μητρώα, δηλ.

$$\exists X \text{ τ.ώ. } X^{-1}AX = \Lambda(A), X^{-1}BX = \Lambda(B).$$

Σχετικά με το γινόμενο: Θυμηθείτε ότι $\det(AB) = \det(BA) = \det(A)\det(B)$. Για τις ιδιοτιμές δεν ισχύει κάτι αντίστοιχο με την 1η ισότητα. Υπάρχει όμως κάτι αντίστοιχο με τη δεύτερη ισότητα.

Στη συνέχεια εξετάζουμε πως συνδέονται οι ιδιοτιμές $\Lambda(AB)$ με τις ιδιοτιμές $\Lambda(BA)$.

Ιδιοτιμές γινομένου μητρώων (Strang, 6.2) II

Χρησιμοποιώντας ομοιότητα, θα δείξουμε πως συνδέονται οι ιδιοτιμές $\Lambda(AB)$ με τις ιδιοτιμές $\Lambda(BA)$. Έστω ότι $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{n\times m}$. Προσοχή, τα μητρώα A, B δεν είναι κατ' ανάγκη τετραγωνικά, το γινόμενό τους όμως πρέπει να είναι! Τότε αν

$$X = \begin{pmatrix} I & A \\ 0 & I \end{pmatrix} \Rightarrow X^{-1} = \begin{pmatrix} I & -A \\ 0 & I \end{pmatrix}$$

$$\begin{pmatrix} I & -A \\ 0 & I \end{pmatrix} \begin{pmatrix} AB & 0_{m,n} \\ B & 0_{n,n} \end{pmatrix} \begin{pmatrix} I & A \\ 0 & I \end{pmatrix} = \begin{pmatrix} 0_{m,m} & 0_{m,n} \\ B & BA \end{pmatrix}$$

Έστω ότι $m \geq n$, τότε

$$(\lambda_1(AB), \ldots, \lambda_n(AB), \lambda_{n+1}(AB), \ldots \lambda_m(AB), \overbrace{0, \ldots 0}^{n}) = (\lambda_1(BA), \ldots, \lambda_n(BA), \underbrace{0, \ldots 0}_{m})$$

επομένως οι ιδιοτιμές του AB είναι ίδιες με τις ιδιοτιμές του BA συν επιπλέον μία μηδενική ιδιοτιμή πολλαπλότητας $m-n$.

$$\Lambda(AB) = \Lambda(BA) \cup \{0\}$$

Μιγαδικά διανύσματα και μητρώα I

Κεφ. 10.2 Strang

Όπως έχουμε πει, τα διανύσματα και τα μητρώα ορίζονται χωρίς πρόβλημα και στο πεδίο των μιγαδικών.

Ισχύουν όλες οι ιδιότητες που έχουμε αναφέρει.

Επέκταση της έννοιας του αναστρόφου:

Αν $a = [\alpha_1, \alpha_2, \ldots, \alpha_n]^\top \in \mathbb{C}^n$, τότε το ανάστροφο είναι (ως συνήθως):

$$a^\top = [\alpha_1, \alpha_2, \ldots, \alpha_n]$$

και το το συζυγές ανάστροφο ή Ερμιτιανό ανάστροφο ορίζεται ως:

$$\begin{aligned} \bar{a}^\top &= [\bar{\alpha}_1, \bar{\alpha}_2, \ldots, \bar{\alpha}_n] \\ &= a^* \end{aligned}$$

Παρόμοια και για τα μητρώα.

Μιγαδικά διανύσματα και μητρώα ΙΙ

Κεφ. 10.2 Strang

Ένα μητρώο καλείται Ερμιτιανό αν $A^* = A$. Αν το μητρώο είναι πραγματικό, τότε όταν λέμε Ερμιτιανό μητρώο εννοούμε συμμετρικό μητρώο. Ένα τετραγωνικό μιγαδικό μητρώο A καλείται ορθομοναδιαίο αν $A^*A = I$

Προσοχή στον ορισμό του εσωτερικού γινομένου (δείτε τις σχετικές διαφάνειες): δηλ. το εσωτερικό γινόμενο

$$\begin{aligned} \langle x, y \rangle &= x^\top y \\ &= x^* y \end{aligned}$$

Ένα μιγαδικό μητρώο για το οποίο $A^*A = I$ καλείται unitary (ορθομοναδιαίο).

Αριστερά ιδιοδιανύσματα

Ποιά είναι η σχέση των ιδιοτιμών/διανυσμάτων του A με αυτά του A^*;

- Αν $Ax = \lambda x$ τότε $x^* A^* = \bar{\lambda} x^*$επομένως το $\bar{\lambda}$ είναι ιδιοτιμή του A^*
- Επομένως οι ιδιοτιμές του A^* είναι οι συζυγείς των ιδιοτιμών του A.

Γενικά δεν υπάρχει απλή σχέση μεταξύ των ιδιοδιανυσμάτων των A, A^*.

- Αν $\bar{\lambda}$ είναι ιδιοτιμή του A^* τότε θα υπάρχει ιδιοδιάνυσμα y ώστε $A^* y = \bar{\lambda} y$
- επομένως $y^* A = \lambda y^*$

Υπόχωροι ιδιοδιανυσμάτων ⇒ Ιδιόχωροι

Ιδιόχωρος ιδιοτιμής:

Ο χώρος που παράγεται από το διάνοιγμα των ιδιοδιανυσμάτων που αντιστοιχούν σε κάποια ιδιοτιμή.

Προσοχή: Οι ιδιόχωροι είναι αναλλοίωτοι ως προς το μητρώο που τους ορίζει:

$$\text{Αν } \mathcal{V} = \text{span}\langle q_{i_1}, \dots, q_{i_k} \rangle \Rightarrow A\mathcal{V} = \mathcal{V}$$

Προσοχή: Στη γενική περίπτωση (διανυσματικού υπόχωρου από διάνοιγμα διανυσμάτων που δεν είναι ιδιοδιανύσματα) δεν ισχύει αυτό. Π.χ. δείτε την περίπτωση, $A = \begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix} \in \mathbb{R}^{2\times 2}$ με μοναδικό ιδιοδιάνυσμα το e_1. Τότε $Ae_1 = e_1 \in \text{span}\langle < e_1 \rangle$ ενώ αν x είναι το διάνυσμα $[\xi_1, \xi_2]^\top$ όπου $\xi_2 \neq 0$, τότε $Ax = [\xi_1 + \xi_2, \xi_2]^\top \notin \text{span}\langle x \rangle$.

Γραμμική ανεξαρτησία και ορθογωνιότητα ιδιοδιανυσμάτων I

Δύο σημαντικές ιδιότητες:

1. Τα ιδιοδιανύσματα ενός μητρώου που αντιστοιχουν σε διαφορετικές ιδιοτιμές, είναι γραμμικά ανεξάρτητα. Απόδειξη: Έστω ότι $Aq_1 = \lambda_1 q_1$, $Aq_2 = \lambda_2 q_2$ και

   $\lambda_1 \neq \lambda_2$. Αν είναι γραμμικά εξαρτημένα τότε για μη μηδενικά α, β ισχύει ότι $\alpha q_1 + q_2 = 0$ Μπορούμε να υποθέσουμε ότι $\alpha \neq 0, \beta = 1$. οπότε, πολλαπλασιάζοντας με το q_2^*,

$$0 = \alpha q_1 + q_2 \Rightarrow 0 = \alpha q_2^* q_1 + q_2^* q_2$$

Πολλαπλασιάζοντας με το $q_2^* A$,

$$\begin{aligned} 0 &= \alpha q_1 + q_2 \Rightarrow 0 = \alpha q_2^* A q_1 + q_2^* A q_2 \\ 0 &= \lambda_1 \alpha q_2^* q_1 + \lambda_2 q_2^* q_2 \end{aligned}$$

Άρα

$$\begin{aligned} 0 &= \alpha q_2^* q_1 + q_2^* q_2 \\ 0 &= \lambda_1 \alpha q_2^* q_1 + \lambda_2 q_2^* q_2 \end{aligned}$$

Μία από τις ιδιοτιμές είναι οπωσδήποτε μη μηδενική, έστω $\lambda_2 \neq 0$. Πολλαπλασιάζοντας την πρώτη εξίσωση με λ_2 και αφαιρώντας

$$0 = (\lambda_2 - -\lambda_1)\overbrace{q_2^* q_2}^{\|q_2\|^2}$$

που είναι αδύνατον από τις υποθέσεις μας.

Γραμμική ανεξαρτησία και ορθογωνιότητα ιδιοδιανυσμάτων II

2. Αν ένα μητρώο είναι συμμετρικό, τότε τα ιδιοδιανύσματα διαφορετικών ιδιοτιμών είναι κάθετα μεταξύ τους.

