ΓΡΑΜΜΙΚΗ ΑΛΓΕΒΡΑ

Ε. Γαλλόπουλος & Ε. Στεφανόπουλος

CEID

Διάλεξη 9

24 Απριλίου 2018

Έννοιες, λέξεις κλειδιά, υπενθύμιση I

Υπενθύμιση της 8ης δ.

- Ορθοκανικοποίηση διανυσμάτων και διαδικασία Gram-Schmidt.
- Παραγοντοποίηση QR μέσω Gram-Schmidt (απλή αναφορά)
- Θεώρημα προβολής
- Γενική θεώρηση προβλημάτων ελαχιστοποίησης
- Προβλήματα και προσεγγίσεις ελαχίστων τετραγώνων

Έννοιες, λέξεις κλειδιά, υπενθύμιση II

Σήμερα θα συζητήσουμε (όροι 9ης δ.):

- Εσωτερικό γινόμενο, νόρμα και ελάχιστα τετράγωνα σε διανυσματικούς χώρους συναρτήσεων.
- Gram-Schmidt σε χώρους πολυωνύμων και βέλτιστη προσέγγιση. συναρτήσεων.
- Gram-Schmidt σε χώρους πολυωνύμων και βέλτιστη προσέγγιση. συναρτήσεων.
- Ενδιαφέρουσες βαθμωτές συναρτήσεις μητρώων: Ίχνος και Ορίζουσα
- αξιωματικός ορισμός ορίζουσας
- τρόποι υπολογισμού ορίζουσας
- ιδιότητες ορίζουσας
- επίλυση συστήματος και αντιστροφή μητρώου μέσω οριζουσών.
- τύπος Sylvester
- επίλυση συστήματος και αντιστροφή μητρώου μέσω οριζουσών.

Υπό συζήτηση ενότητες

Προσέγγιση συναρτήσεων με ελάχιστα τετράγωνα: Εισαγωγή I

- Γνωρίζουμε ότι πολλές κατηγορίες συναρτήσεων αποτελούν δ.χ.
- Παραδείγματα: α) $\mathbb{P}_n$: τα πολυώνυμα βαθμού $\leq n$. (ο δ.χ. έχει διάσταση $n+1$). β) $C[-1,1]$: χώρος συνεχών συναρτήσεων στο κλειστό διάστημα $[-1,1]$. γ) Τα πολυώνυμα οποιουδήποτε βαθμού (οι δ.χ. είναι απειροδιάστατοι).
- Για ολοκληρώσιμες συναρτήσεις f, g στο διάστημα $[a, b]$ και θετική συνάρτηση w, το ορισμένο ολοκλήρωμα

$$\langle f, g \rangle := \int_a^b f(t)g(t)w(t)dt$$

  ορίζει εσωτερικό γινόμενο.
- Επομένως μπορούμε να ορίσουμε και καθετότητα συναρτήσεων. Για παράδειγμα, αν $w(x) = 1$, οι συναρτήσεις $\{1, x, \frac{1}{2}(3x^2 - 1)\}$ είναι κάθετες μεταξύ τους. Επίσης είναι γραμμικά ανεξάρτητες και αποτελούν βάση για τον $\mathbb{P}_2[-1,1]$ που είναι υπόχωρος του $C[-1,1]$.
- Μπορούμε να εφαρμόσουμε τη διαδικασία Gram-Schmidt σε πολυώνυμα και να κατασκευάσουμε Ορθογώνια Πολυώνυμα (σε κάποιο διάστημα και ως προς δεδομένη συνάρτηση w.)

Προσέγγιση συναρτήσεων με ελάχιστα τετράγωνα: Εισαγωγή ΙΙ

Ερώτημα Δίνεται συνάρτηση f από κάποιο χώρο συναρτήσεων και θέλουμε να βρούμε βέλτιστη προσέγγισή της ως προς το κριτήριο των ελαχίστων τετραγώνων από κάποιον επιλεγμένο γραμμικό υπόχωρο, π.χ. τα πολυώνυμα, με βάση το παραπάνω εσωτερικό γινόμενο.

Επίλυση όπως πριν, θα αξιοποιήσουμε το ότι η διαφορά της υπό προσέγγιση συνάρτησης από τη βέλτιστη προσέγγιση θα είναι κάθετη σε όλες τις συναρτήσεις του υποχώρου απ' όπου «χτίζουμε» την προσέγγιση.

Ιδέα

Από το θεώρημα προβολής, αν αφαιρέσουμε από τη συνάρτηση το γραμμικό συνδυασμό που παράγει τη συνάρτηση βέλτιστης προσέγγισης, το κατάλοιπο θα είναι κάθετο σε όλες τις συναρτήσεις ϕ_j, επομένως $\langle f - \sum_{j=1}^{n} \gamma_j \phi_j, \phi_i \rangle = 0$ για $i = 1, \ldots, n$. Επομένως

$$\begin{pmatrix} \langle f, \phi_1 \rangle \\ \vdots \\ \langle f, \phi_n \rangle \end{pmatrix} = \underbrace{\begin{pmatrix} \langle \phi_1, \phi_1 \rangle & \cdots & \langle \phi_1, \phi_n \rangle \\ \vdots & & \vdots \\ \langle \phi_n, \phi_1 \rangle & \cdots & \langle \phi_n, \phi_n \rangle \end{pmatrix}}_{\text{Μητρώο Gram}} \begin{pmatrix} \gamma_1 \\ \vdots \\ \gamma_n \end{pmatrix}.$$

Παράδειγμα

Να υπολογιστεί η βέλτιστη προσέγγιση της συνάρτησης x^2 από τον υπόχωρο $\mathrm{span}\langle 1, x\rangle$ με εσωτερικό γινόμενο $\langle f, g\rangle := \int_{-1}^{1} f(x)g(x)dx$.
Έστω $\phi(x) = \gamma_1 + \gamma_2 x$ η ζητούμενη προσέγγιση, τότε αν $r(x) = x^2 - \phi(x)$ θα πρέπει από το θεώρημα ορθογωνιότητας $\langle r, 1\rangle = \langle r, x\rangle = 0$. Επομένως

$$\begin{aligned}
\langle x^2, 1\rangle &= \int_{-1}^{1} x^2 dx = \frac{2}{3} \\
\langle x^2, x\rangle &= \int_{-1}^{1} x^3 dx = 0 \\
\langle 1, 1\rangle &= 2, \langle 1, x\rangle = 0 \\
\langle x, 1\rangle &= 0, \langle x, x\rangle = \frac{2}{3}
\end{aligned}$$

$$\begin{pmatrix} \frac{2}{3} \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 \\ 0 & \frac{2}{3} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \end{pmatrix}. \Rightarrow \gamma_1 = \frac{1}{3}, \gamma_2 = 0.$$

΄Αρα η βέλτιστη προσέγγιση είναι $\phi(x) = \frac{1}{3}$.

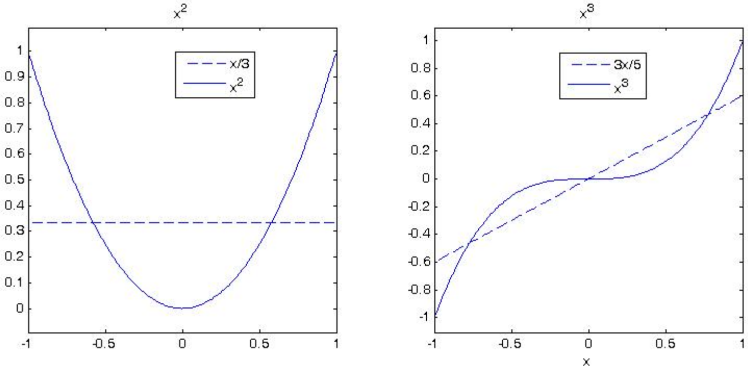
Προσέξτε επίσης ότι $\langle 1, x \rangle = 0$ δηλ, οι συναρτήσεις $\phi_1(x) = 1$ και $\phi_2(x) = x$ είναι κάθετες ως προς το παραπάνω εσωτερικό γινόμενο.

Προσοχή: Με τον ίδιο τρόπο αλλά λύνοντας το σύστημα

$$\begin{pmatrix} 0 \\ \frac{2}{5} \\ 0 \end{pmatrix} = \begin{pmatrix} 2 & 0 & \frac{2}{3} \\ 0 & \frac{2}{3} & 0 \\ \frac{2}{3} & 0 & \frac{2}{5} \end{pmatrix} \begin{pmatrix} \gamma_1 \\ \gamma_2 \\ \gamma_3 \end{pmatrix}. \Rightarrow \gamma_1 = 0, \gamma_1 = \frac{3}{5}, \gamma_3 = 0.$$

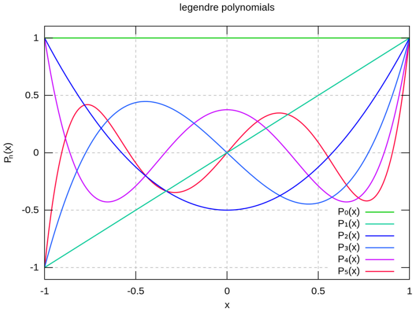
προκύπτει ότι η βέλτιστη προσέγγιση του x^3 από το χώρο πολυωνύμων $\langle 1, x, x^2 \rangle$ ως προς παραπάνω εσωτερικό γινόμενο είναι $\phi(x) = \frac{3}{5}x$.

Οπτικοποίηση των βέλτιστων προσεγγίσεων



Ορθογώνια πολυώνυμα Legendre

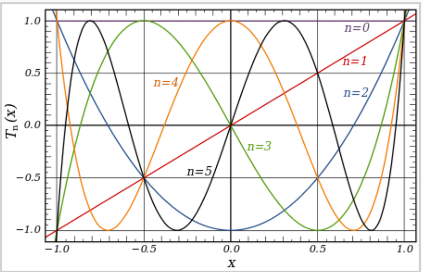
n	$P_n(x)$
0	1
1	x
2	$\frac{1}{2}(3x^2-1)$
3	$\frac{1}{2}(5x^3-3x)$
4	$\frac{1}{8}(35x^4-30x^2+3)$
5	$\frac{1}{8}(63x^5-70x^3+15x)$
6	$\frac{1}{16}(231x^6-315x^4+105x^2-5)$
7	$\frac{1}{16}(429x^7-693x^5+315x^3-35x)$
8	$\frac{1}{128}(6435x^8-12012x^6+6930x^4-1260x^2+35)$
9	$\frac{1}{128}(12155x^9-25740x^7+18018x^5-4620x^3+315x)$
10	$\frac{1}{256}(46189x^{10}-109395x^8+90090x^6-30030x^4+3465x^2-63)$



- Παράγονται με διαδικασία Gram-Schmidt επί των $\{1, x, x^2, \ldots\}$ με το εσωτερικό γινόμενο $\int_{-1}^{1} f(x)g(x)\mathrm{d}x$
- Ισχύει ότι $\int_{-1}^{1} P_n(x)P_m(x)dx = \frac{2}{2n+1}\delta_{nm}$ (δέλτα Kronecker)
- Υπάρχουν πολλές οικογένειες ορθογωνίων πολυωνύμων (π.χ. Chebyshev, Gegenbauer, Legendre, Laguerre, Hermite...).
- Η βασική διαφορά τους αφορά στη συνάρτηση βάρους.

Ορθογώνια πολυώνυμα Chebyshev

$T_0(x) = 1$
$T_1(x) = x$
$T_2(x) = 2x^2 - 1$
$T_3(x) = 4x^3 - 3x$
$T_4(x) = 8x^4 - 8x^2 + 1$
$T_5(x) = 16x^5 - 20x^3 + 5x$
$T_6(x) = 32x^6 - 48x^4 + 18x^2 - 1$
$T_7(x) = 64x^7 - 112x^5 + 56x^3 - 7x$
$T_8(x) = 128x^8 - 256x^6 + 160x^4 - 32x^2 + 1$
$T_9(x) = 256x^9 - 576x^7 + 432x^5 - 120x^3 + 9x.$
$T_{10}(x) = 512x^{10} - 1280x^8 + 1120x^6 - 400x^4 + 50x^2 - 1.$
$T_{11}(x) = 1024x^{11} - 2816x^9 + 2816x^7 - 1232x^5 + 220x^3 - 11x.$



The first few Chebyshev polynomials of the first kind in the domain $-1 < x < 1$: The flat T_0, T_1, T_2, T_3, T_4 and T_5.

- Παράγονται με διαδικασία Gram-Schmidt επί των $\{1, x, x^2, \dots\}$ με το εσωτερικό γινόμενο $\int_{-1}^{1} f(x)g(x)\frac{1}{\sqrt{1-x^2}}\mathrm{d}x$
- Ισχύει ότι $\int_{-1}^{1} T_n(x)T_m(x)dx = 0$ αν $m \neq n$.
- Ένα σημαντικό στοιχείο των ορθογωνίων πολυωνύμων που αξίζει να αναφερθεί (εκτός ύλης) είναι ότι μπορούν να παραχθούν με "κοντή" αναδρομική σχέση (τριών όρων), π.χ. τα Chebyshev: $T_{n+1}(x) = 2xT_n(x) - T_{n-1}(x)$ όπου $T_0(x) = 1, T_1(x) = x$.

Ειδικές τιμές μητρώων

Παρόλο που ένα μητρώο $A \in \mathbb{R}^{m\times n}$ περιέχει $m \times n$ αριθμούς, σε αυτό αντιστοιχούν επίσης ορισμένες ειδικές τιμές συνδεδεμένες με αυτό. Αυτές είναι βαθμωτές συναρτήσεις μητρώου ($f : \mathbb{R}^{n\times n}\mathbb{R}$), και αποκαλύπτουν σημαντικές πληροφορίες για το μητρώο.

είδος	*συμβολισμός*	*παρατηρήσεις*
διαστάσεις	m, n	αριθμ. γραμμών και αριθμ. στηλών
αραιότητα	$\mathrm{nnz}(A)$	πλήθος μη μηδενικών
τάξη	$\mathrm{rank}(A)$	μέγιστος αριθμ. γ.α. γραμμών (ή στηλών)
νόρμα	$\|A\|$	μετρική «βάρους» , π.χ. $\sup_{x\neq 0} \frac{\|Ax\|}{\|x\|}$
ίχνος	$\mathrm{trace}(A)$	$(m = n)$???
ορίζουσα	$\det(A)$	$(m = n)$???
ιδιοτιμές	???	$(m = n)$???
ιδιάζουσες τιμές	???	???

Ίχνος

τετραγωνικού μητρώου

- Χαρακτηριστική τιμή για κάθε τετραγωνικό μητρώο ίση με το άθροισμα των διαγώνιων στοιχείων του.
- Βαθμωτή συνάρτηση

$$\mathrm{trace} : \mathbb{R}^{n\times n} \longrightarrow \mathbb{R}$$

τέτοια ώστε

$$\mathrm{trace}(A) := \alpha_{1,1} + \alpha_{2,2} + \cdots + \alpha_{n,n}$$

Ιδιότητες

1. $\mathrm{trace}(\gamma A + \delta B) = \gamma \mathrm{trace}(A) + \delta \mathrm{trace}(B)$
2. $\mathrm{trace}(AB) = \mathrm{trace}(BA)$

Σημαντικό επακόλουθο: $\mathrm{trace}(Q^{-1}AQ) = \mathrm{trace}(A)$

Ορίζουσα

τετραγωνικού μητρώου

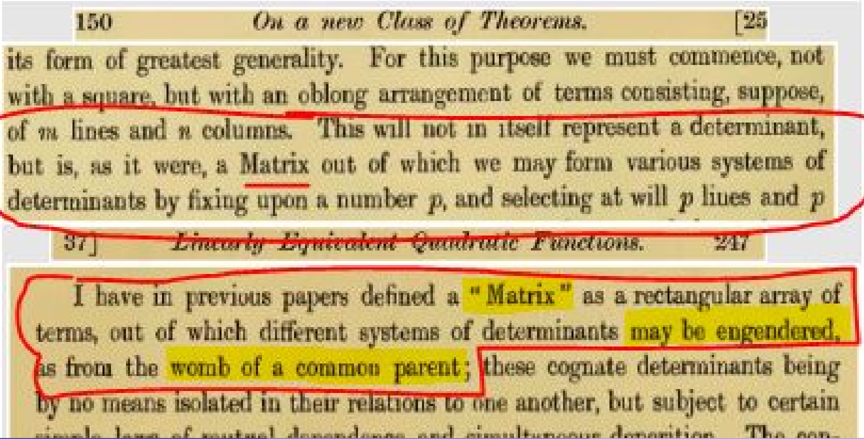
- Μοναδικό χαρακτηριστικό μέγεθος (βαθμωτός) που υπάρχει για κάθε τετραγωνικό μητρώο.

$$\det : \mathbb{R}^{n \times n} \longrightarrow \mathbb{R}$$

- «Συμπυκνώνει» πληροφορίες σχετικά με το μητρώο.
- Η πιο γνωστή: Ένα μητρώο A είναι αντιστρέψιμο αν και μόνον αν $\det(A) \neq 0$.
- Υπολογιστικά θέματα: Ο πιο πρακτικός τρόπος υπολογισμού του για μέτριο ή μεγάλο n είναι ως παραπροϊόν της παραγοντοποίησης LU με το αντίστοιχο κόστος.
- Έχει χρήσιμη γεωμετρική ερμηνεία ως «όγκος» στο n-διάστατο χώρο.

Ιστορικά

- Η μελέτη των οριζουσών προηγήθηκε της Γραμμικής Άλγεβρας: Ο Gauss (1801) χρησιμοποίησε τον όρο εννοώντας την διακρίνουσα πολυωνύμων 4ου βαθμού. Ο όρος με τη σημερινή έννοια εισήχθη από τον Cauchy (1812).
- Ο Sylvester «βάπτισε» τις "μήτρες" για να αναδείξει ότι «γεννούν» ορίζουσες.


150 *On a new Class of Theorems.* [25

its form of greatest generality. 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

37] *Linearly Equivalent Quadratic Functions.* 247

I have in previous papers 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; these cognate determinants being by no means isolated in their relations to one another, but subject to certain simple laws of mutual dependence and simultaneous deposition. The con-

Τύπος ορίζουσας όταν $n = 2$

$n = 2$

$$A = \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} \Rightarrow \det(A) = \begin{vmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{vmatrix} := \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}.$$

Στη συνέχεια θα χρησιμοποιήσουμε έναν αξιωματικό ορισμό της ορίζουσας από τον οποίο θα προκύψει αυτός και άλλο τύποι υπολογισμού.

Παρατήρηση: Να επαληθεύσετε την παρακάτω παραγοντοποίηση.

$$\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ \alpha_{21}\alpha_{11}^{-1} & 1 \end{pmatrix} \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ 0 & \alpha_{22} - \alpha_{21}\alpha_{11}^{-1}\alpha_{12} \end{pmatrix}$$

Βοηθητικοί ορισμοί

- Έστω δ.χ. $U_1, \ldots, U_n$ και W επί του ιδίου σώματος $\mathcal{F}$ και έστω η απεικόνιση

$$f : U_1 \times U_2 \times \cdots \times U_n \to W$$

τότε η απεικόνιση ονομάζεται πλειογραμμική (ή n-γραμμική) αν είναι γραμμική ως πρός κάθε όρισμα αν τα υπόλοιπα ορίσματα μείνουν αμετάβλητα, π.χ. αν για $i = 1, \ldots, n$, και $u_j \in U_j$, η απεικόνιση

$$x \to f\big(u_1, \ldots, u_{i-1}, x, u_{i+1}, \ldots, u_n\big)$$

είναι γραμμική.

- Έστω μια μετάθεση $\pi = (\pi_1, \ldots, \pi_n)$ των αριθμών 1 ως n. Ονομάζουμε φυσική διάταξη την n-άδα $(1, 2, \ldots, n)$. Τότε πρόσημο της μετάθεσης ονομάζεται ο αριθμός

$$\sigma(\pi) = \begin{cases} +1 & \text{αν η } \pi \text{ επανέρχεται στη φυσική διάταξη} \\ & \text{με άρτιο πλήθος εναλλαγών} \\ -1 & \text{αν η } \pi \text{ επανέρχεται στη φυσική διάταξη} \\ & \text{με περιττό πλήθος εναλλαγών.} \end{cases}$$

Ορίζοντας την ορίζουσα (αξιωματικά)

Η ορίζουσα $\det(A)$ ενός πραγματικού τετραγωνικού μητρώου $A \in \mathbb{R}^{n\times n}$ είναι συνάρτηση με πεδίο ορισμού το σύνολο των τετραγωνικών μητρώων και πεδίο τιμών τους βαθμωτούς (παρακάτω το σώμα $\mathbb{K}$ μπορεί να είναι το $\mathbb{R}$ ή το $\mathbb{C}$):

$$\det : \mathbb{K}^{n\times n} \longrightarrow \mathbb{K}$$

που έχει τις ακόλουθες 3 ιδιότητες[α΄] $A = [a_1, \cdots, a_n] \in \mathbb{R}^{n\times n}$:

1. Αν $a_i = a_j$ για $i \neq j$ τότε $\det([a_1, \ldots, a_n]) = 0$ (εκφυλισμός).
2. $\det([e_1, \ldots, e_n]) = 1$ (κανονικότητα).
3. Η det είναι γραμμική ως προς κάθε μία από τις n στήλες:

$$\det([a_1, \ldots, a_{j-1}, \gamma a + \delta b, a_{j+1}, \ldots, a_n]) = \gamma \det([a_1, \ldots, a_{j-1}, a, a_{j+1}, \ldots, a_n]) + \delta \det([a_1, \ldots, a_{j-1}, b, a_{j+1}, \ldots, a_n])$$

Παρατήρηση: Συναρτήσεις πολλών μεταβλητών που είναι γραμμικές ως πρός κάθε μεταβλητή λέγονται πλειογραμμικές. Για παράδειγμα το εσωτερικό γινόμενο πραγματικών διανυσμάτων (χαρακτηρίζεται και ως διγραμμικό).

[α΄] Αναφερόμαστε στις στήλες, αντίστοιχες διατυπώσεις υπάρχουν και ως προς τις γραμμές).

Σημαντικά επακόλουθα

Έστω το μητρώο (σε μορφή στηλών) $A = [a_1, \cdots, a_n] \in \mathbb{R}^{n\times n}$.

- Αν εναλλάξουμε 2 στήλες ενός μητρώου, η ορίζουσα του νέου μητρώου έχει ίδιο μέτρο αλλά αντίθετο πρόσημο. Γι' αυτό η ορίζουσα χαρακτηρίζεται ως εναλλασσόμενη απεικόνιση (alternating map).

Σημαντικά επακόλουθα

Έστω το μητρώο (σε μορφή στηλών) $A = [a_1, \cdots, a_n] \in \mathbb{R}^{n\times n}$.

- Αν εναλλάξουμε 2 στήλες ενός μητρώου, η ορίζουσα του νέου μητρώου έχει ίδιο μέτρο αλλά αντίθετο πρόσημο. Γι' αυτό η ορίζουσα χαρακτηρίζεται ως εναλλασσόμενη απεικόνιση (alternating map).
- δείγμα απόδειξης

$$\begin{aligned} 0 &= \det([a_1+a_2, a_1+a_2]) = \det([a_1,a_1]) + \det([a_1,a_2]) + \det([a_2,a_1]) + \det([a_2,a_2]) \\ 0 &= \det([a_1,a_2]) + \det([a_2,a_1]) \end{aligned}$$

Σημαντικά επακόλουθα

Έστω το μητρώο (σε μορφή στηλών) $A = [a_1, \cdots, a_n] \in \mathbb{R}^{n\times n}$.

- Αν εναλλάξουμε 2 στήλες ενός μητρώου, η ορίζουσα του νέου μητρώου έχει ίδιο μέτρο αλλά αντίθετο πρόσημο. Γι΄ αυτό η ορίζουσα χαρακτηρίζεται ως εναλλασσόμενη απεικόνιση (alternating map).
- δείγμα απόδειξης

$$\begin{aligned} 0 &= \det([a_1+a_2, a_1+a_2]) = \det([a_1,a_1]) + \det([a_1,a_2]) + \det([a_2,a_1]) + \det([a_2,a_2]) \\ 0 &= \det([a_1,a_2]) + \det([a_2,a_1]) \end{aligned}$$

- Αν οι στήλες μητρώου είναι γραμμικά εξαρτημένες, η ορίζουσα είναι 0.

Σημαντικά επακόλουθα

Έστω το μητρώο (σε μορφή στηλών) $A = [a_1, \cdots, a_n] \in \mathbb{R}^{n\times n}$.