Απόδειξη: (δοκιμάστε - η απόδειξη σε λίγο)

Ιδιοτιμές ερμιτιανών και πραγματικών συμμετρικών μητρώων

$$\begin{aligned} Ax &= \lambda x, A = A^\top \in \mathbb{R}^{n\times n} \Rightarrow x^* A^\top = \bar{\lambda} x^* \\ \Rightarrow (x^* A)x &= \|x\|^2 \bar{\lambda} = x^*(Ax) = \lambda \|x\|^2 \end{aligned}$$

$$\begin{aligned} Ax &= \lambda x, A = A^* \in \mathbb{R}^{n\times n} \Rightarrow x^* A^* = \bar{\lambda} x^* \\ \Rightarrow (x^* A)x &= \|x\|^2 \bar{\lambda} = x^*(Ax) = \lambda \|x\|^2 \end{aligned}$$

άρα, επειδή $\|x\| \neq 0$,

$$\lambda = \bar{\lambda} \Rightarrow \lambda \in \mathbb{R}.$$

Οι ιδιοτιμές των ερμιτιανών και των πραγματικών συμμετρικών μητρώων είναι όλες πραγματικές. Επίσης αν $A = A^\top \in \mathbb{R}^{n\times n}$ τότε όλα τα ιδιοδιανύσματα είναι πραγματικά.

Ορθογωνιότητα ιδιοδιανυσμάτων

συμμετρικών πραγματικών μητρώων

Έστω διαφορετικές ιδιοτιμές $\lambda \neq \mu$ και

$$\begin{aligned} Ax = \lambda x, \quad Ay = \mu y \\ y^\top Ax = \lambda y^\top x = y^\top A^\top x \\ (Ay)^\top x = \mu y^\top x \Rightarrow \mu(y^\top x) = \lambda(y^\top x) \end{aligned}$$

οπότε είτε $\lambda = \mu$ ή $y^\top x = 0$.

Τα ιδιοδιανύσματα ενός πραγματικού συμμετρικού μητρώου που αντιστοιχούν σε διαφορετικές ιδιοτιμές είναι πάντοτε κάθετα μεταξύ τους

Ορθογωνιότητα ιδιοδιανυσμάτων

συμμετρικών πραγματικών μητρώων

Τι γίνεται με τα ιδιοδιανύσματα που αντιστοιχούν σε πολλαπλές ιδιοτιμές;

Κάθε πραγματικό συμμετρικό ή ερμιτιανό μητρώο διαθέτει ακριβώς n ιδιοδιανύσματα που είναι κάθετα μεταξύ τους. Δηλαδή διαθέτει ένα πλήρες σύνολο ορθογωνίων ιδιοδιανυσμάτων ακόμα και αν υπάρχουν ιδιοτιμές με πολλαπλότητα.

$$
\begin{aligned}
A[q_1,\dots,q_n] &= [q_1,\dots,q_n]\Lambda \Leftrightarrow [q_1,\dots,q_n]^\top A[q_1,\dots,q_n] = \Lambda \\
Q^\top A Q &= \Lambda \\
A &= \lambda_1 q_1 q_1^\top + \cdots \lambda_n q_n q_n^\top \quad \text{φασματικό ανάπτυγμα}
\end{aligned}
$$

Επίλυση συστημάτων μέσω φασματικού αναπτύγματος και ερμηνεία

Αν γνωρίζουμε τα X, Λ ώστε $X^{-1}AX = \Lambda$ και αναζητούμε το x τ.ώ. $Ax = b$,

$$\begin{aligned} Ax &= b \\ X^{-1}AXX^{-1}x &= X^{-1}b \end{aligned}$$

οπότε

$$\begin{aligned} x &= X\Lambda^{-1}X^{-1}b \\ &= \frac{1}{\lambda_1}x_1(y_1^*b) + \frac{1}{\lambda_2}x_2(y_2^*b) + \cdots \frac{1}{\lambda_n}x_n(y_n^*b) \end{aligned}$$

<u>ΠΡΟΣΟΧΗ</u> Ο τρόπος αυτός επίλυσης σπάνια συμφέρει ...
πιο χρήσιμο είναι η πληροφορία που παρέχει αναδεικνύοντας ορισμένα χαρακτηριστικά της λύσης! π.χ. πως μεγεθύνονται οι παράγοντες που αντιστοιχούν σε όρους όπου $y_j^*b \neq 0$ και η απόλυτη τιμή $|\lambda_j|$ είναι πολύ μικρή (αν είναι 0, το μητρώο δεν αντιστρέφεται).

Σύνοψη: Πώς διευκολύνουν τα συμμετρικά (πραγματικά) και ερμιτιανά (μιγαδικά) μητρώα;

Αν $A = A^\top \in \mathbb{R}^{n\times n}$ ή $A = A^* \in \mathbb{C}^{n\times n}$ τότε

- όλες οι ιδιοτιμές είναι πραγματικές,
- υπάρχουν n γραμμικά ιδιοδιανύσματα, $\{q_1, \ldots, q_n\}$,
- τα ιδιοδιανύσματα είναι κάθετα μεταξύ τους, άρα $Q^\top Q = I$ ή $Q^* Q = I$,
- τα ιδιοδιανύσματα είναι πραγματικά αν το μητρώο είναι πραγματικό,
- το μητρώο είναι διαγωνιοποιήσιμο,
- $Q^\top A Q = \Lambda$ ή $Q^* A Q = \Lambda$.

Μία ενδιαφέρουσα αναγωγή

Αφού τα συμμετρικά μητρώα έχουν όλες αυτές τις καλές ιδιότητες, έχει ενδιαφέρον το εξής:
Κάθε τετραγωνικό μητρώο $\mathbb{R}^{n\times n}$ μπορεί να γραφτεί ως

$$A = \frac{A+A^\top}{2} + \frac{A-A^\top}{2},$$

Ο πρώτος όρος είναι συμμετρικός και ο δεύτερος αντισυμμετρικός:

$$\left(\frac{A-A^\top}{2}\right)^\top = -\left(\frac{A-A^\top}{2}\right)$$

Ενδιαφέρον: Οι ιδιοτιμές των αντισυμμετρικών μητρώων είναι όλες φανταστικές.

Συμμετρικά θετικά ορισμένα μητρώα

Πολλά μητρώα στις εφαρμογές έχουν την ιδιότητα που θα συζητήσουμε τώρα.
Το αντίστοιχο των «θετικών αριθμών» στην περίπτωση των μητρώων!

Ορισμός

Δίνεται συμμετρικό $A \in \mathbb{R}^{n \times n}$. Τότε Οι παρακάτω ιδιότητες είναι ισοδύναμες:

1. Οι ιδιοτιμές του A είναι όλες θετικές.
2. Για $k = 1, \ldots, n$, $\det A_{1:k,1:k} > 0$.
3. Για κάθε $x \neq 0$, $x^\top A x > 0$.
4. Αν $A = LU$ είναι η παραγοντοποίηση LU του A, όλοι οι οδηγοί (τα διαγώνια στοιχεία του U) είναι θετικοί.

Αν ισχύει μία από αυτές, το μητρώο καλείται συμμετρικό θετικά ορισμένο (ΣΘΟ).

Όταν ένα μητρώο είναι ΣΘΟ τότε

- όλα τα στοιχεία στη διαγώνιο είναι θετικά.
- είναι αντιστρέψιμο.
- υπάρχει κάτω τριγωνικό L τέτοιο ώστε $A = LL^\top$.

Διαγωνιοποίηση(επανάληψη)

Δύο σημαντικές περιπτώσεις για δοθέν $A \in \mathbb{C}^{n \times n}$:

1. είναι διαγωνιοποιήσιμο $\Rightarrow$ υπάρχουν n γ.α. ιδιοδιανύσματα
2. δεν είναι διαγωνιοποιήσιμο (ελλειμματικό) $\Rightarrow$ δεν υπάρχουν n γ.α. ιδιοδιανύσματα

<u>ΠΡΟΣΟΧΗ</u> Αν διαγωνιοποιήσιμο, τότε σε ορισμένες (καλές) περιπτώσεις τα ιδιοδιανύσματα είναι ορθογώνια (αν πραγματικά) ή ορθομοναδιαία (αν μιγαδικά), οπότε

$$Q^\top AQ = \Lambda, \;\; \text{ή} \;\; Q^* AQ = \Lambda$$

Κανονική μορφή Jordan

Τι κάνουμε όταν το μητρώο δεν είναι διαγωνιοποιήσιμο; Δεν υπάρχει παραγοντοποίηση $A = Q\Lambda Q^{-1}$.