- Αν εναλλάξουμε 2 στήλες ενός μητρώου, η ορίζουσα του νέου μητρώου έχει ίδιο μέτρο αλλά αντίθετο πρόσημο. Γι' αυτό η ορίζουσα χαρακτηρίζεται ως εναλλασσόμενη απεικόνιση (alternating map).
- δείγμα απόδειξης

$$\begin{aligned} 0 &= \det([a_1+a_2, a_1+a_2]) = \det([a_1,a_1]) + \det([a_1,a_2]) + \det([a_2,a_1]) + \det([a_2,a_2]) \\ 0 &= \det([a_1,a_2]) + \det([a_2,a_1]) \end{aligned}$$

- Αν οι στήλες μητρώου είναι γραμμικά εξαρτημένες, η ορίζουσα είναι 0.
- δείγμα απόδειξης

$$\det([a_1, a_2, \gamma a_1 + \delta a_2]) = \gamma \det([a_1,a_2,a_1]) + \delta \det([a_1,a_2,a_2]) = 0$$

Σημαντικά επακόλουθα

Έστω το μητρώο (σε μορφή στηλών) $A = [a_1, \cdots, a_n] \in \mathbb{R}^{n\times n}$.

- Αν εναλλάξουμε 2 στήλες ενός μητρώου, η ορίζουσα του νέου μητρώου έχει ίδιο μέτρο αλλά αντίθετο πρόσημο. Γι' αυτό η ορίζουσα χαρακτηρίζεται ως εναλλασσόμενη απεικόνιση (alternating map).
- δείγμα απόδειξης

$$\begin{aligned} 0 &= \det([a_1+a_2, a_1+a_2]) = \det([a_1,a_1]) + \det([a_1,a_2]) + \det([a_2,a_1]) + \det([a_2,a_2]) \\ 0 &= \det([a_1,a_2]) + \det([a_2,a_1]) \end{aligned}$$

- Αν οι στήλες μητρώου είναι γραμμικά εξαρτημένες, η ορίζουσα είναι 0.
- δείγμα απόδειξης

$$\det([a_1, a_2, \gamma a_1 + \delta a_2]) = \gamma\det([a_1,a_2,a_1]) + \delta\det([a_1,a_2,a_2]) = 0$$

- Αν προσθέσουμε δύο στήλες η ορίζουσα παραμένει αμετάβλητη.

Σημαντικά επακόλουθα

Έστω το μητρώο (σε μορφή στηλών) $A = [a_1, \cdots, a_n] \in \mathbb{R}^{n\times n}$.

- Αν εναλλάξουμε 2 στήλες ενός μητρώου, η ορίζουσα του νέου μητρώου έχει ίδιο μέτρο αλλά αντίθετο πρόσημο. Γι' αυτό η ορίζουσα χαρακτηρίζεται ως εναλλασσόμενη απεικόνιση (alternating map).
- δείγμα απόδειξης

$$\begin{aligned} 0 &= \det([a_1 + a_2, a_1 + a_2]) = \det([a_1, a_1]) + \det([a_1, a_2]) + \det([a_2, a_1]) + \det([a_2, a_2]) \\ 0 &= \det([a_1, a_2]) + \det([a_2, a_1]) \end{aligned}$$

- Αν οι στήλες μητρώου είναι γραμμικά εξαρτημένες, η ορίζουσα είναι 0.
- δείγμα απόδειξης

$$\det([a_1, a_2, \gamma a_1 + \delta a_2]) = \gamma \det([a_1, a_2, a_1]) + \delta \det([a_1, a_2, a_2]) = 0$$

- Αν προσθέσουμε δύο στήλες η ορίζουσα παραμένει αμετάβλητη.
- δείγμα απόδειξης

$$\det([a_1, a_1 + a_2]) = \det([a_1, a_1]) + \det([a_1, a_2]) = \det([a_1, a_2])$$

Ορίζουσα μητρώου 2×2

(Ποιά είναι επιτέλους ;;;)

Ορίζουσα μητρώου 2×2

(Ποιά είναι επιτέλους;;;)

Μπορούμε να την υπολογίσουμε βάσει των παραπάνω ιδιοτήτων:

$$\begin{aligned}\det([a_1, a_2]) &= \det([\alpha_{11}e_1 + \alpha_{21}e_2, \alpha_{12}e_1 + \alpha_{22}e_2]) \\ &= \alpha_{11}\det([e_1, \alpha_{12}e_1 + \alpha_{22}e_2]) + \alpha_{21}\det([e_2, \alpha_{12}e_1 + \alpha_{22}e_2])\end{aligned}$$

Ορίζουσα μητρώου 2×2

(Ποιά είναι επιτέλους;;;)

Μπορούμε να την υπολογίσουμε βάσει των παραπάνω ιδιοτήτων:

$$\begin{aligned}\det([a_1, a_2]) &= \det([\alpha_{11}e_1 + \alpha_{21}e_2, \alpha_{12}e_1 + \alpha_{22}e_2]) \\ &= \alpha_{11}\det([e_1, \alpha_{12}e_1 + \alpha_{22}e_2]) + \alpha_{21}\det([e_2, \alpha_{12}e_1 + \alpha_{22}e_2])\end{aligned}$$

$$= \alpha_{11}\alpha_{12}\det([e_1, e_1]) + \alpha_{11}\alpha_{22}\det([e_1, e_2]) + \alpha_{21}\alpha_{12}\det([e_2, e_1]) + \alpha_{21}\alpha_{22}\det([e_2, e_2])$$

$$\det([a_1, a_2]) = \alpha_{11}\alpha_{22} + \alpha_{21}\alpha_{12}\det([e_2, e_1]) = \boldsymbol{\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}}$$

Προσοχή: πολλοί όροι μηδενίζονται (λόγω «εκφυλισμού»).

- Όταν $n = 2$, αντί για $4 = 2^2$ όρους, έχουμε μόνον 2.
- με $n = 3$, αντί για $27 = 3^3$ όρους έχουμε 6 ...
- ... γενικά, αντί για για n^n όρους, έχουμε $n!$.

Υπολογισμός ορίζουσας από αλγεβρικά συμπληρώματα

Ανάπτυγμα Laplace

$$\det(A) = \alpha_{11}\det(A_{11}) + (-1)^{1+2}\alpha_{12}\det(A_{12}) + \cdots + (-1)^{1+n}\alpha_{1n}\det(A_{1n})$$

- Προσωρινός συμβολισμός Με $A_{ij} \in \mathbb{R}^{(n-1)\times(n-1)}$ συμβολίζεται το μητρώο που προκύπτει αν διαγράψουμε από το A τη γραμμή i και τη στήλη j. Αποκαλείται ελάσσον (minor) του στοιχείου α_{ij}.
- Το $\psi_{i,j} := (-1)^{i+j}\det(A_{ij})$ λέγεται αλγεβρικό συμπλήρωμα του στοιχείου στη θέση (i,j).

Γενικά ισχύει ότι μπορούμε να εκφράσουμε την ορίζουσα ως προς οποιαδήποτε γραμμή (ή στήλη):

$$\det(A) = \alpha_{i,1}\psi_{i,1} + \alpha_{i,2}\psi_{i,2} + \cdots + \alpha_{i,n}\psi_{i,n}$$

Προσοχή: Στον παραπάνω τύπο εκφράσαμε την ορίζουσα βάσει οριζουσών μεγέθους $(n-1)\times(n-1)$.
Σύσταση: Στην εφαρμογή του τύπου, αξίζει να επιλέγουμε τη γραμμή ή στήλη με τα περισσότερα μηδενικά!

Τύπος για $n = 3$

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} := \alpha_{11} \begin{vmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{vmatrix} - \alpha_{12} \begin{vmatrix} \alpha_{21} & \alpha_{23} \\ \alpha_{31} & \alpha_{33} \end{vmatrix} + \alpha_{13} \begin{vmatrix} \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{vmatrix}$$

Άρα

Τύπος για $n = 3$

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} := \alpha_{11} \begin{vmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{vmatrix} - \alpha_{12} \begin{vmatrix} \alpha_{21} & \alpha_{23} \\ \alpha_{31} & \alpha_{33} \end{vmatrix} + \alpha_{13} \begin{vmatrix} \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{vmatrix}$$

Άρα

$$\det(A) = \alpha_{11}\alpha_{22}\alpha_{33} - \alpha_{11}\alpha_{23}\alpha_{32} - \alpha_{12}\alpha_{21}\alpha_{33} + \alpha_{12}\alpha_{23}\alpha_{31} + \alpha_{13}\alpha_{21}\alpha_{32} - \alpha_{13}\alpha_{22}\alpha_{31}$$

Τύπος για $n = 3$

$$\begin{vmatrix} \alpha_{11} & \alpha_{12} & \alpha_{13} \\ \alpha_{21} & \alpha_{22} & \alpha_{23} \\ \alpha_{31} & \alpha_{32} & \alpha_{33} \end{vmatrix} := \alpha_{11}\begin{vmatrix} \alpha_{22} & \alpha_{23} \\ \alpha_{32} & \alpha_{33} \end{vmatrix} - \alpha_{12}\begin{vmatrix} \alpha_{21} & \alpha_{23} \\ \alpha_{31} & \alpha_{33} \end{vmatrix} + \alpha_{13}\begin{vmatrix} \alpha_{21} & \alpha_{22} \\ \alpha_{31} & \alpha_{32} \end{vmatrix}$$

Άρα

$$\det(A) = \alpha_{11}\alpha_{22}\alpha_{33} - \alpha_{11}\alpha_{23}\alpha_{32} - \alpha_{12}\alpha_{21}\alpha_{33} + \alpha_{12}\alpha_{23}\alpha_{31} + \alpha_{13}\alpha_{21}\alpha_{32} - \alpha_{13}\alpha_{22}\alpha_{31}$$

Προσέξτε:

$$\det(A) = \alpha_{1{\color{red}1}}\alpha_{2{\color{red}2}}\alpha_{3{\color{red}3}} - \alpha_{1{\color{red}1}}\alpha_{2{\color{red}3}}\alpha_{3{\color{red}2}} - \alpha_{1{\color{red}2}}\alpha_{2{\color{red}1}}\alpha_{3{\color{red}3}} + \alpha_{1{\color{red}2}}\alpha_{2{\color{red}3}}\alpha_{3{\color{red}1}} + \alpha_{1{\color{red}3}}\alpha_{2{\color{red}1}}\alpha_{3{\color{red}2}} - \alpha_{1{\color{red}3}}\alpha_{2{\color{red}2}}\alpha_{3{\color{red}1}}$$

Μεγάλος τύπος ορίζουσας

... και πρόσημο μετάθεσης

Προσοχή: Για μία μετάθεση π των στοιχείων $(1,\ldots,n)$, δηλ. $(\pi(1),\ldots,\pi(n))$, ορίζουμε ως πρόσημο της μετάθεσης, $\sigma(\pi)$, μία τιμή ± 1. Η τιμή θα είναι +1 αν η μετάθεση π προέρχεται από άρτιο αριθμό εναλλαγών ή -1 αν προέρχεται από περιττό αριθμό εναλλαγών. Το αντίστοιχο μητρώο μετάθεσης P_π θα είναι εκείνο για το οποίο $P_\pi \begin{pmatrix} 1 \\ 2 \\ \vdots \\ n \end{pmatrix} = \begin{pmatrix} \pi(1) \\ \pi(2) \\ \cdots \\ \pi(n) \end{pmatrix}$. Παρατηρήστε ότι το P_π είναι μετάθεση του ταυτοτικού, επομένως $\det(P_\pi) = \pm 1$. Η τιμή του συμπίπτει με το πρόσημο της μετάθεσης.

Μεγάλος τύπος (γιατί έχει $n!$ όρους)

$\det(A) = \sum_{\pi \in \mathcal{S}} \sigma(\pi) \alpha_{1\pi(1)} \cdots \alpha_{n\pi(n)}$ όπου $\mathcal{S}$ είναι το σύνολο των $n!$ μεταθέσεων των στοιχείων $(1,\ldots,n)$ και $\sigma(\pi)$ είναι το πρόσημο της μετάθεσης π.

Ορίζουσες μητρώων με ειδική δομή

Τριγωνικά: Η ορίζουσα κάθε τριγωνικού μητρώου είναι ίση με το γινόμενο όλων των στοιχείων της διαγωνίου του.