Πόσο απλό μπορούμε να κάνουμε γενικό μητρώο A με μετασχηματισμό ομοιότητας;

Για κάθε $A \in \mathbb{C}^{n\times n}$ με ιδιοτιμές $\lambda_1, \ldots, \lambda_n$, υπάρχει μητρώο $X \in \mathbb{C}^{n\times n}$ τέτοιο ώστε

$$X^{-1}AX = J = \begin{pmatrix} J_1 & & & \\ & J_2 & & \\ & & \ddots & \\ & & & J_q \end{pmatrix}$$

- q είναι το πλήθος των γ.α. ιδιοδιανυσμάτων του A.
- Σε κάθε ιδιοτιμή αντιστοιχούν όσα υπομητρώα Jordan είναι η γεωμετρική πολλαπλότητά της.
- Το άθροισμα των διαστάσεων των υπομητρώων Jordan για μία ιδιοτιμή είναι ίσο με την αλγεβρική πολλαπλότητά της.
- Κάθε J_i αποκαλείται απλό (υπο)μητρώο Jordan και έχει τη μορφή

$$J_i = \begin{pmatrix} \lambda_i & 1 & 0 & \cdots \\ 0 & \lambda_i & 1 & 0 \\ \vdots & \ddots & \ddots & \ddots \\ 0 & \cdots & 0 & \lambda_i \end{pmatrix} \in \mathbb{C}^{k_i\times k_i}, \quad \text{όπου } \sum_{i=1}^{q} k_i = n.$$

Παρατηρήσεις

Καλά νέα 1) Η μορφή Jordan αποκαλύπτει τη φασματική δομή του μητρώου.
2) Αποτελεί μία σημαντική κανονική μορφή μητρώου (υπάρχουν και άλλες).

Κακά νέα Αν και ενδιαφέρουσα από μαθηματικής άποψης, υπάρχουν εγγενείς δυσκολίες στον υπολογισμό της μορφής Jordan.

Τι γίνεται στην πράξη Αναζητούμε άλλες κανονικές μορφές.

Παρατηρήσεις

Καλά νέα 1) Η μορφή Jordan αποκαλύπτει τη φασματική δομή του μητρώου.
2) Αποτελεί μία σημαντική κανονική μορφή μητρώου (υπάρχουν και άλλες).

Κακά νέα Αν και ενδιαφέρουσα από μαθηματικής άποψης, υπάρχουν εγγενείς δυσκολίες στον υπολογισμό της μορφής Jordan.

Τι γίνεται στην πράξη Αναζητούμε άλλες κανονικές μορφές.

Οι χρησιμότερες κανονικές μορφές

Παραγοντοποίηση Schur: Για κάθε τετραγωνικό A, υπάρχει ορθομοναδιαίο Q τέτοιο ώστε $Q^*AQ = R$ είναι άνω τριγωνικό. Το R λέγεται η μορφή Schur του A.

Παραγοντοποίηση ιδιαζουσών τιμών (SVD) Για οποιοδήποτε μητρώο A υπάρχουν πάντα ορθομοναδιαία U, V τέτοια ώστε το $U^*AV = \Sigma$ να είναι διαγώνιο (με μη αρνητικά στοιχεία!)

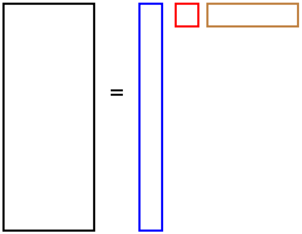
ΠΑΡΑΓΟΝΤΟΠΟΙΗΣΕΙΣ ΜΗΤΡΩΟΥ

Αναφερόμαστε στη δυνατότητα να εκφράσουμε δοθέν μητρώο $A \in \mathbb{C}^{m\times n}$ ως γινόμενο

$$A = WKH, \text{ όπου } k \le n, W = \mathbb{C}^{m\times k}, Y \in \mathbb{C}^{k\times k}, H \in \mathbb{C}^{k\times n}$$

για κάποια κατάλληλη τιμή k.

π.χ. αν $n < m$ και υπάρχει $k \le n$ τ.ώ.



- Η επιλογή κατάλληλου $k \le \min(m, n)$ μπορεί να είναι και αυτή μέρος του προβλήματος).
- Πολλές φορές, μπορεί να θέλουμε να χρησιμοποιήσουμε k τ.ώ. να ισχύει έχουμε ικανοποιητική προσέγγιση $A \approx WKH$.

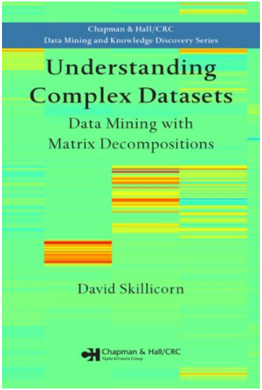
ΠΑΡΑΓΟΝΤΟΠΟΙΗΣΕΙΣ ΜΗΤΡΩΟΥ

Αναφερόμαστε στη δυνατότητα να εκφράσουμε δοθέν μητρώο $A \in \mathbb{C}^{m\times n}$ ως γινόμενο

$$A = WKH,\ \text{όπου } k \leq n, W = \mathbb{C}^{m\times k}, Y \in \mathbb{C}^{k\times k}, H \in \mathbb{C}^{k\times n}$$

για κάποια κατάλληλη τιμή k.

- Η επιλογή κατάλληλου $k \leq \min(m,n)$ μπορεί να είναι και αυτή μέρος του προβλήματος).
- Πολλές φορές, μπορεί να θέλουμε να χρησιμοποιήσουμε k τ.ώ. να ισχύει έχουμε ικανοποιητική προσέγγιση $A \approx WKH$.
- Τέτοιου τύπου προβλήματα εμφανίζονται σε πολλές εφαρμογές: π.χ. Ανάκτηση Πληροφορίας, Επεξεργασία Σημάτων και Εικόνας, Ανάλυση Δεδομένων.



Γνωστές παραγοντοποιήσεις

Γνωρίζουμε (κεφ. 3): Αν $\operatorname{rank}(A) = r$ τότε μπορούμε να βρούμε (από την ΑΓΚΜ) παράγοντες $B \in \mathbb{C}^{m\times r}, C \in \mathbb{C}^{r\times n}$ τ.ώ. $A = BC$

[1] Δείτε τον εκτενή κατάλογο στο σύγγραμμα του Strang.

Γνωστές παραγοντοποιήσεις

Γνωρίζουμε (κεφ. 3): Αν $\mathrm{rank}(A) = r$ τότε μπορούμε να βρούμε (από την ΑΓΚΜ) παράγοντες $B \in \mathbb{C}^{m\times r}, C \in \mathbb{C}^{r\times n}$ τ.ώ. $A = BC$

Παραδείγματα[1] $A = WKH$ όπου $A \in \mathbb{R}^{m\times n}$. Θέσαμε $r = \mathrm{rank}(A)$ (προσέξτε, ορισμένες παραγοντοποιήσεις προϋποθέτουν τετραγωνικό A):

παραγοντοποίηση	W	K	H
$A = BC$ (τάξης)	$B \in \mathbb{R}^{m\times r}$, r στήλες οδηγών A		$C \in \mathbb{R}^{r\times n}$, μη μηδεν. γραμμές ΑΓΚΜ
$A = QR$ $A = Q_1 R_1$	$Q \in \mathbb{R}^{m\times m}$ ορθογώνιο $Q_1 \in \mathbb{R}^{m\times r}$, ΟΚ στήλες		$R \in \mathbb{R}^{m\times n}$ άνω τριγωνικό $R_1 \in \mathbb{R}^{r\times r}$ άνω τριγωνικό
$A = QTQ^*$ (Schur)	Q ορθομοναδιαίο	άνω τριγωνικό T	Q^*
$A = XJX^{-1}$ (ΑΠΙ)	X αντιστρέψιμο	μορφή Jordan J	X^{-1}
$A = X\Lambda X^{-1}$ διαγ/σιμο (ΑΠΙ)	X αντιστρέψιμο	διαγώνιο Λ	X^{-1}
$A = Q\Lambda Q^\top$ συμμετρικό (ΑΠΙ)	Q ορθογώνιο	διαγώνιο Λ	$Q^\top$

Πρόκληση: Να διαγωνιοποιήσουμε οποιοδήποτε μητρώο (ακόμα και μη τετραγωνικό) με ορθογώνιους μετασχηματισμούς.

[1] Δείτε τον εκτενή κατάλογο στο σύγγραμμα του Strang.

Χτίζοντας τη Διάσπαση Ιδιαζουσών Τιμών (SVD) I

Έστω $A \in \mathbb{R}^{m\times n}$ όπου $\mathrm{rank}(A) = r$. Θεωρούμε ότι διαθέτουμε ορθοκανονικές βάσεις για τους 4 θεμελιώδεις υπόχωρους (π.χ. μέσω της ΑΓΚΜ και στη συνέχεια Gram-Schmidt):

$$\begin{aligned}
\mathrm{range}(A) &= \mathrm{span}\{u_1, \dots, u_r\}, \text{ ορίζουμε } U_1 \in \mathbb{R}^{m\times r} \\
\mathrm{null}(A^\top) &= \mathrm{span}\{u_{r+1}, \dots, u_m\}, \text{ ορίζουμε } U_2 \in \mathbb{R}^{m\times (m-r)} \\
\mathrm{range}(A^\top) &= \mathrm{span}\{v_1, \dots, v_r\}, \text{ ορίζουμε } V_1 \in \mathbb{R}^{n\times r} \\
\mathrm{null}(A) &= \mathrm{span}\{v_{r+1}, \dots, v_m\}, \text{ ορίζουμε } V_2 \in \mathbb{R}^{n\times (n-r)}
\end{aligned}$$

και θέτουμε $U = [U_1, U_2], V = [V_1, V_2]$.
Τότε αν χρησιμοποιήσουμε τα παραπάνω,

$$\begin{aligned}
U^\top A V &= \begin{pmatrix} U_1^\top \\ U_2^\top \end{pmatrix} A \begin{pmatrix} V_1 & V_2 \end{pmatrix} \\
&= \begin{pmatrix} U_1^\top A V_1 & U_1^\top A V_2 \\ U_2^\top A V_1 & U_2^\top A V_2 \end{pmatrix} = \begin{pmatrix} C & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix}
\end{aligned}$$

όπου C είναι $r \times r$.