Κατά πλοκάδες τριγωνικά Αν A, D τετραγωνικά μητρώα, τότε

$$\det\begin{pmatrix} A & B \\ 0 & D \end{pmatrix} = \det(A)\det(D)$$

Ειδική περίπτωση αν $\hat{A} \in \mathbb{R}^{(n-1)\times(n-1)}$,

$$\det\left(\begin{array}{c|ccc} 1 & \star & \cdots & \star \\ \hline 0_{n-1,1} & \hat{A} & & \end{array}\right) = \det(\hat{A})$$

Προσοχή: Γενικά

$$\begin{aligned}\det\begin{pmatrix} A & B \\ C & D \end{pmatrix} &\neq \det(A)\det(D) \text{ καθώς επίσης} \\ &\neq \det(A)\det(D) - \det(B)\det(C)\end{aligned}$$

Άλλες χρήσιμες ιδιότητες

- Το A είναι αντιστρέψιμο αν και μόνον αν $\det(A) \neq 0$.
- $\det(A^{-1}) = 1/\det(A^\top)$ αν $\det(A) \neq 0$.
- $\det(A) = \det(A^\top)$
- $\det(AB) = \det(A)\det(B)$

<u>ΠΡΟΣΟΧΗ</u> Η τελευταία ιδιότητα αποκαλύπτει έναν τρόπο υπολογισμού του $\det(A)$ που έχει περίπου όσες πράξεις χρειάζεται η παραγοντοποίηση LU.
<u>Πώς;</u> Αν $A = PLU$ τότε

$$\begin{aligned} \det(A) &= \det(P)\underbrace{\det(L)}_{=1}\det(U) \\ &= (-1)^\kappa \det(U) = (-1)^\kappa \upsilon_{11}\upsilon_{22}\cdots\upsilon_{nn}. \end{aligned}$$

όπου κ είναι ο αριθμός των εναλλαγών που μετατρέπουν το I σε P. Το $(-1)^\kappa$ θα είναι το πρόσημο της αντίστοιχης μετάθεσης.

Κριτική

Γιατί δεν λύνουμε με Cramer;

- Ο υπολογισμός της ορίζουσας με τους παραπάνω τύπους είναι ακριβός
- ... εκτός αν την υπολογίσουμε μέσω LU
- ... οπότε δεν υπάρχει λόγος να χρησιμοποιήσουμε τον κανόνα
- ... γιατί από την LU βρίσκουμε τη λύση με μπρος και πίσω αντικατάσταση, πιο φθηνά!
- Εκτός αν επικρατούν ειδικές συνθήκες! (π.χ. κάτω τριγωνικό μητρώο)

Ορίζουσες και τάξη μητρώου

Ενδιαφέρον: Ένα μητρώο $A \in \mathbb{R}^{m\times n}$ έχει τάξη ακριβώς r ανν και μόνον αν το μεγαλύτερο τετραγωνικό (υπο)μητρώο με μη μηδενική ορίζουσα είναι $r \times r$.

Κανόνας Cramer

και επίλυση γραμμικών συστημάτων

Αν $Ax = b$ τότε κάθε στοιχείο του x μπορεί να υπολογιστεί με τον κανόνα Cramer που εκφράζεται ως εξής:

Έστω οι στήλες $a_1, \ldots, a_n$ τ.ώ. $\det([a_1, \ldots, a_n]) \neq 0$. Έστω επίσης βαθμωτοί $\xi_1, \ldots, \xi_n$ τ.ώ.

$$x = \xi_1 e_1 + \cdots + \xi_n e_n.$$

Τότε για κάθε $j = 1, \ldots, n$ ισχύει ότι

$$\xi_j = \frac{\det([a_1, \ldots, a_{j-1}, b, a_{j+1}, \ldots, a_n])}{\det([a_1, \ldots, a_n])}$$

Τύπος αντιστρόφου

Θεώρημα

Αν το A είναι αντιστρέψιμο, τότε

$$\begin{aligned}(A^{-1})_{i,j} &= (-1)^{i+j}\frac{\det(A_{j,i})}{\det(A)}\\ &= \frac{\psi_{j,i}}{\det(A)}\end{aligned}$$

Παραδείγματα οριζουσών

εμφανίζονται συχνά στις εφαρμογές

<u>Μητρώο τάξης 1</u> Έστω ότι από τα διανύσματα (στήλες) u, v κατασκευάζουμε το μητρώο $A = uv^\top \in \mathbb{R}^{n\times n}$. Ποιά θα είναι η ορίζουσα;

Παραδείγματα οριζουσών

εμφανίζονται συχνά στις εφαρμογές

Μητρώο τάξης 1 Έστω ότι από τα διανύσματα (στήλες) u, v κατασκευάζουμε το μητρώο $A = uv^\top \in \mathbb{R}^{n\times n}$. Ποιά θα είναι η ορίζουσα;

Απάντηση: 0 (γιατί;)

Παραδείγματα οριζουσών

εμφανίζονται συχνά στις εφαρμογές

Μητρώο τάξης 1 Έστω ότι από τα διανύσματα (στήλες) u, v κατασκευάζουμε το μητρώο $A = uv^\top \in \mathbb{R}^{n\times n}$. Ποιά θα είναι η ορίζουσα;

Απάντηση: 0 (γιατί;)

Μητρώο προβολής σε (γνήσιο) υπόχωρο;

Απάντηση: 0 (γιατί;)

Παραδείγματα οριζουσών

εμφανίζονται συχνά στις εφαρμογές

Μητρώο τάξης 1 Έστω ότι από τα διανύσματα (στήλες) u, v κατασκευάζουμε το μητρώο $A = uv^\top \in \mathbb{R}^{n\times n}$. Ποιά θα είναι η ορίζουσα;

Απάντηση: 0 (γιατί;)

Μητρώο προβολής σε (γνήσιο) υπόχωρο;

Απάντηση: 0 (γιατί;)

Ορθογώνιο μητρώο;

Απάντηση: $\det(A) = \pm 1$ (γιατί;)

Παραδείγματα οριζουσών

εμφανίζονται συχνά στις εφαρμογές

Μητρώο τάξης 1 Έστω ότι από τα διανύσματα (στήλες) u, v κατασκευάζουμε το μητρώο $A = uv^\top \in \mathbb{R}^{n\times n}$. Ποιά θα είναι η ορίζουσα;

Απάντηση: 0 (γιατί;)

Μητρώο προβολής σε (γνήσιο) υπόχωρο;

Απάντηση: 0 (γιατί;)

Ορθογώνιο μητρώο;

Απάντηση: $\det(A) = \pm 1$ (γιατί;)

Ταυτοτικό συν μητρώο ‘τάξης-1’ $A = I + uv^\top$. Ποιά θα είναι η ορίζουσα;

Παραδείγματα οριζουσών

εμφανίζονται συχνά στις εφαρμογές

Μητρώο τάξης 1 Έστω ότι από τα διανύσματα (στήλες) u, v κατασκευάζουμε το μητρώο $A = uv^\top \in \mathbb{R}^{n\times n}$. Ποιά θα είναι η ορίζουσα;

Απάντηση: 0 (γιατί;)

Μητρώο προβολής σε (γνήσιο) υπόχωρο;

Απάντηση: 0 (γιατί;)

Ορθογώνιο μητρώο;

Απάντηση: $\det(A) = \pm 1$ (γιατί;)

Ταυτοτικό συν μητρώο 'τάξης-1' $A = I + uv^\top$. Ποιά θα είναι η ορίζουσα;

Απάντηση: $\det(A) = 1 + v^\top u$. (γιατί;)

Παραδείγματα οριζουσών

εμφανίζονται συχνά στις εφαρμογές

Μητρώο τάξης 1 Έστω ότι από τα διανύσματα (στήλες) u, v κατασκευάζουμε το μητρώο $A = uv^\top \in \mathbb{R}^{n\times n}$. Ποιά θα είναι η ορίζουσα;

Απάντηση: 0 (γιατί;)

Μητρώο προβολής σε (γνήσιο) υπόχωρο;

Απάντηση: 0 (γιατί;)

Ορθογώνιο μητρώο;

Απάντηση: $\det(A) = \pm 1$ (γιατί;)

Ταυτοτικό συν μητρώο 'τάξης-1' $A = I + uv^\top$. Ποιά θα είναι η ορίζουσα;

Απάντηση: $\det(A) = 1 + v^\top u$. (γιατί;)

$$B = \begin{pmatrix} 1 & -v^\top \\ u & I \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ u & I \end{pmatrix} \begin{pmatrix} 1 & -v^\top \\ 0 & I + uv^\top \end{pmatrix}, \text{ και } C = \begin{pmatrix} I & u \\ -v^\top & 1 \end{pmatrix} = \begin{pmatrix} I & 0 \\ -v^\top & 1 \end{pmatrix} \begin{pmatrix} I & u \\ 0 & 1 + v^\top u \end{pmatrix}$$

και αν $P = \begin{pmatrix} 0_{1,n} & 1 \\ I_n & 0_{n,1} \end{pmatrix}$ τότε $\det(P) = (-1)^{n+2}$ και $P^\top BP = C$ άρα $\det(P^\top BP) = \det(B) = \det(C)$. Από τις παραγοντοποιήσεις, $\det(I + uv^\top) = 1 + v^\top u$.

Γενίκευση και παράδειγμα

- **Ταυτότητα Sylvester** Γενικότερα ισχύει ότι αν $C \in \mathbb{R}^{m\times n}, D \in \mathbb{R}^{n\times m}$

$$\det(I_m + CD) = \det(I_n + DC)$$

- Συνιστάται να την αποδείξετε!

Παράδειγμα: Να υπολογίσετε την ορίζουσα του

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

Επίλυση: Παρατηρούμε ότι το μητρώο μπορεί να γραφτεί ως $A = 2I + 2ee^\top$ όπου $e = (1, 1, 1, 1)^\top$. Επομένως

$$\det(A) = \det(2(I + ee^\top)) = 2^4\det(I + ee^\top) = 2^4(1 + e^\top e) = 16 \cdot 5 = 80.$$

☺ Θα ήταν αρκετά πιο χρονοβόρο αν δουλεύατε απευθείας με το A.

Οι ορίζουσες στη σύγχρονη έρευνα

How to find a good submatrix

S. A. Goreinov, I. V. Oseledets, D. V. Savostyanov,
E. E. Tyrtyshnikov, N. L. Zamarashkin

October 17, 2008

Abstract

Pseudoskeleton approximation and some other problems require the knowledge of sufficiently well-conditioned submatrix in a large-scale matrix. The quality of a submatrix can be measured by modulus of its determinant, also known as volume. In this paper we discuss a search algorithm for the maximum-volume submatrix which already proved to be useful in several matrix and tensor approximation algorithms. We investigate the behavior of this algorithm on random matrices and present some its applications, including maximization of a bivariate functional.

Από τις ορίζουσες στις permanents

Για φανατικούς !

$$\mathrm{perm}(A) := \sum_{\pi \in \mathcal{S}} \alpha_{1\pi(1)} \cdots \alpha_{n\pi(n)}$$

όπου $\mathcal{S}$ είναι το σύνολο των $n!$ μεταθέσεων των στοιχείων $(1, \ldots, n)$.

Ενώ η ορίζουσα μπορεί να υπολογιστεί με $O(n^3)$ πράξεις, μέσω της παραγοντοποίησης LU, ο υπολογισμός της $\mathrm{perm}(A)$ απαιτεί πολύ μεγάλο πλήθος πράξεων (μεγαλύτερο από πολυωνυμικό)! Ο πιο γνωστός αλγόριθμος (Ryser) για τον ακριβή υπολογισμό απαιτεί $O(n2^n)$ πράξεις !!!
Ερευνητικό πρόβλημα : Επινόηση γρήγορων αλγορίθμου για την αποτελεσματική προσέγγιση της permanent σε πολυωνυμικό χρόνο.