Χτίζοντας τη Διάσπαση Ιδιαζουσών Τιμών (SVD) II

- $U_1^\top AV_2 = 0_{r,n-r}$ γιατί V_2 είναι βάση του $\mathrm{null}(A)$.
- $U_2^\top AV_1 = 0_{m-r,r}$ γιατί U_2 είναι βάση του $\mathrm{null}(A^\top)$.
- Συνεπάγεται ότι $U_2^\top AV_2 = 0_{m-r,n-r}$.
- Το $C \in \mathbb{R}^{r\times r}$ είναι $r \times r$ και έχει τάξη r γιατί $\mathrm{rank}(A) = r = \mathrm{rank}(U^\top AV) = \mathrm{rank}(C)$. Επομένως θα είναι αντιστρέψιμο.
- Αν ονομάσουμε το $U^\top AV = R$ τότε $A = URV^\top$. Η παραγοντοποίηση αυτή ενίοτε ονομάζεται παραγοντοποίηση *URV* (εκτός ύλης!)
- Ένα σημαντικό στοιχείο είναι ότι οι βάσεις U_1, V_1 μπορούν να επιλεγούν έτσι ώστε το $C = U_1^\top AV_1$ να είναι διαγώνιο με θετικά στοιχεία στη διαγώνιο. Η παραγοντοποίηση που προκύπτει λέγεται παραγοντοποίηση ιδιαζουσών τιμών (SVD.)

Απόδειξη ύπαρξης του SVD

Δίνεται $A \in \mathbb{R}^{m \times n}$.

- Εξετάζουμε το μητρώο $A^\top A$: Είναι συμμετρικό θετικά ημιορισμένο,

Απόδειξη ύπαρξης του SVD

Δίνεται $A \in \mathbb{R}^{m\times n}$.

- Εξετάζουμε το μητρώο $A^\top A$: Είναι συμμετρικό θετικά ημιορισμένο,
- επομένως υπάρχουν n μη αρνητικές ιδιοτιμές

Απόδειξη ύπαρξης του SVD

Δίνεται $A \in \mathbb{R}^{m\times n}$.

- Εξετάζουμε το μητρώο $A^\top A$: Είναι συμμετρικό θετικά ημιορισμένο,
- επομένως υπάρχουν n μη αρνητικές ιδιοτιμές

$$\sigma_1^2 \geq \cdots, \geq \sigma_r^2 > \sigma_{r+1}^2 = \cdots = \sigma_n^2$$

και ΟΚ ιδιοδιανύσματα $v_1, ..., v_r, v_{r+1}, \ldots, v_n$. Ονομάζουμε όπως πριν τα V_1, V_2.

Απόδειξη ύπαρξης του SVD

Δίνεται $A \in \mathbb{R}^{m\times n}$.

- Εξετάζουμε το μητρώο $A^\top A$: Είναι συμμετρικό θετικά ημιορισμένο,
- επομένως υπάρχουν n μη αρνητικές ιδιοτιμές

$$\sigma_1^2 \geq \cdots, \geq \sigma_r^2 > \sigma_{r+1}^2 = \cdots = \sigma_n^2$$

  και ΟΚ ιδιοδιανύσματα $v_1, \ldots, v_r, v_{r+1}, \ldots, v_n$. Ονομάζουμε όπως πριν τα V_1, V_2.
- Άρα $A^\top A V_1 = V_1 S^2$ όπου $S = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$

Απόδειξη ύπαρξης του SVD

Δίνεται $A \in \mathbb{R}^{m\times n}$.

- Εξετάζουμε το μητρώο $A^\top A$: Είναι συμμετρικό θετικά ημιορισμένο,
- επομένως υπάρχουν n μη αρνητικές ιδιοτιμές

$$\sigma_1^2 \geq \cdots, \geq \sigma_r^2 > \sigma_{r+1}^2 = \cdots = \sigma_n^2$$

  και ΟΚ ιδιοδιανύσματα $v_1, ..., v_r, v_{r+1}, \ldots, v_n$. Ονομάζουμε όπως πριν τα V_1, V_2.
- Άρα $A^\top A V_1 = V_1 S^2$ όπου $S = \mathrm{diag}(\sigma_1, ..., \sigma_r)$
- άρα $V_1^\top A^\top A V_1 = V_1^\top V_1 S^2 = S^2$

Απόδειξη ύπαρξης του SVD

Δίνεται $A \in \mathbb{R}^{m\times n}$.

- Εξετάζουμε το μητρώο $A^\top A$: Είναι συμμετρικό θετικά ημιορισμένο,
- επομένως υπάρχουν n μη αρνητικές ιδιοτιμές

$$\sigma_1^2 \geq \cdots, \geq \sigma_r^2 > \sigma_{r+1}^2 = \cdots = \sigma_n^2$$

  και ΟΚ ιδιοδιανύσματα $v_1, ..., v_r, v_{r+1}, \ldots, v_n$. Ονομάζουμε όπως πριν τα V_1, V_2.
- Άρα $A^\top A V_1 = V_1 S^2$ όπου $S = \mathrm{diag}(\sigma_1, ..., \sigma_r)$
- άρα $V_1^\top A^\top A V_1 = V_1^\top V_1 S^2 = S^2$
- άρα $S^{-1} V_1^\top A^\top A V_1 S^{-1} = I$ (*).

Απόδειξη ύπαρξης του SVD

Δίνεται $A \in \mathbb{R}^{m \times n}$.

- Εξετάζουμε το μητρώο $A^\top A$: Είναι συμμετρικό θετικά ημιορισμένο,
- επομένως υπάρχουν n μη αρνητικές ιδιοτιμές

$$\sigma_1^2 \geq \cdots, \geq \sigma_r^2 > \sigma_{r+1}^2 = \cdots = \sigma_n^2$$

  και ΟΚ ιδιοδιανύσματα $v_1, \ldots, v_r, v_{r+1}, \ldots, v_n$. Ονομάζουμε όπως πριν τα V_1, V_2.
- Άρα $A^\top A V_1 = V_1 S^2$ όπου $S = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$
- άρα $V_1^\top A^\top A V_1 = V_1^\top V_1 S^2 = S^2$
- άρα $S^{-1} V_1^\top A^\top A V_1 S^{-1} = I$ (*).
- Επίσης $A^\top A V_2 = V_2 0 = 0$ άρα $V_2^\top A^\top A V_2 = 0$ άρα $A V_2 = 0$.

Απόδειξη ύπαρξης του SVD

Δίνεται $A \in \mathbb{R}^{m \times n}$.

- Εξετάζουμε το μητρώο $A^\top A$: Είναι συμμετρικό θετικά ημιορισμένο,
- επομένως υπάρχουν n μη αρνητικές ιδιοτιμές

$$\sigma_1^2 \geq \cdots, \geq \sigma_r^2 > \sigma_{r+1}^2 = \cdots = \sigma_n^2$$

και ΟΚ ιδιοδιανύσματα $v_1, ..., v_r, v_{r+1}, \ldots, v_n$. Ονομάζουμε όπως πριν τα V_1, V_2.

- Άρα $A^\top A V_1 = V_1 S^2$ όπου $S = \mathrm{diag}(\sigma_1, ..., \sigma_r)$
- άρα $V_1^\top A^\top A V_1 = V_1^\top V_1 S^2 = S^2$
- άρα $S^{-1} V_1^\top A^\top A V_1 S^{-1} = I$ (*).
- Επίσης $A^\top A V_2 = V_2 0 = 0$ άρα $V_2^\top A^\top A V_2 = 0$ άρα $A V_2 = 0$.
- Ορίζουμε $U_1 = A V_1 S^{-1}$. Τότε από την (*) φαίνεται ότι $U_1^\top U_1 = I$. Άρα οι στήλες του U_1 είναι ΟΚ.

Απόδειξη ύπαρξης του SVD

Δίνεται $A \in \mathbb{R}^{m\times n}$.

- Εξετάζουμε το μητρώο $A^\top A$: Είναι συμμετρικό θετικά ημιορισμένο,
- επομένως υπάρχουν n μη αρνητικές ιδιοτιμές

$$\sigma_1^2 \geq \cdots, \geq \sigma_r^2 > \sigma_{r+1}^2 = \cdots = \sigma_n^2$$

  και ΟΚ ιδιοδιανύσματα $v_1, ..., v_r, v_{r+1}, \ldots, v_n$. Ονομάζουμε όπως πριν τα V_1, V_2.
- Άρα $A^\top A V_1 = V_1 S^2$ όπου $S = \mathrm{diag}(\sigma_1, ..., \sigma_r)$
- άρα $V_1^\top A^\top A V_1 = V_1^\top V_1 S^2 = S^2$
- άρα $S^{-1} V_1^\top A^\top A V_1 S^{-1} = I$ (*).
- Επίσης $A^\top A V_2 = V_2 0 = 0$ άρα $V_2^\top A^\top A V_2 = 0$ άρα $A V_2 = 0$.
- Ορίζουμε $U_1 = A V_1 S^{-1}$. Τότε από την (*) φαίνεται ότι $U_1^\top U_1 = I$. Άρα οι στήλες του U_1 είναι ΟΚ.
- Επιλέγουμε U_2 ώστε το $[U_1, U_2]$ να είναι ορθογώνιο.

Απόδειξη ύπαρξης του SVD

Δίνεται $A \in \mathbb{R}^{m \times n}$.

- Εξετάζουμε το μητρώο $A^\top A$: Είναι συμμετρικό θετικά ημιορισμένο,
- επομένως υπάρχουν n μη αρνητικές ιδιοτιμές

$$\sigma_1^2 \geq \cdots, \geq \sigma_r^2 > \sigma_{r+1}^2 = \cdots = \sigma_n^2$$

  και ΟΚ ιδιοδιανύσματα $v_1, ..., v_r, v_{r+1}, \ldots, v_n$. Ονομάζουμε όπως πριν τα V_1, V_2.
- Άρα $A^\top A V_1 = V_1 S^2$ όπου $S = \mathrm{diag}(\sigma_1, ..., \sigma_r)$
- άρα $V_1^\top A^\top A V_1 = V_1^\top V_1 S^2 = S^2$
- άρα $S^{-1} V_1^\top A^\top A V_1 S^{-1} = I$ (*).
- Επίσης $A^\top A V_2 = V_2 0 = 0$ άρα $V_2^\top A^\top A V_2 = 0$ άρα $AV_2 = 0$.
- Ορίζουμε $U_1 = AV_1 S^{-1}$. Τότε από την (*) φαίνεται ότι $U_1^\top U_1 = I$. Άρα οι στήλες του U_1 είναι ΟΚ.
- Επιλέγουμε U_2 ώστε το $[U_1, U_2]$ να είναι ορθογώνιο.
- Επίσης $U_1 = AV_1 S^{-1}$ άρα $U_1^\top A V_1 = S$ και $U_2^\top A V_1 = U_2^\top U_1 S = 0$

Απόδειξη ύπαρξης του SVD

Δίνεται $A \in \mathbb{R}^{m\times n}$.

- Εξετάζουμε το μητρώο $A^\top A$: Είναι συμμετρικό θετικά ημιορισμένο,
- επομένως υπάρχουν n μη αρνητικές ιδιοτιμές

$$\sigma_1^2 \geq \cdots, \geq \sigma_r^2 > \sigma_{r+1}^2 = \cdots = \sigma_n^2$$

και ΟΚ ιδιοδιανύσματα $v_1, ..., v_r, v_{r+1}, \ldots, v_n$. Ονομάζουμε όπως πριν τα V_1, V_2.

- Άρα $A^\top A V_1 = V_1 S^2$ όπου $S = \mathrm{diag}(\sigma_1, ..., \sigma_r)$
- άρα $V_1^\top A^\top A V_1 = V_1^\top V_1 S^2 = S^2$
- άρα $S^{-1} V_1^\top A^\top A V_1 S^{-1} = I$ (*).
- Επίσης $A^\top A V_2 = V_2 0 = 0$ άρα $V_2^\top A^\top A V_2 = 0$ άρα $AV_2 = 0$.
- Ορίζουμε $U_1 = AV_1 S^{-1}$. Τότε από την (*) φαίνεται ότι $U_1^\top U_1 = I$. Άρα οι στήλες του U_1 είναι ΟΚ.
- Επιλέγουμε U_2 ώστε το $[U_1, U_2]$ να είναι ορθογώνιο.
- Επίσης $U_1 = AV_1 S^{-1}$ άρα $U_1^\top A V_1 = S$ και $U_2^\top A V_1 = U_2^\top U_1 S = 0$
- άρα

$$U^\top A V = (U_1, U_2)^\top A (V_1, V_2) = \begin{pmatrix} U_1^\top A V_1 & U_1^\top A V_2 \\ U_2^\top A V_1 & U_2^\top A V_2 \end{pmatrix} = \begin{pmatrix} S & 0_{r,n-r} \\ 0_{m-r,r} & 0_{m-r,n-r} \end{pmatrix}$$

και

Διάσπαση ιδιάζουσας τιμής (SVD)

Κάθε μητρώο $A \in \mathbb{R}^{m\times n}$ όπου $m \geq n$ μπορεί να παραγοντοποιηθεί ως γινόμενο δύο ορθογώνιων $U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{n\times n}$ και ενός πραγματικού, μη αρνητικού και διαγώνιου μητρώου $\Sigma \in \mathbb{R}^{m\times n}$ ως εξής

$$
\begin{aligned}
A &= U\Sigma V^\top = U\begin{pmatrix} \hat{\Sigma} \\ 0 \end{pmatrix} V^\top, \\
& \text{όπου } \; U^\top U = I_m, V^\top V = I_n, \hat{\Sigma} \in \mathbb{R}^{n\times n}, \\
& \text{ και } \hat{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) \geq 0.
\end{aligned}
$$

Επίσης, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ ενώ $\sigma_{r+1} = \cdots = \sigma_n = 0$. Μετά από τις απαραίτητες τροποποιήσεις, υπάρχει αντίστοιχη παραγοντοποίηση όταν $m \leq n$.

Από το SVD μαθαίνουμε σχεδόν τα πάντα για το μητρώο

Διάσπαση ιδιάζουσας τιμής (SVD)

Κάθε μητρώο $A \in \mathbb{R}^{m\times n}$ όπου $m \geq n$ μπορεί να παραγοντοποιηθεί ως γινόμενο δύο ορθογώνιων $U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{n\times n}$ και ενός πραγματικού, μη αρνητικού και διαγώνιου μητρώου $\Sigma \in \mathbb{R}^{m\times n}$ ως εξής

$$\begin{aligned} A &= U\Sigma V^\top = U\begin{pmatrix}\hat{\Sigma}\\ 0\end{pmatrix}V^\top, \\ &\text{όπου } U^\top U = I_m, V^\top V = I_n, \hat{\Sigma} \in \mathbb{R}^{n\times n}, \\ &\text{και } \hat{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) \geq 0. \end{aligned}$$

Επίσης, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ ενώ $\sigma_{r+1} = \cdots = \sigma_n = 0$. Μετά από τις απαραίτητες τροποποιήσεις, υπάρχει αντίστοιχη παραγοντοποίηση όταν $m \leq n$.

Από το SVD μαθαίνουμε σχεδόν τα πάντα για το μητρώο

Παραγοντοποιήσεις

Παραδείγματα[2] $A = WKH$ όπου $A \in \mathbb{R}^{m\times n}$. Θέσαμε $r = \mathrm{rank}(A)$ (προσέξτε, ορισμένες παραγοντοποιήσεις προϋποθέτουν τετραγωνικό A):

παραγοντοποίηση	W	K	H
$A = BC$ (τάξης)	$B \in \mathbb{R}^{m\times r}$, r στήλες οδηγών A		$C \in \mathbb{R}^{r\times n}$, μη μηδεν. γραμμές ΑΓΚΜ
$A = QR$ $A = Q_1 R_1$	$Q \in \mathbb{R}^{m\times m}$ ορθογώνιο $Q_1 \in \mathbb{R}^{m\times r}$, ΟΚ στήλες		$R \in \mathbb{R}^{m\times n}$ άνω τριγωνικό $R_1 \in \mathbb{R}^{r\times r}$ άνω τριγωνικό
$A = QTQ^*$ (Schur)	Q ορθομοναδιαίο	άνω τριγωνικό T	Q^*
$A = XJX^{-1}$ (ΑΠΙ)	X αντιστρέψιμο	μορφή Jordan J	X^{-1}
$A = X\Lambda X^{-1}$ διαγ/σιμο (ΑΠΙ)	X αντιστρέψιμο	διαγώνιο Λ	X^{-1}
$A = Q\Lambda Q^\top$ συμμετρικό (ΑΠΙ)	Q ορθογώνιο	διαγώνιο Λ	$Q^\top$

Πρόκληση: Να διαγωνιοποιήσουμε οποιοδήποτε μητρώο (ακόμα και μη τετραγωνικό) με ορθογώνιους μετασχηματισμούς.

[2] Δείτε τον εκτενή κατάλογο στο σύγγραμμα του Strang.