The International Journal of Robotics Research
000(00):1–16

DOI:doi number
http://mms.sagepub.com

Localization from Semantic Observations via the Matrix Permanent

Nikolay Atanasov* **Menglong Zhu** **Kostas Daniilidis** **George J. Pappas**
GRASP Laboratory, University of Pennsylvania, Philadelphia, PA 19104, USA

Abstract

Most approaches to robot localization rely on low-level geometric features such as points, lines, and planes. In this paper, we use object recognition to obtain semantic information from the robot's sensors and consider the task of localizing the robot within a prior map of landmarks, which are annotated with semantic labels. As object recognition algorithms miss detections and produce false alarms, correct data association between the detections and the landmarks on the map is central to the semantic localization problem. Instead of the traditional vector-based representation, we propose a sensor model, which encodes the semantic observations via random finite sets and enables a unified treatment of missed detections, false alarms, and data association. Our second contribution is to reduce the problem of computing the likelihood of a set-valued observation to the problem of computing a matrix permanent. It is this crucial transformation that allows us to solve the semantic localization problem with a polynomial-time approximation to the set-based Bayes filter. Finally, we address the active semantic localization problem, in which the observer's trajectory is planned in order to improve the accuracy and efficiency of the localization process. The performance of our approach is demonstrated in simulation and in real environments using deformable-part-model-based object detectors. Robust global localization from semantic observations

Down with Determinants!

Sheldon Axler

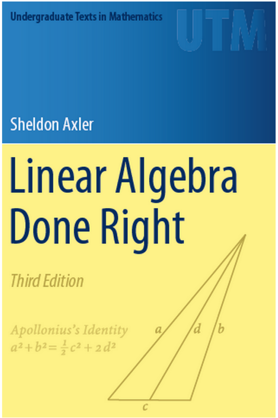
This paper was published in the *American Mathematical Monthly* 102 (1995), 139-154.

In 1996 this paper received the Lester R. Ford Award for expository writing from the Mathematical Association of America.

Abstract: This paper shows how linear algebra can be done better without determinants. The standard proof that a square matrix

complexification discussed earlier.

Now we are ready for the formal definition. The *determinant* of T, denoted $\det T$, is defined to be the product of the eigenvalues of T, counting multiplicity. This definition would not be possible with the traditional approach to eigenvalues, because that method uses determinants to prove that eigenvalues exist. With the techniques used here, we already know (by Theorem 3.11(a)) that T has $\dim V$ eigenvalues, counting multiplicity. Thus our simple definition makes sense.



This paper focuses on showing that determinants should be banished from much of the theoretical part of linear algebra. Determinants are also useless in the computational part of linear algebra. For example, Cramer's rule for solving systems of linear equations is already worthless for 10×10 systems, not to mention the much larger systems often encountered in the real world. Many computer programs efficiently calculate eigenvalues numerically—none of them uses determinants. To emphasize the point, let me quote a numerical analyst. Henry Thacher, in a review (*SIAM News*, September 1988) of the *Turbo Pascal Numerical Methods Toolbox*, writes,

> I find it hard to conceive of a situation in which the numerical value of a determinant is needed: Cramer's rule, because of its inefficiency, is completely impractical, while the magnitude of the determinant is an indication of neither the condition of the matrix nor the accuracy of the solution.

A Randomized Algorithm for Approximating the Log Determinant of a Symmetric Positive Definite Matrix

Christos Boutsidis [*] Petros Drineas [†] Prabhanjan Kambadur [‡]
Eugenia-Maria Kontopoulou [§] Anastasios Zouzias [¶]

Abstract

We introduce a novel algorithm for approximating the logarithm of the determinant of a symmetric positive definite (SPD) matrix. The algorithm is randomized and approximates the traces of a small number of matrix powers of a specially constructed matrix, using the method of Avron and Toledo [AT11]. From a theoretical perspective, we present additive and relative error bounds for our algorithm. Our additive error bound works for any SPD matrix, whereas our relative error bound works for SPD matrices whose eigenvalues lie in the interval $(\theta_1, 1)$, with $0 < \theta_1 < 1$; the latter setting was proposed in [HMS15]. From an empirical perspective, we demonstrate that a C++ implementation of our algorithm can approximate the logarithm of the determinant of large matrices very accurately in a matter of seconds.

Αλγεβρικό πρόβλημα ιδιοτιμών

Στις επόμενες διαλέξεις επικεντρωνόμαστε στο εξής πρόβλημα:

Δοθέντος $A \in \mathbb{R}^{n \times n}$ ή $A \in \mathbb{C}^{n \times n}$ να βρεθούν τα ζεύγη $\lambda \in \mathbb{C}, x \in \mathbb{C}^n$ τέτοια ώστε

$$Ax = \lambda x.$$

ή αντίστοιχα,

Δοθέντος $A \in \mathbb{R}^{n \times n}$ ή $A \in \mathbb{C}^{n \times n}$ να βρεθούν τα ζεύγη $\lambda \in \mathbb{C}, y \in \mathbb{C}^n$ τέτοια ώστε

$$y^* A = \lambda y^*.$$

όπου $y^* = \bar{y}^\top$ και με $\bar{y}$ συμβολίζεται το διάνυσμα με τιμές τις συζυγείς του y. Αν $y \in \mathbb{R}^n$ τότε $y^* = y^\top$.

Ιδιοτιμές και ιδιοδιανύσματα I

Ορισμός

Έστω ότι $A \in \mathbb{C}^{n\times n}$. Τότε κάθε μιγαδικός αριθμός λ για τον οποίο το μητρώο $(\lambda I - A)$ είναι μη αντιστρέψιμο ονομάζεται **ιδιοτιμή του A**.

- Έπεται ότι η διάσταση του μηδενόχωρου $\dim \mathrm{null}(\lambda I - A) \geq 1$. Κάθε διάνυσμά $x \in \mathrm{null}(\lambda I - A)$ ονομάζεται (δεξιό) **ιδιοδιάνυσμα του A** (για την ιδιοτιμή λ).
- Έπεται ότι η διάσταση του μηδενόχωρου $\dim \mathrm{null}(\lambda I - A)^* \geq 1$. Κάθε διάνυσμά $y \in \mathrm{null}(\lambda I - A)^*$ ονομάζεται (αριστερό) **ιδιοδιάνυσμα του A** (για την ιδιοτιμή λ), οπότε $y^* A = \lambda y^*$.

Οι ιδιοτιμές και τα ιδιοδιανύσματα ενός μητρώου αναδεικνύουν πολλά σημαντικά ζητήματα για ένα μητρώο και για τους διανυσματικούς (υπο)χώρους που συνδέονται με αυτό και χρησιμοποιούνται πολύ στις εφαρμογές.

Ιδιοτιμές και ιδιοδιανύσματα II

Ερωτήματα:

- Υπάρχουν ιδιοτιμές; Πόσες υπάρχουν;
- Υπάρχουν ιδιοδιανύσματα; Πόσα ιδιοδιανύσματα (ή καλύτερα, ποιές είναι οι διαστάσεις των αντίστοιχων μηδενοχώρων;)
- Πού και πώς εντοπίζονται (ιδιοδιανύσματα, ιδιοτιμές); Πώς υπολογίζονται;
- Τι χαρακτηριστικά έχουν και γιατί μας ενδιαφέρουν;

Το μηδέν ΜΠΟΡΕΙ να είναι ιδιοτιμή, το μηδενικό διάνυσμα ΔΕΝ θεωρείται ιδιοδιάνυσμα.

SIAM REVIEW
Vol. 48, No. 3, pp. 569–581

The $25,000,000,000 Eigenvector: The Linear Algebra behind Google*

Kurt Bryan†
Tanya Leise‡

Abstract. Google's success derives in large part from its PageRank algorithm, which ranks the importance of web pages according to an eigenvector of a weighted link matrix. Analysis of the PageRank formula provides a wonderful applied topic for a linear algebra course. Instructors may assign this article as a project to more advanced students or spend one or two lectures presenting the material with assigned homework from the exercises. This material also complements the discussion of Markov chains in matrix algebra. Maple and *Mathematica* files supporting this material can be found at www.rose-hulman.edu/~bryan.

Key words. linear algebra, PageRank, eigenvector, stochastic matrix

Τι λέει και τι γράφει ο κόσμος για τις ιδιοτιμές και τα ιδιοδιανύσματα;

Google eigenvalues

Ιστός Βίντεο Εικόνες Βιβλία Ειδήσεις Περισσότερα Εργ

Περίπου 5.820.000 αποτελέσματα (0,12 δευτερόλεπτα)

Eigenvalues and eigenvectors - Wikipedia, the free ...
en.wikipedia.org/.../**Eigenvalues**_and_eige... Μετάφραση αυτής της σελίδας
Eigenvalues and eigenvectors have many applications in both pure and applied mathematics. They are used in matrix factorization, in quantum mechanics, and ...
Eigenvalue algorithm - Eigenface - Square matrix - Quadratic eigenvalue problem

Google eigenvectors

Ιστός Βίντεο Εικόνες Ειδήσεις Χάρτες Περισσότερα Εργ

Περίπου 2.480.000 αποτελέσματα (0,15 δευτερόλεπτα)


Mathematics FOR Quantum Mechanics
An Introductory Survey of Operators, Eigenvalues, and Linear Vector Spaces
John David Jackson
LOOK INSIDE!
Eigenvalues of Matrices
Françoise Chatelin
C·L·A·S·S·I·C·S
In Applied Mathematics
siam
Free Vibrations of Beams and Frames
Eigenvalues and Eigenfunctions
Igor A. Karnovsky
Olga I. Lebed
McGRAW-HILL ENGINEERING REFERENCE
Probability and Its Applications
Mu-Fa Chen
Eigenvalues, Inequalities, and Ergodic Theory
Springer
Numerical Methods for Large Eigenvalue Problems
Revised Edition
Yousef Saad
Geometric Control Theory and Sub-Riemannian Geometry
Springer
Eigenvalues of Symmetric Five-Diagonal Matrices
L. W. Ehrlich
The Asymptotic Distribution of Eigenvalues of Partial Differential Operators
Antoine Henrot
Extremum Problems for Eigenvalues of Elliptic Operators
Numerical Analysis 2000, Volume 3
Linear Algebra - Linear Systems and Eigenvalues
North-Holland
On the Eigenvalues of the Manakov System
LAMBERT
Jan Lang
David Edmunds
Eigenvalues, Embeddings and Generalised Trigonometric Functions
Springer
Volume IX
Surveys in Differential Geometry
Eigenvalues of Laplacians and other geometric operators
Eigenvalues in Uniform Waveguides Using Transmission Line Equivalences
COLLOQUIUM PUBLICATIONS
Random Matrices, Frobenius Eigenvalues, and Monodromy
London Mathematical Society
Lecture Note Series
CAMBRIDGE UNIVERSITY PRESS
Isaac Elishakoff
EIGENVALUES OF INHOMOGENEOUS STRUCTURES
Unusual Closed-Form Solutions
CRC PRESS
MEMOIRS of the American Mathematical Society
Kuznetsov's Trace Formula and the Heck Eigenvalues of Maass Forms
A. Knightly
C. Li
American Mathematical Society
DIGITALBookmobile
CALCULUS OF VARIATIONS
AND PARTIAL DIFFERENTIAL EQUATIONS
G. Prodi (Ed.)
Eigenvalues of Non-Linear Problems
Springer
CIME
The Symmetric Eigenvalue Problem
Beresford N. Parlett
MULTIPARAMETER EIGENVALUE PROBLEMS
STURM-LIOUVILLE THEORY
F.V. ATKINSON
ANGELO B. MINGARELLI
CRC Press
EIGENVALUES OF MATRICES
Françoise Chatelin
WILEY

Παράδειγμα

Από το $A = \begin{pmatrix} -3 & 2 \\ -12 & 7 \end{pmatrix}$ κατασκευάζουμε το παραμετροποιημένο

$$A(\lambda) = \lambda I - A = \begin{pmatrix} \lambda+3 & -2 \\ 12 & \lambda-7 \end{pmatrix}$$

Οι ιδιοτιμές του A είναι όλες εκείνες οι τιμές λ (ενδεχομένως μιγαδικές) για τις οποίες το $A(\lambda)$ δεν είναι αντιστρέψιμο. Στην περίπτωσή μας, μπορούμε να επαληθεύσουμε ότι τα $A(1), A(3)$ δεν είναι αντιστρέψιμα. Επίσης

$$A(1)u_1 = 0, A(3)u_2 = 0$$

και λύνοντας τα αντίστοιχα ομογενή συστήματα προκύπτει ότι (ειδικές) λύσεις τους είναι οι $u_1 = (1,2)^\top$ και $u_2 = (1,3)^\top$. Προφανώς ισχύει ότι $\mathrm{null}(A(1)) = \mathrm{span}\{u_1\}$ και $\mathrm{null}(A(3)) = \mathrm{span}\{u_2\}$. Επομένως, ως ιδιοδιάνυσμα του A για την ιδιοτιμή 1 μπορούμε να επιλέξουμε οποιοδήποτε διάνυσμα του $\mathrm{null}(A(1))$. Συνήθως, επιλέγουμε διανύσματα που έχουν κανονικοποιηθεί με κάποιον τρόπο, π.χ. τέτοια ώστε $\|u_1\|_2 = \|u_2\|_2 = 1$.

<u>Προσοχή:</u> Το παραπάνω παράδειγμα δεν εξηγεί με ποιόν τρόπο επιλέξαμε τα $A(1)$ και $A(3)$. Ούτε γιατί αυτά είναι τα μόνα δύο μη αντιστρέψιμα μητρώα του τύπου $\lambda I - A$. Ο σχεδιασμός συστηματικών μεθόδων αναζήτησης και υπολογισμού τους είναι ένα από τα βασικά ζητούμενα όχι μόνον αυτού του κεφαλαίου, αλλά και πολλών βιβλίων και της σημερινής έρευνας.

Ένα χαρακτηριστικό πολυώνυμο για κάθε μητρώο

Χαρακτηριστικά πολυώνυμα και ιδιοτιμές

Οι πληροφορίες που θέλουμε εξάγονται από ένα ειδικό πολυώνυμο που υπολογίζεται από την ορίζουσα $\det(\lambda I - A)$ ως προς τη μεταβλητή λ.

Εξετάζουμε το $\det(\lambda I - A)$:

- Αν

$$
\begin{aligned}
\lambda I - A &= \begin{pmatrix} \lambda - \alpha_{11} & -\alpha_{12} \\ -\alpha_{21} & \lambda - \alpha_{22} \end{pmatrix} \\
\det(\lambda I - A) &= (-\alpha_{11} + \lambda)(-\alpha_{22} + \lambda) - \alpha_{21}\alpha_{12} \\
&= \lambda^2 - (\alpha_{11} + \alpha_{22})\lambda + \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}
\end{aligned}
$$

Ένα χαρακτηριστικό πολυώνυμο για κάθε μητρώο

Χαρακτηριστικά πολυώνυμα και ιδιοτιμές

Οι πληροφορίες που θέλουμε εξάγονται από ένα ειδικό πολυώνυμο που υπολογίζεται από την ορίζουσα $\det(\lambda I - A)$ ως προς τη μεταβλητή λ.

Εξετάζουμε το $\det(\lambda I - A)$:

- Αν

$$\begin{aligned} \lambda I - A &= \begin{pmatrix} \lambda - \alpha_{11} & -\alpha_{12} \\ -\alpha_{21} & \lambda - \alpha_{22} \end{pmatrix} \\ \det(\lambda I - A) &= (-\alpha_{11} + \lambda)(-\alpha_{22} + \lambda) - \alpha_{21}\alpha_{12} \\ &= \lambda^2 - (\alpha_{11} + \alpha_{22})\lambda + \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12} \end{aligned}$$

- Το $\det(A - \lambda I)$ είναι πολυώνυμο 2ου βαθμού ως προς λ.

Ένα χαρακτηριστικό πολυώνυμο για κάθε μητρώο

Χαρακτηριστικά πολυώνυμα και ιδιοτιμές

Οι πληροφορίες που θέλουμε εξάγονται από ένα ειδικό πολυώνυμο που υπολογίζεται από την ορίζουσα $\det(\lambda I - A)$ ως προς τη μεταβλητή λ.

Εξετάζουμε το $\det(\lambda I - A)$:

- Αν

$$\begin{aligned}
\lambda I - A &= \begin{pmatrix} \lambda - \alpha_{11} & -\alpha_{12} \\ -\alpha_{21} & \lambda - \alpha_{22} \end{pmatrix} \\
\det(\lambda I - A) &= (-\alpha_{11} + \lambda)(-\alpha_{22} + \lambda) - \alpha_{21}\alpha_{12} \\
&= \lambda^2 - (\alpha_{11} + \alpha_{22})\lambda + \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}
\end{aligned}$$

- Το $\det(A - \lambda I)$ είναι πολυώνυμο 2ου βαθμού ως προς λ.
- Προσέξτε τους συντελεστές των 2 μεγαλύτερων δυνάμεων λ^2, λ και τη σταθερά:

$$\begin{aligned}
\gamma_2 &= 1, \\
\gamma_1 &= -(\alpha_{11} + \alpha_{22}) \\
\gamma_0 &= \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}
\end{aligned}$$

Ορισμοί με ορίζουσες

Χαρακτηριστικό πολυώνυμο και ιδιοτιμές

- *Χαρακτηριστικό πολυώνυμο* ενός μητρώου $A \in \mathbb{R}^{n\times n}$, ονομάζεται το πολυώνυμο βαθμού n με το οποίο ισούται η ορίζουσα

$$\begin{aligned} p(\lambda) &= \det(\lambda I - A) \text{ ή «ισοδύναμα»} \\ \hat{p}(\lambda) &= \det(A - \lambda I) = (-1)^n p(\lambda). \end{aligned}$$

Ορισμοί με ορίζουσες

Χαρακτηριστικό πολυώνυμο και ιδιοτιμές

- *Χαρακτηριστικό πολυώνυμο* ενός μητρώου $A \in \mathbb{R}^{n\times n}$, ονομάζεται το πολυώνυμο βαθμού n με το οποίο ισούται η ορίζουσα

$$\begin{aligned} p(\lambda) &= \det(\lambda I - A) \text{ ή «ισοδύναμα»} \\ \hat{p}(\lambda) &= \det(A - \lambda I) = (-1)^n p(\lambda). \end{aligned}$$

- Αν γράψουμε

$$p(\lambda) = \lambda^n + \gamma_{n-1}\lambda^{n-1} + \cdots + \gamma_1\lambda + \gamma_0$$

ισχύει ότι $\gamma_{n-1} = -\mathrm{trace}(A)$, $\gamma_0 = (-1)^n \det(A)$

Ορισμοί με ορίζουσες

Χαρακτηριστικό πολυώνυμο και ιδιοτιμές

- *Χαρακτηριστικό πολυώνυμο* ενός μητρώου $A \in \mathbb{R}^{n\times n}$, ονομάζεται το πολυώνυμο βαθμού n με το οποίο ισούται η ορίζουσα

$$\begin{aligned} p(\lambda) &= \det(\lambda I - A) \text{ ή «ισοδύναμα»} \\ \hat{p}(\lambda) &= \det(A - \lambda I) = (-1)^n p(\lambda). \end{aligned}$$

- Αν γράψουμε

$$p(\lambda) = \lambda^n + \gamma_{n-1}\lambda^{n-1} + \cdots + \gamma_1\lambda + \gamma_0$$

  ισχύει ότι $\gamma_{n-1} = -\mathrm{trace}(A)$, $\gamma_0 = (-1)^n \det(A)$
- προσέξτε ότι αν το μέγεθος του μητρώου είναι περιττό, οι συντελεστές των p και $\hat{p}$ έχουν διαφορετικό πρόσημο.

Ορισμοί με ορίζουσες

Χαρακτηριστικό πολυώνυμο και ιδιοτιμές

- *Χαρακτηριστικό πολυώνυμο* ενός μητρώου $A \in \mathbb{R}^{n\times n}$, ονομάζεται το πολυώνυμο βαθμού n με το οποίο ισούται η ορίζουσα

$$\begin{aligned} p(\lambda) &= \det(\lambda I - A) \text{ ή «ισοδύναμα»} \\ \hat{p}(\lambda) &= \det(A - \lambda I) = (-1)^n p(\lambda). \end{aligned}$$

- Αν γράψουμε

$$p(\lambda) = \lambda^n + \gamma_{n-1}\lambda^{n-1} + \cdots + \gamma_1\lambda + \gamma_0$$

  ισχύει ότι $\gamma_{n-1} = -\mathrm{trace}(A)$, $\gamma_0 = (-1)^n \det(A)$
- προσέξτε ότι αν το μέγεθος του μητρώου είναι περιττό, οι συντελεστές των p και $\hat{p}$ έχουν διαφορετικό πρόσημο.
- Οι ρίζες του χ.π. ονομάζονται ΙΔΙΟΤΙΜΕΣ (του μητρώου).

Ορισμοί με ορίζουσες

Χαρακτηριστικό πολυώνυμο και ιδιοτιμές

- *Χαρακτηριστικό πολυώνυμο* ενός μητρώου $A \in \mathbb{R}^{n\times n}$, ονομάζεται το πολυώνυμο βαθμού n με το οποίο ισούται η ορίζουσα

$$\begin{aligned} p(\lambda) &= \det(\lambda I - A) \text{ ή «ισοδύναμα»} \\ \hat{p}(\lambda) &= \det(A - \lambda I) = (-1)^n p(\lambda). \end{aligned}$$

- Αν γράψουμε

$$p(\lambda) = \lambda^n + \gamma_{n-1}\lambda^{n-1} + \cdots + \gamma_1\lambda + \gamma_0$$

  ισχύει ότι $\gamma_{n-1} = -\mathrm{trace}(A)$, $\gamma_0 = (-1)^n \det(A)$
- προσέξτε ότι αν το μέγεθος του μητρώου είναι περιττό, οι συντελεστές των p και $\hat{p}$ έχουν διαφορετικό πρόσημο.
- Οι ρίζες του χ.π. ονομάζονται ΙΔΙΟΤΙΜΕΣ (του μητρώου).
- Το πολυώνυμο έχει n ρίζες και άρα n ιδιοτιμές. Μερικές ή όλες μπορεί να είναι μεταξύ τους ίσες (πολλαπλές). Το σύνολο των ιδιοτιμών λέγεται φάσμα του A.

Μια κρίσιμη ιδιότητα του Χ.Π. - αφετηρία της θεωρίας μητρώων (Cayley' 1858)

Έστω ότι $A \in \mathbb{R}^{2\times 2}$, οπότε

$$\det(\lambda I - A) \quad = \quad \lambda^2 - (\alpha_{11} + \alpha_{22})\lambda + \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}.$$

Αντικαθιστούμε το λ με το A:

$$\begin{aligned}
p(A) &= A^2 - (\alpha_{11}+\alpha_{22})A + (\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12})I \\
&= \begin{pmatrix} \alpha_{11}^2 + \alpha_{12}\alpha_{21} & \alpha_{11}\alpha_{12} + \alpha_{12}\alpha_{22} \\ \alpha_{21}\alpha_{11} + \alpha_{22}\alpha_{21} & \alpha_{21}\alpha_{12} + \alpha_{22}^2 \end{pmatrix} - (\alpha_{11}+\alpha_{22}) \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} + \begin{pmatrix} (\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}) & 0 \\ 0 & (\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}) \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}
\end{aligned}$$

Μια κρίσιμη ιδιότητα του Χ.Π. - αφετηρία της θεωρίας μητρώων (Cayley' 1858)

Έστω ότι $A \in \mathbb{R}^{2\times 2}$, οπότε

$$\det(\lambda I - A) \quad = \quad \lambda^2 - (\alpha_{11} + \alpha_{22})\lambda + \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}.$$

Αντικαθιστούμε το λ με το A:

$$\begin{aligned} p(A) &= A^2 - (\alpha_{11} + \alpha_{22})A + (\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12})I \\ &= \begin{pmatrix} \alpha_{11}^2 + \alpha_{12}\alpha_{21} & \alpha_{11}\alpha_{12} + \alpha_{12}\alpha_{22} \\ \alpha_{21}\alpha_{11} + \alpha_{22}\alpha_{21} & \alpha_{21}\alpha_{12} + \alpha_{22}^2 \end{pmatrix} - (\alpha_{11} + \alpha_{22}) \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} + \begin{pmatrix} (\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}) & 0 \\ 0 & (\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}) \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

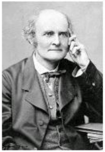
Φαίνεται ότι $p(A) = 0$. Αυτή η ιδιότητα ισχύει για **οποιοδήποτε τετραγωνικό μητρώο,** $A \in \mathbb{C}^{n\times n}$.

Θεώρημα Cayley-Hamilton

Έστω το χαρακτηριστικό πολυώνυμο $p(\lambda) = \det(\lambda I - A)$. Τότε

$$\boxed{p(A) = 0_{n\times n}}.$$

A Memoir on the Theory of Matrices
Author(s): Arthur Cayley
Source: *Philosophical Transactions of the Royal Society of London*, Vol. 148 (1858), pp. 17-37
Published by: The Royal Society
Stable URL: http://www.jstor.org/stable/108649
Accessed: 03/05/2010 17:45



II. *A Memoir on the Theory of Matrices.* By Arthur Cayley, *Esq., F.R.S.*

Received December 10, 1857,—Read January 14, 1858.