Παραγοντοποιήσεις

Παραδείγματα[2] $A = WKH$ όπου $A \in \mathbb{R}^{m\times n}$. Θέσαμε $r = \mathrm{rank}(A)$ (προσέξτε, ορισμένες παραγοντοποιήσεις προϋποθέτουν τετραγωνικό A):

παραγοντοποίηση	W	K	H
$A = BC$ (τάξης)	$B \in \mathbb{R}^{m\times r}$, r στήλες οδηγών A		$C \in \mathbb{R}^{r\times n}$, μη μηδεν. γραμμές ΑΓΚΜ
$A = QR$ $A = Q_1 R_1$	$Q \in \mathbb{R}^{m\times m}$ ορθογώνιο $Q_1 \in \mathbb{R}^{m\times r}$, ΟΚ στήλες		$R \in \mathbb{R}^{m\times n}$ άνω τριγωνικό $R_1 \in \mathbb{R}^{r\times r}$ άνω τριγωνικό
$A = QTQ^*$ (Schur)	Q ορθομοναδιαίο	άνω τριγωνικό T	Q^*
$A = XJX^{-1}$ (ΑΠΙ)	X αντιστρέψιμο	μορφή Jordan J	X^{-1}
$A = X\Lambda X^{-1}$ διαγ/σιμο (ΑΠΙ)	X αντιστρέψιμο	διαγώνιο Λ	X^{-1}
$A = Q\Lambda Q^\top$ συμμετρικό (ΑΠΙ)	Q ορθογώνιο	διαγώνιο Λ	$Q^\top$
$A = U\Sigma V^\top$ (SVD) $A = U_1\hat{\Sigma}V_1^\top$ (SVD)	$U \in \mathbb{R}^{m\times m}$ ορθογώνιο $U_1 \in \mathbb{R}^{m\times r}$ ΟΚ στήλες	διαγ. $\Sigma \geq 0 \in \mathbb{R}^{m\times n}$, με r θετικά στοιχεία διαγ. $\hat{\Sigma} \in \mathbb{R}^{r\times r}$, με r θετικά στοιχεία	$V \in \mathbb{R}^{n\times n}$ ορθογώνιο $V_1 \in \mathbb{R}^{n\times r}$ ΟΚ γραμμές

Πρόκληση: Να διαγωνιοποιήσουμε οποιοδήποτε μητρώο (ακόμα και μη τετραγωνικό) με ορθογώνιους μετασχηματισμούς.

[2] Δείτε τον εκτενή κατάλογο στο σύγγραμμα του Strang.

Παραγοντοποιήσεις

Παραδείγματα[2] $A = WKH$ όπου $A \in \mathbb{R}^{m\times n}$. Θέσαμε $r = \mathrm{rank}(A)$ (προσέξτε, ορισμένες παραγοντοποιήσεις προϋποθέτουν τετραγωνικό A):

παραγοντοποίηση	W	K	H
$A = BC$ (τάξης)	$B \in \mathbb{R}^{m\times r}$, r στήλες οδηγών A		$C \in \mathbb{R}^{r\times n}$, μη μηδεν. γραμμές ΑΓΚΜ
$A = QR$ $A = Q_1 R_1$	$Q \in \mathbb{R}^{m\times m}$ ορθογώνιο $Q_1 \in \mathbb{R}^{m\times r}$, ΟΚ στήλες		$R \in \mathbb{R}^{m\times n}$ άνω τριγωνικό $R_1 \in \mathbb{R}^{r\times r}$ άνω τριγωνικό
$A = QTQ^*$ (Schur)	Q ορθομοναδιαίο	άνω τριγωνικό T	Q^*
$A = XJX^{-1}$ (ΑΠΙ)	X αντιστρέψιμο	μορφή Jordan J	X^{-1}
$A = X\Lambda X^{-1}$ διαγ/σιμο (ΑΠΙ)	X αντιστρέψιμο	διαγώνιο Λ	X^{-1}
$A = Q\Lambda Q^\top$ συμμετρικό (ΑΠΙ)	Q ορθογώνιο	διαγώνιο Λ	$Q^\top$
$A = U\Sigma V^\top$ (SVD) $A = U_1\hat{\Sigma}V_1^\top$ (SVD)	$U \in \mathbb{R}^{m\times m}$ ορθογώνιο $U_1 \in \mathbb{R}^{m\times r}$ ΟΚ στήλες	διαγ. $\Sigma \geq 0 \in \mathbb{R}^{m\times n}$, με r θετικά στοιχεία διαγ. $\hat{\Sigma} \in \mathbb{R}^{r\times r}$, με r θετικά στοιχεία	$V \in \mathbb{R}^{n\times n}$ ορθογώνιο $V_1 \in \mathbb{R}^{n\times r}$ ΟΚ γραμμές
$A = UP$ (polar)	$U \in \mathbb{R}^{m\times m}$ unitary	διαγ. $P \in \mathbb{R}^{m\times n}$, ερμιτιανό ημιορισμένο	

Πρόκληση: Να διαγωνιοποιήσουμε οποιοδήποτε μητρώο (ακόμα και μη τετραγωνικό) με ορθογώνιους μετασχηματισμούς.

[2] Δείτε τον εκτενή κατάλογο στο σύγγραμμα του Strang.

Διάσπαση ιδιάζουσας τιμής (SVD)

Κάθε μητρώο $A \in \mathbb{R}^{m\times n}$ όπου $m \geq n$ μπορεί να παραγοντοποιηθεί ως γινόμενο δύο ορθογώνιων $U \in \mathbb{R}^{m\times m}, V \in \mathbb{R}^{n\times n}$ και ενός πραγματικού, μη αρνητικού και διαγώνιου μητρώου $\Sigma \in \mathbb{R}^{m\times n}$ ως εξής

$$
\begin{aligned}
A &= U\Sigma V^\top = U\begin{pmatrix}\hat{\Sigma}\\ 0\end{pmatrix}V^\top,\\
& \text{όπου } \; U^\top U = I_m, V^\top V = I_n, \hat{\Sigma} \in \mathbb{R}^{n\times n},\\
& \text{ και } \hat{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_n) \geq 0.
\end{aligned}
$$

Επίσης, $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r > 0$ ενώ $\sigma_{r+1} = \cdots = \sigma_n = 0$. Μετά από τις απαραίτητες τροποποιήσεις, υπάρχει αντίστοιχη παραγοντοποίηση όταν $m \leq n$.

Από το SVD μαθαίνουμε σχεδόν τα πάντα για το μητρώο

Ορολογία και παρατηρήσεις

Έστω ότι $A = U\Sigma V^\top$ είναι η SVD του A.

- Τα διαγώνια στοιχεία του A λέγονται ιδιάζουσες τιμές του A. Θεωρούμε ότι $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_{\min(m,n)}$.
- Οι στήλες του U ονομάζονται αριστερά **ιδιάζοντα** διανύσματα. Προσέξτε ότι είναι τα ορθοκανονικά ιδιοδιανυσματα του $AA^\top$.
- Οι στήλες του V ονομάζονται δεξιά **ιδιάζοντα** διανύσματα. Προσέξτε ότι είναι τα ορθοκανονικά ιδιοδιανυσματα του $A^\top A$.

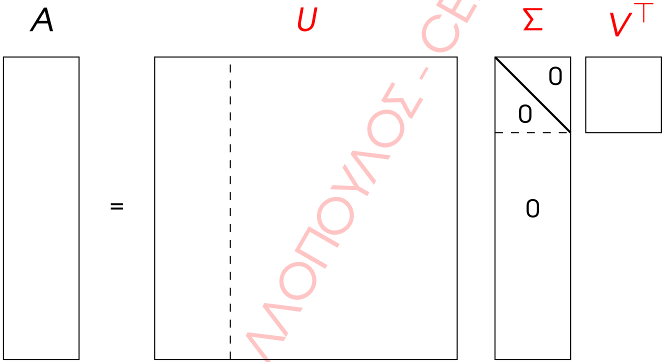
τάξη του A ο αριθμός των μη μηδενικών στοιχείων της διαγωνίου του Σ.