The term matrix might be used in a more general sense, but in the present memoir I consider only square and rectangular matrices, and the term matrix used without qualification is to be understood as meaning a square matrix; in this restricted sense, a set of quantities arranged in the form of a square, *e. g.*

$$\left(\begin{array}{lll} a\,, & b\,, & c \\ a'\,, & b'\,, & c' \\ a''\,, & b''\,, & c'' \end{array}\right)$$

is said to be a matrix. The notion of such a matrix arises naturally from an abbreviated notation for a set of linear equations, viz. the equations

$$\begin{aligned} X &= ax + by + cz, \\ Y &= a'x + b'y + c'z, \\ Z &= a''x + b''y + c''z, \end{aligned}$$

may be more simply represented by

$$(X, Y, Z) = \left(\begin{array}{lll} a\,, & b\,, & c \\ a'\,, & b'\,, & c' \\ a''\,, & b''\,, & c'' \end{array}\right)(x, y, z),$$

ΠΡΟΣΟΧΗ

Η $\det$ είναι συνάρτηση όλων των στοιχείων του μητρώου, π.χ. για $n = 2$, όλων των

$$\alpha_{11} - \lambda, \alpha_{12}, \alpha_{21}, \alpha_{22} - \lambda$$

Επομένως ΔΕΝ μπορούμε να πούμε ότι $p(A) = \det(A - AI)\det(0) = 0$.

Δείτε την Wikipedia για μερικές αποδείξεις του θεωρήματος (υπάρχουν πολλές!)
Θα έχουμε την ευκαιρία να τη συζητήσουμε σε επόμενη διάλεξη ....

Μια (ακόμα) ιδιομορφία των μητρώων

... επιπλέον των $AB \neq BA$, $AB = 0$ ακόμα και αν $A \neq 0, B \neq 0$, ... που έπεται από το θεώρημα Cayley-Hamilton:

Για τις δυνάμεις $A^n = -\gamma_{n-1}A^{n-1} - \gamma_{n-2}A^{n-2} - \cdots + \gamma_1 A + \gamma_0 I.$

Δηλ. για κάθε μητρώο $A \in \mathbb{R}^{n\times n}$, οι δυνάμεις μεγαλύτερες από $n-1$ μπορούν να γραφτούν ως γραμμικός συνδυασμός χαμηλοτέρων δυνάμεων!

Για το αντίστροφο (αν υπάρχει)

$$\begin{aligned} 0 &= A^n + \gamma_{n-1}A^{n-1} + \cdots + \gamma_1 A + \gamma_0 I \\ &= A^{-1}\left(A^n + \gamma_{n-1}A^{n-1} + \cdots + \gamma_1 A + \gamma_0 I\right) \end{aligned}$$

επομένως

$$A^{-1} = -\frac{1}{\gamma_0}\left(A^{n-1} + \gamma_{n-1}A^{n-2} + \cdots + \gamma_1 I\right)$$

Ιδιοτιμές χωρίς ορίζουσες (κατά S. Axler)

Για κάθε $A \in \mathbb{C}^{n\times n}$, ορίζεται ως ιδιοτιμή κάθε βαθμωτός λ για τον οποίο το μητρώο $A-\lambda I$ δεν είναι αντιστρέψιμο και ιδιοχώρος του A που αντιστοιχεί στο λ ο μηδενόχωρος του $A-\lambda I$. **Θα δείξουμε ότι υπάρχει τουλάχιστον μια ιδιοτιμή.**

- Δίνεται $A \in \mathbb{R}^{n\times n}$ και μη μηδενικό $x \in \mathbb{R}^n$. Τότε τα $n+1$ διανύσματα
$$x, Ax, A^2x, ..., A^nx$$
είναι οπωσδήποτε γραμμικά εξαρτημένα (γιατί;).
- Επομένως υπάρχουν συντελεστές $\gamma_0, \cdots, \gamma_n$ όχι όλοι μηδέν τέτοιοι ώστε
$$0 \quad = \quad \gamma_0 x + \gamma_1 Ax + \cdots \gamma_n A^n x = p(A)x$$
- Δεν γνωρίζουμε ακριβώς το βαθμό του πολυωνύμου! Μπορεί να είναι μικρότερος του n.
- Αφού $p(A)x = 0$ και $x \neq 0$, το μητρώο $p(A)$ δεν είναι αντιστρέψιμο.
- Αντικαθιστώντας το A με βαθμωτή μεταβλητή ζ, $p(\zeta) = 0$, και έστω ότι $\mathrm{deg} p = m \leq n$ και ότι οι ρίζες του είναι $\rho_1, ..., \rho_m \in \mathbb{C}$. Τότε
$$p(z) = \gamma(\zeta-\rho_1)(\zeta-\rho_2)(\zeta-\rho_m).$$
όπου γ είναι ο συντελεστής του μεγιστοβάθμιου όρου.
- Έπεται ότι ένας τουλάχιστον από τους παράγοντες του
$$p(A) \quad = \quad \gamma(A-\rho_1 I)(A-\rho_2 I), (A-\rho_m I)$$
δεν είναι αντιστρέψιμος. Αν είναι ο παράγοντας $A-\rho_j I$, το ρ_j θα είναι ιδιοτιμή.

Ανακεφαλαίωση

Βασικός ορισμός

Για ένα μητρώο $A \in \mathbb{R}^{n\times n}$ και μεταβλητή λ, η ορίζουσα $p(\lambda) = \det(\lambda I - A)$ είναι πολυώνυμο βαθμού n. Λέγεται χαρακτηριστικό πολυώνυμο και αν το γράψουμε

$$p(\lambda) \quad = \quad \lambda^n + \gamma_{n-1}\lambda^{n-1} + \cdots + \gamma_1\lambda + \gamma_0$$

τότε $\gamma_{n-1} = -\mathrm{trace}(A)$ και $\gamma_0 = (-1)^n \det(A)$.

Προσοχή:

- Το πολυώνυμο έχει ακριβώς n ρίζες $\lambda_1, \cdots \lambda_n \in \mathbb{C}^n$ που αποκαλούμε **ιδιοτιμές** του A.
- Για κάθε ιδιοτιμή, λ, ο μηδενόχωρος $\mathrm{null}(\lambda I - A)$ λέγεται **ιδιόχωρος** και έχει διάσταση τουλάχιστον 1. Κάθε διάνυσμα βάσης του χώρου ονομάζεται **ιδιοδιάνυσμα**.
- Ισχύει ότι $p(A) = 0$ **(θεώρημα Cayley-Hamilton)**.
- τα ιδιοζεύγη (ιδιοτιμές και ιδιοδιανύσματα) μητρώων παίζουν πολύ σημαντικό ρόλο σε πληθώρα εφαρμογών.

Πολλαπλασιασμός με βαθμωτό

Έστω ότι $Ax = \lambda x$ και έστω βαθμωτός $\gamma \neq 0$. Τότε

$$\gamma(Ax) \quad = \quad \gamma(\lambda x).$$

Δύο ερμηνείες:

$(\gamma A)x = (\gamma\lambda)x$: αν πολλαπλασιάσουμε το μητρώο με βαθμωτό, το ιδιοδιάνυσμα παραμένει αμετάβλητο και η ιδιοτιμή πολλαπλασιάζεται με το βαθμωτό.

$A(\gamma x) = \lambda(\gamma x)$: αν πολλαπλασιάσουμε το ιδιοδιάνυσμα με βαθμωτό, έχουμε πάλι ιδιοδιάνυσμα και η ιδιοτιμή παραμένει αμετάβλητη.

Επομένως, όλα τα διανύσματα συγγραμικά με το ιδιοδιάνυσμα είναι και αυτά ιδιοδιανύσματα. Συνήθως κανονικοποιούμε ώστε να λαμβάνουμε ως "εκπρόσωπο της κατεύθυνσης" το διάνυσμα που έχει μήκος 1.

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου.

.

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου.
2. Υπολογισμός των ριζών του (που είναι οι ιδιοτιμές).

.

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου.
2. Υπολογισμός των ριζών του (που είναι οι ιδιοτιμές).

.

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου. Δαπανηρή και αριθμητικά προβληματική για μεγάλα προβλήματα.
2. Υπολογισμός των ριζών του (που είναι οι ιδιοτιμές).

Για κάθε ιδιοτιμή λ_j, συνήθως ενδιαφέρει ένα σύνολο από γραμμικά ανεξάρτητα ιδιοδιανύσματα που είναι βάση για το $\mathrm{null}(\lambda_j I - A)$. Αυτά είναι ειδικές λύσεις του $(\lambda_j I - A)x = 0$. Συνήθως τα διανύσματα επιλέγονται κανονικοποιημένα.

.

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου. Δαπανηρή και αριθμητικά προβληματική για μεγάλα προβλήματα.
2. Υπολογισμός των ριζών του (που είναι οι ιδιοτιμές). Αν $n \geq 5$ τότε δεν υπάρχει αναλυτικός τύπος υπολογισμού των ριζών και πρέπει να χρησιμοποιηθεί (επαναληπτικός) αλγόριθμος προσέγγισης ριζών πολυωνύμου[1].
3. Για κάθε ιδιοτιμή, επίλυση του $(A - \lambda I)x = 0$ και επιλογή των ιδιοδιανυσμάτων (ανεύρεση των ειδικών λύσεων).

.

[1] Η αδυναμία αυτή αποτελεί ένα βασικό εύρημα των Μαθηματικών του 19ου αιώνα (Abel, Rufffini, και Galois)

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου. Δαπανηρή και αριθμητικά προβληματική για μεγάλα προβλήματα.
2. Υπολογισμός των ριζών του (που είναι οι ιδιοτιμές). Αν $n \geq 5$ τότε δεν υπάρχει αναλυτικός τύπος υπολογισμού των ριζών και πρέπει να χρησιμοποιηθεί (επαναληπτικός) αλγόριθμος προσέγγισης ριζών πολυωνύμου[1].
3. Για κάθε ιδιοτιμή, επίλυση του $(A - \lambda I)x = 0$ και επιλογή των ιδιοδιανυσμάτων (ανεύρεση των ειδικών λύσεων).

.

[1] Η αδυναμία αυτή αποτελεί ένα βασικό εύρημα των Μαθηματικών του 19ου αιώνα (Abel, Rufffini, και Galois)

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου.
2. Υπολογισμός των ριζών του (που είναι οι ιδιοτιμές).

Τα παραπάνω είναι μόνο για μικρά προβλήματα και δεν χρησιμοποιείται: Στην πράξη (π.χ. στη MATLAB) χρησιμοποιούνται ειδικοί αλγόριθμοι (συνάρτηση `eig`) για την εύρεση των ιδιοτιμών (πχ. αλγόριθμος QR. Μάλιστα, το χ.π. υπολογίζεται μετά από κλήση στην `eig` για τον υπολογισμό των ιδιοτιμών και με υπολογισμό των συντελεστών της δυναμομορφής από την παραγοντοποιημένη μορφή. Δηλαδή, η διαδικασία που ακολουθείται έχει την αντίθετη φορά από αυτήν που χρησιμοποιούμε εδώ[1].

[1] Η αριθμητική επίλυση του ΑΠΙ είναι ένα σημαντικό επιστημονικό αντικείμενο των περιοχών της Υπολογιστικής Γραμμικής Άλγεβρας και της Αριθμητικής Ανάλυσης.

Παραδείγματα

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

$p(\lambda) = (\lambda - 1)(\lambda - 3)$, $\Lambda(A) = \{1, 3\}$,
$\lambda_1 = 1, \mu_1 = 1$, $\lambda_2 = 3, \mu_2 = 1$.