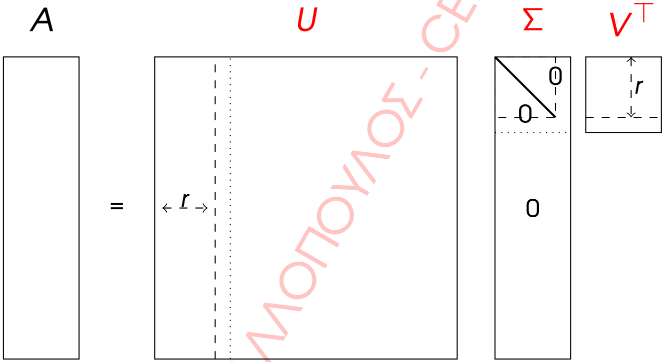
ΟΚ βάση για $\mathrm{range}(A)$ οι στήλες του $U_1 = [u_1, \ldots, u_r]$

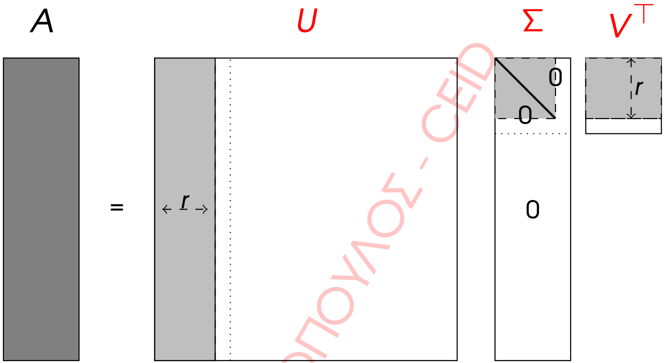
ΟΚ βάση για $\mathrm{null}(A)$ οι στήλες του $V_2 = [v_{r+1}, \ldots, v_n]$

ΟΚ βάση για $\mathrm{range}(A^\top)$ οι στήλες του $V_1 = [v_1, \ldots, v_r]$

ΟΚ βάση για $\mathrm{null}(A^\top)$ οι στήλες του $U_2 = [u_{r+1}, \ldots, u_m]$

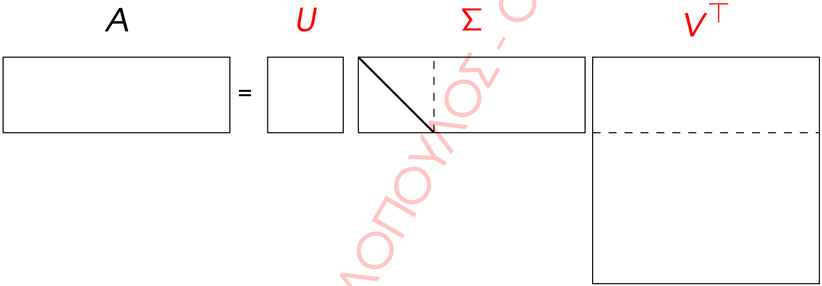

A
$=$
U
Σ
$V^\top$
0
0
0


A
=
U
r
Σ
0
0
0
$V^\top$
r



$U_1 = U_{:,1:r}$	$U_2 = U_{:,r+1:m}$	$V_1 = V_{:,1:r}$	$V_2 = V_{:,r+1:n}$
$\mathrm{range}(A)$	$\mathrm{null}(A^\top)$	$\mathrm{range}(A^\top)$	$\mathrm{null}(A)$

$$\boxed{A = U_1 \hat{\Sigma} V_1^\top}$$


A
=
U
Σ
$V^\top$

Strang, σελ. 449

Θα σας πω άμεσα τη γνώμη μου: Η SVD είναι το αποκορυφωμα αυτού του μαθήματος της γραμμικής άλγεβρας. Τη σκέφτομαι ως το τελικό βήμα στο Θεμελιώδες Θεώρημα. Αρχικά εμφανίζονται οι διαστάσεις των τεσσάρων υποχώρων. Στη συνέχεια η ορθογωνιότητά τους. Έπειτα οι ορθοκανονικές βάσεις οι οποίες διαγωνιοποιούν τον A. Όλα εμπεριέχονται στον τύπο $A = U\Sigma V^\top$.

Ανάπτυγμα σε ιδιάζουσες τριπλέτες

Αν το μητρώο είναι τάξης r και θέσουμε $\hat{\Sigma}_{r,r}$ το $r \times r$ διαγώνιο μητρώο $\mathrm{diag}[\sigma_1, \cdots, \sigma_r]$ (όλα μη μηδενικά) τότε

$$\begin{aligned} A &= [U_1, U_2] \begin{pmatrix} \hat{\Sigma}_{r,r} & 0_{r,n-r} \\ 0_{m-r} & 0_{m-r,n-r} \end{pmatrix} \begin{pmatrix} V_1^\top \\ V_2^\top \end{pmatrix} \\ &= U_1 \Sigma_{r,r} V_1^\top \\ &= \sigma_1 u_1 v_1^\top + \sigma_2 u_2 v_2^\top + \cdots + \sigma_r u_r v_r^\top \end{aligned}$$

Αποκαλύπτεται ότι το μητρώο τάξης r είναι άθροισμα r ανεξάρτητων μητρώων πρώτης τάξης. Λόγω της διάταξης των σ_j και της ορθογωνιότητας των u_j και v_j αποδεικνύεται ότι οι παρακάτω προσεγγίσεις είναι «καλές» (η ποιότητα αυξάνει καθώς μεγαλώνει το $k = 1, \ldots, r$):

$$\begin{aligned} A &\approx \sigma_1 u_1 v_1^\top \\ \vdots &\;\;\vdots\;\; \vdots \\ A &\approx \sigma_1 u_1 v_1^\top + \sigma_k u_k v_k^\top \end{aligned}$$

Ιδιαίτερα χρήσιμο για να προσεγγίζουμε (συμπιέζουμε) μητρώα και ό,τι αυτά αναπαριστούν.

Υπολογισμός του SVD

μέσω ιδιοδιασπάσεων

Για μικρά προβλήματα, αξιοποιούμε τη σχέση του SVD του A με τις ιδιοτιμές και ιδιοδιανύσματα των $A^\top A$ ή και $AA^\top$.
Αν $A = U\Sigma V^\top \in \mathbb{R}^{m\times n}$, τότε,

$$A^\top A \;=\; V\Sigma^\top \Sigma V^\top,\;\; AA^\top = U\Sigma\Sigma^\top U^\top \quad \text{καθώς } U^\top U = I_m, V^\top V = I_n$$

επομένως αν θεωρήσουμε ότι $m \geq n$ και ονομάσουμε $\mu_1, \ldots, \mu_n$ τις ιδιοτιμές του $A^\top A$ (τις υπολογίζουμε με κάποιον τρόπο), τότε

1. $\sigma_i = \sqrt{\mu_i}$
2. Τα δεξιά ιδιάζοντα διανύσματα του A είναι τα ιδιοδιανύσματα του $A^\top A$.
3. Ισχύει επίσης ότι τα αριστερά ιδιάζοντα διανύσματα του A είναι τα ιδιοδιανύσματα του $A^\top A$.
4. Για τις μη μηδενικές ιδιάζουσες τιμές ισχύει ότι

$$\frac{1}{\sigma_i} A v_i = u_i$$

και μπορούμε να υπολογίσουμε τα αριστερά ιδιάζοντα διανύσματα με οικονομικό τρόπο από τη σχέση.

Προσοχή: Το θέμα του υπολογισμού SVD όπως και το αλγεβρικό πρόβλημα ιδιοτιμών είναι πολύ ευρύ. Η μέθοδος που αναφέραμε είναι κυρίως για χρήση «στο χαρτί».

Προσέγγιση μειωμένης τάξης με SVD

Θεώρημα Eckart & Schmidt

Έστω $A \in \mathbb{R}^{m\times n}$ τέτοιο ώστε $A = U\Sigma V^\top$. Τότε το μητρώο τάξης $k < \min(m,n)$, έστω B, που προσεγγίζει το A με τρόπο ώστε να ελαχιστοποιείται η νόρμα Frobenius $\|A-B\|_F$ είναι το μητρώο

$$B = U^{(k)} S_k (V^{(k)})^\top$$

όπου

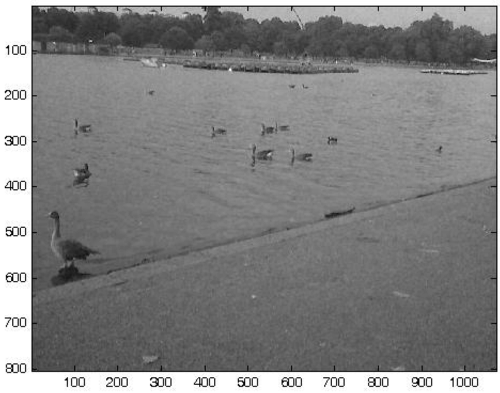
$$\begin{aligned} U^{(k)} &= [u_1, \dots, u_k], \\ V^{(k)} &= [v_1, \dots, v_k], \\ S^{(k)} &= \mathrm{diag}(\sigma_1, \dots, \sigma_k) \end{aligned}$$

Τότε

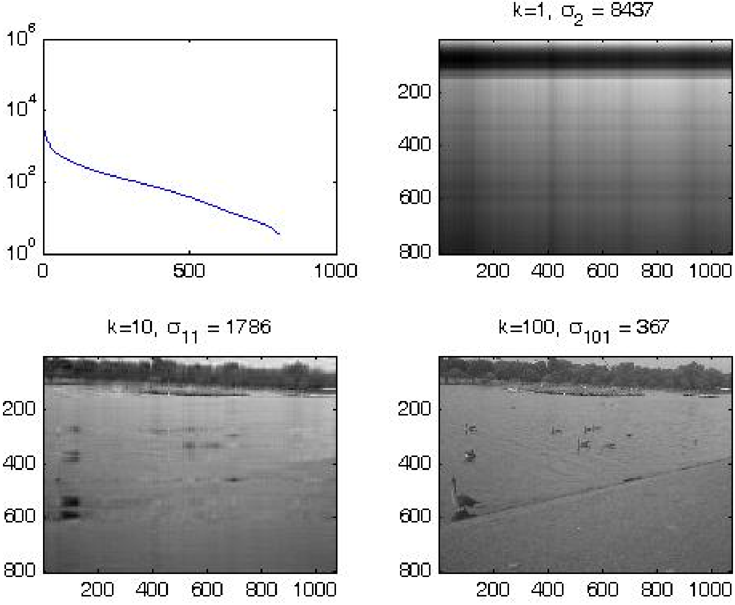
$$\|A-B\|_F^2 = \sigma_{k+1}^2 + \cdots + \sigma_r^2$$

Παράδειγμα προσέγγισης

Η παρακάτω εικόνα κωδικοποιείται με μητρώο $X \in \mathbb{R}^{804\times 1072}$ του οποίου μπορούμε να υπολογίσουμε το SVD.



Προσεγγίσεις $U^{(k)} S_k (V^{(k)})^\top$ για $k = 1, 10, 100$. Το πρώτο σχήμα δείχνει τη διακύμανση των 804 ιδιαζουσών τιμών.



Αντιστροφή με SVD και η έννοια του ψευδοαντιστρόφου

Όταν το A είναι τετραγωνικό και αντιστρέψιμο

$$\begin{aligned} A^{-1} &= (U\Sigma V^\top)^{-1} = (V^\top)^{-1}\Sigma^{-1}U^{-1} \\ &= V\Sigma^{-1}U^\top \\ &= \frac{1}{\sigma_1}v_1u_1^\top + \frac{1}{\sigma_2}v_2u_2^\top + \cdots + \frac{1}{\sigma_n}v_nu_n^\top \end{aligned}$$

<u>Μπορούμε να επεκτείνουμε την έννοια του αντιστρόφου;</u>

Αντιστροφή με SVD και η έννοια του ψευδοαντιστρόφου

Όταν το A είναι τετραγωνικό και αντιστρέψιμο

$$\begin{aligned} A^{-1} &= (U\Sigma V^\top)^{-1} = (V^\top)^{-1}\Sigma^{-1}U^{-1} \\ &= V\Sigma^{-1}U^\top \\ &= \frac{1}{\sigma_1}v_1u_1^\top + \frac{1}{\sigma_2}v_2u_2^\top + \cdots + \frac{1}{\sigma_n}v_nu_n^\top \end{aligned}$$

Μπορούμε να επεκτείνουμε την έννοια του αντιστρόφου; Αν το A δεν είναι αντιστρέψιμο, τότε θα είναι μειωμένης τάξης, δηλ. $r < n$. Επομένως στις παραπάνω εκφράσεις, θα υπάρξει αστοχία τους μηδενικούς όρους $\sigma_{r+1}, \ldots, \sigma_n$ του $\hat{\Sigma}$:

$$\Sigma^{-1} = \begin{pmatrix} \frac{1}{\sigma_1} & & & & \\ & \ddots & & & \\ & & \frac{1}{\sigma_r} & & \\ & & & ? & \\ & & & & \ddots \end{pmatrix}$$

Ψευδοαντίστροφο μητρώο

Επεκτείνουμε τον ορισμό του αντιστρόφου για το $\hat{\Sigma}$ θέτοντας (προσέξτε το διαφορετικό σύμβολο-υπερδείκτη):

$$\Sigma^{\dagger} = \begin{pmatrix} \frac{1}{\sigma_1} & & & & \\ & \ddots & & & \\ & & \frac{1}{\sigma_r} & & \\ & & & 0 & \\ & & & & \ddots \end{pmatrix}$$

Τότε $A^{\dagger} = \frac{1}{\sigma_1} v_1 u_1^{\top} + \frac{1}{\sigma_2} v_2 u_2^{\top} + \cdots + \frac{1}{\sigma_r} v_r u_r^{\top}$, άρα

$$A^{\dagger} = V \Sigma^{\dagger} U^{\top}$$

Το $A^{\dagger}$ ονομάζεται **ψευδοαντίστροφο** του A. Όπως ορίστηκε, το ψευδοαντίστροφο μητρώο είναι μοναδικό για οποιοδήποτε μητρώο $A \in \mathbb{R}^{m \times n}$.

Επιλεγμένες ιδιότητες ψευδοαντιστρόφου

- $AA^\dagger A = A$, $A^\dagger AA^\dagger = A^\dagger$
- $(AA^\dagger)^\top = AA^\dagger$ και $(A^\dagger A)^\top = A^\dagger A$.
- αν A είναι πλήρους τάξης στηλών, $A^\dagger = (A^\top A)^{-1}A^\top$.
- .... και $A^\dagger A = I$ (αριστερό αντίστροφο).
- αν είναι πλήρους τάξης γραμμών, $A^\dagger = A^\top(AA^\top)^{-1}$,
- ... και $AA^\dagger = I$ (δεξιό αντίστροφο).
- Η επίλυση ελαχίστων τετραγώνων του $Ax = b$ είναι $x_{\mathrm{LS}} = A^\dagger b$.

Παράδειγμα

Να υπολογίσετε το ψευδοαντίστροφο του

$$A = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \\ -2 & -1 & -1 \\ 4 & 2 & 2 \end{pmatrix}$$

Παράδειγμα

Να υπολογίσετε το ψευδοαντίστροφο του

$$A = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix} \begin{pmatrix} 2 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 2 & 1 & 1 \\ -2 & -1 & -1 \\ 4 & 2 & 2 \end{pmatrix}$$

Προσέξτε ότι $A = ab^\top$ οπότε $\sigma_1 > 0, \sigma_2 = \sigma_3 = 0$.

$$A = \|a\|_2 \|b\|_2 uv^\top, \text{ όπου } u_1 = \frac{a}{\|a\|_2}, v_1 = \frac{b}{\|b\|_2}$$

επομένως $\|u_1\|_2 = \|v_1\|_2 = 1$. (Στην περίπτωσή μας, $\|a\|_2 = \|b\|_2 = \sqrt{6}$).
Επομένως ισχύει ότι η SVD του

$$A = \underbrace{(\|a\|_2 \|b\|_2)}_{\sigma_1} u_1 v_1^\top.$$

Άρα

$$A^\dagger = \frac{1}{6} v_1 u_1^\top = \frac{1}{6} \frac{b}{\|b\|_2} \frac{a^\top}{\|a\|_2} = \frac{1}{36} ba^\top = \frac{1}{36} \begin{pmatrix} 2 & -2 & 4 \\ 1 & -1 & 2 \\ 1 & -1 & 2 \end{pmatrix}$$

Παράδειγμα

Να υπολογίσετε το ψευδοαντίστροφο του

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

Όπως και πριν: $A = ab^\top$ οπότε $\sigma_1 > 0, \sigma_2 = \sigma_3 = 0$.

$$A = \|a\|_2 \|b\|_2 uv^\top, \text{ όπου } u_1 = \frac{a}{\|a\|_2}, v_1 = \frac{b}{\|b\|_2}$$

επομένως $\|u_1\|_2 = \|v_1\|_2 = 1$. (Στην περίπτωσή μας, $\|a\|_2 = \|b\|_2 = \sqrt{2}$).
Επομένως ισχύει ότι η SVD του

$$A = \underbrace{(\|a\|_2 \|b\|_2)}_{\sigma_1} u_1 v_1^\top.$$

Άρα

$$A^\dagger = \frac{1}{2} v_1 u_1^\top = \frac{1}{2} \frac{b}{\|b\|_2} \frac{a^\top}{\|a\|_2} = \frac{1}{4} ba^\top = \begin{pmatrix} \frac{1}{4} & -\frac{1}{4} & 0 \\ \frac{1}{4} & -\frac{1}{4} & 0 \end{pmatrix}$$

Ψευδοαντίστροφο

Moore–Penrose pseudoinverse

From Wikipedia, the free encyclopedia

In mathematics, and in particular linear algebra, a **pseudoinverse** A^+ of a matrix A is a generalization of the inverse matrix.[1] The most widely known type of matrix pseudoinverse is the **Moore–Penrose pseudoinverse**, which was independently described by E. H. Moore[2] in 1920, Arne Bjerhammar[3] in 1951 and Roger Penrose[4] in 1955. Earlier, Fredholm had introduced the concept of a pseudoinverse of integral operators in 1903. When referring to a matrix, the term pseudoinverse, without further specification, is often used to indicate the Moore–Penrose pseudoinverse. The term generalized inverse is sometimes used as a synonym for pseudoinverse.

A common use of the Moore–Penrose pseudoinverse (hereafter, just pseudoinverse) is to compute a 'best fit' (least squares) solution to a system of linear equations that lacks a unique solution (see below under § Applications). Another use is to find the minimum (Euclidean) norm solution to a system of linear equations with multiple solutions. The pseudoinverse facilitates the statement and proof of results in linear algebra.

The pseudoinverse is defined and unique for all matrices whose entries are real or complex numbers. It can be computed using the singular value decomposition.