Παραδείγματα

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

$p(\lambda) = (\lambda - 1)(\lambda - 3), \Lambda(A) = \{1, 3\},$
$\lambda_1 = 1, \mu_1 = 1, \lambda_2 = 3, \mu_2 = 1.$

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

$p(\lambda) = (\lambda - 1)^2, \Lambda(A) = \{1\},$
$\lambda_1 = 1, \mu_1 = 2.$

Παραδείγματα

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

$p(\lambda) = (\lambda - 1)(\lambda - 3), \Lambda(A) = \{1, 3\},$
$\lambda_1 = 1, \mu_1 = 1, \lambda_2 = 3, \mu_2 = 1.$

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

$p(\lambda) = (\lambda - 1)^2, \Lambda(A) = \{1\},$
$\lambda_1 = 1, \mu_1 = 2.$

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

$p(\lambda) = (\lambda - 1)(\lambda - 2)^2,$
$\Lambda(A) = \{1, 2\}, \lambda_1 = 1, \mu_1 = 1,$
$\lambda_2 = 2, \mu_2 = 2.$

Παραδείγματα

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

$p(\lambda) = (\lambda - 1)(\lambda - 3)$, $\Lambda(A) = \{1, 3\}$,
$\lambda_1 = 1, \mu_1 = 1$, $\lambda_2 = 3, \mu_2 = 1$.

$$A = \begin{pmatrix} 4 & 2i & 0 & -2i \\ 2i & 4 & -2i & 0 \\ 0 & -2i & 4 & 2i \\ -2i & 0 & 2i & 4 \end{pmatrix}$$

$p(\lambda) = (\lambda - 4)^2(\lambda - 4 - 4i)(\lambda - 4 + 4i)$,
$\Lambda(A) = \{2, 4 + 4i, 4 - 4i\}$, $\lambda_1 = 4, \mu_1 = 2$,
$\lambda_2 = 4 + 4i, \mu_2 = 1$, $\lambda_3 = 4 - 4i, \mu_3 = 1$.

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

$p(\lambda) = (\lambda - 1)^2$, $\Lambda(A) = \{1\}$,
$\lambda_1 = 1, \mu_1 = 2$.

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

$p(\lambda) = (\lambda - 1)(\lambda - 2)^2$,
$\Lambda(A) = \{1, 2\}$, $\lambda_1 = 1, \mu_1 = 1$,
$\lambda_2 = 2, \mu_2 = 2$.

Παραδείγματα

Συμβολισμοί: για κάθε διακριτή ιδιοτιμή λ_j, η πολλαπλότητά της συμβολίζεται με μ_j.

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

$$p(\lambda) = \lambda^4 - 8\lambda^3 + 16\lambda^2 = \lambda^2(\lambda - 4)^2$$

επομένως

$$\lambda_1 = 0, \mu_1 = 2, \lambda_2 = 4, \mu_2 = 2.$$

και

$$\Lambda(A) = \{0, 4\}$$

Συνοδευτικό μητρώο

Από τα πολυώνυμα στα μητρώα

Το $p(z) = \alpha_0 + \alpha_1 \lambda + \cdots + \alpha_{n-1}\lambda^{n-1} + \lambda^n$ είναι το χ.π. των

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & 0 & \cdots & -\alpha_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -\alpha_{n-1} \end{pmatrix}$$ καθώς και του $A^\top$.

Ορισμός και χρήση

Το A ονομάζεται συνοδευτικό μητρώο (companion matrix) του πολυωνύμου p.

(υπολ. ριζών πολυωνύμου βαθμού n) $\equiv$ (υπολ. ιδιοτιμών συνοδευτικού μητρώου)

Επαληθεύστε ότι

1. ότι κάθε ρίζα ζ του p είναι ιδιοτιμή του A
2. ... με αντίστοιχο αριστερό ιδιοδιάνυσμα το $(1, \zeta, \zeta^2, \ldots, \zeta^{n-1})$.

Παραδείγματα

$p(\lambda) = \lambda^2 + 2\lambda + 1$ και

$$A = \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix} \Rightarrow \det(\lambda I - A) = \lambda(\lambda + 2) + 1 = \lambda^2 + 2\lambda + 1$$

$p(\lambda) = \lambda^3 + 3\lambda^2 + 2\lambda + 1$ και

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & -2 \\ 0 & 1 & -3 \end{pmatrix} \Rightarrow \det(\lambda I - A) = \lambda(\lambda(\lambda + 3) + 2) + 1 = \lambda^3 + 3\lambda^2 + 2\lambda + 1$$

Ιδιοδιανύσματα

$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, Βρήκαμε ότι $\lambda_1 = 1, \mu_1 = 1, \lambda_2 = 3, \mu_2 = 1$.

Θέτουμε $A(1) = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}$ και επιλύουμε $A(1)x = 0$:

Η τάξη $\mathrm{rank}(A(1)) = 1$, και βρίσκουμε από τις ειδικές λύσεις ότι μητρώο μηδενοχώρου για το $A(1)$ είναι το $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ Κανονικοποιούμε και επιλέγουμε ιδιοδιάνυσμα για το λ_1: $x_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

Θέτουμε $A(3) = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ και επιλύουμε $A(3)x = 0$: Η τάξη $\mathrm{rank}(A(3)) = 1$, και βρίσκουμε από τις ειδικές λύσεις ότι μητρώο μηδενοχώρου για το $A(3)$ είναι το $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ Κανονικοποιούμε και επιλέγουμε ιδιοδιάνυσμα για το λ_2: $x_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Παρατηρήσεις Ι

Αν πολλαπλασιάσουμε $A(x_1, x_2) = (\lambda_1 x_1, \lambda_2 x_2)$

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

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

Δείτε ότι αν $X = (x_1, x_2)$ και $\mathrm{fl} = \mathrm{diag}(\lambda_1, \lambda_2)$, τότε

$$AX = X\Lambda$$

Το X είναι αντιστρέψιμο γιατί οι ιδιοτιμές είναι διαφορετικές (βλ. ▸ επόμενες διαφάνειες.) Είναι επίσης ορθογώνιο (λόγω συμμετρίας του A, θα το δούμε και αυτό αργότερα).

Παρατηρήσεις II

Ισχύει επομένως ότι

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

Δηλαδή

$$X^{-1}AX = X^{\top}AX = \Lambda.$$

Προσοχή

Ο μετασχηματισμός $A \to X^{-1}AX$ λέγεται μετασχηματισμός ομοιότητας του A με το X. Όταν το X είναι ορθογώνιο, όπως εδώ, χαρακτηρίζεται ως ορθογώνιος μετασχηματισμός ομοιότητας. Στην περίπτωση που εξετάζουμε, η εφαρμογή του μετασχηματισμού, μετατρέπει το A σε διαγώνιο μητρώο, που περιέχει τις ιδιοτιμές. Λέμε ότι το μητρώο είναι διαγωνιοποιήσιμο. Ένα μητρώο με n γραμμικά ανεξάρτητα ιδιοδιανύσματα διαγωνιοποιείται με το μητρώο των ιδιοδιανυσμάτων του.

Ιδιοδιανύσματα I

Στη συνέχεια συμβολίζουμε $A(\lambda) = \lambda I - A$.

$$A = \begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix}$$, Βρήκαμε ότι $\lambda_1 = 0, \mu_1 = 2, \lambda_2 = 4, \mu_2 = 2$.

Θέτουμε $A(0) = -A$ και επιλύουμε $-Ax = 0$:

Η τάξη $\mathrm{rank}(A(0)) = 2$, και βρίσκουμε από τις ειδικές λύσεις ότι μητρώο μηδενοχώρου για το $A(0)$ είναι το $\begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$ Κανονικοποιούμε και επιλέγουμε ιδιοδιανύσματα για το $\lambda_1 = 0$: $x_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, x_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}$

Ιδιοδιανύσματα II

Θέτουμε $A(4) = 4I - A$ και επιλύουμε $(4I - A)x = 0$:

Η τάξη $\mathrm{rank}(A(4)) = 2$, και βρίσκουμε από τις ειδικές λύσεις ότι μητρώο μηδενοχώρου για το $A(4)$ είναι το $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$

Κανονικοποιούμε και επιλέγουμε ιδιοδιανύσματα για το $\lambda_2 = 4$: $x_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$,

$x_4 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$

Παρατηρήστε ότι τα ιδιοδιανύσματα είναι κάθετα μεταξύ τους.

Ιδιοδιανύσματα III

Για το $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. βρήκαμε ότι $\Lambda(A) = \{1\}, \lambda_1 = 1, \mu_1 = 2$.

Θέτουμε $A(1) = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}$ και επιλύουμε $A(1)x = 0$:

Η τάξη $\mathrm{rank}(A(1)) = 1$, επομένως η διάσταση του $\mathrm{null}(A(1)) = 1$.

Παρατηρούμε ότι $\mathrm{null}(A(1)) < \mu_1$.

Από τις ειδικές λύσεις βρίσκουμε ότι το μητρώο μηδενοχώρου για το $A(1)$ είναι το $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ Κανονικοποιούμε, οπότε το ιδιοδιάνυσμα για το λ_1 είναι $x_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Προσοχή: Όλα τα ιδιοδιανύσματα του A για το $\lambda_1 = 1$ είναι πολλαπλάσια του x_1. Δεν υπάρχουν άλλα! Εφόσον η μόνη διακριτή ιδιοτιμή είναι το $\lambda_1 = 1$, αυτό είναι το μοναδικό ιδιοδιάνυσμα του A.

Γραμμική ανεξαρτησία ιδιοδιανυσμάτων

Πρόταση

Τα ιδιοδιανύσματα που αντιστοιχούν σε διαφορετικές ιδιοτιμές είναι γραμμικά ανεξάρτητα.

Απόδειξη[2] Έστω ότι $Ax_1 = \lambda_1 x_1, Ax_2 = \lambda_2 x_2$ όπου $\lambda_1 \neq \lambda_2$. Τότε αν υπάρχουν μη μηδενικά γ_1, γ_2 τ.ώ.

$$0 \quad = \quad \gamma_1 x_1 + \gamma_2 x_2 = \gamma_1 A x_1 + \gamma_2 A x_2 = \gamma_1 \lambda_1 x_1 + \gamma_2 \lambda_2 x_2.$$

Ισχύει επίσης ότι $0 = \lambda_2(\gamma_1 x_1 + \gamma_2 x_2) = \lambda_2 \gamma_1 x_1 + \lambda_2 \gamma_2 x_2$ επομένως

$$\gamma_1 \lambda_1 x_1 + \cancel{\gamma_2 \lambda_2 x_2} \quad = \quad \lambda_2 \gamma_1 x_1 + \cancel{\lambda_2 \gamma_2 x_2}$$

δηλ. $0 = \gamma_1 \lambda_1 x_1 - \lambda_2 \gamma_1 x_1$ επομένως $\lambda_1 = \lambda_2$, άρα άτοπο.

Πόρισμα

Αν ένα μητρώο $A \in \mathbb{C}^{n \times n}$ έχει n διαφορετικές ιδιοτιμές, τότε θα έχει n γραμμικά ανεξάρτητα ιδιοδιανύσματα. Κατά συνέπεια, τότε τα ιδιοδιανύσματα παρέχουν μία βάση (ιδιοβάση) με πολλές ευχάριστες ιδιότητες.

[2] Δείτε την πλήρη απόδειξη στο βιβλίο του Strang:

Διαγωνιοποίηση μητρώου I

Έστω ότι γνωρίζουμε n ιδιοζεύγη (λ_j, x_j) του $A \in \mathbb{C}^{n\times n}$. Τότε

$$Ax_1 = \lambda_1 x_1, Ax_2 = \lambda_2 x_2, \ldots, Ax_n = \lambda_n x_n$$

Συλλέγουμε και διατυπώνουμε με μητρώα:

$$A[x_1, x_2, \ldots, x_n] = [x_1 \ldots, x_n] \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

Επομένως

$$AX = X\Lambda,$$

όπου $\Lambda = \mathrm{diag}([\lambda_1, \ldots, \lambda_n])$ και $X = [x_1 \ldots, x_n]$.

Διαγωνιοποίηση μητρώου II

ΠΡΟΣΕΞΤΕ Αν το X είναι αντιστρέψιμο,

$$X^{-1}AX = \Lambda,$$

δηλ. χρησιμοποιώντας τα μητρώα X (με στήλες τα δεξιά ιδιοδιανύσματα) και X^{-1} (με στήλες του X^{-*} τα αριστερά ιδιοδιανύσματα) διαγωνιοποιήσαμε το A.

Κάθε μητρώο για το οποίο υπάρχουν n γραμμικά ανεξάρτητα ιδιοδιανύσματα χαρακτηρίζεται ως διαγωνιοποιήσιμο, ειδάλλως λέγεται μη διαγωνιοποιήσιμο.

Ιδιοδιανύσματα: Δεξιά και Αριστερά

Αν $A \in \mathbb{C}^{n\times n}$ και $\lambda \in \Lambda(A)$ τότε

$$\det(\lambda I - A) = 0$$

Αν συμβολίσουμε με r την τάξη του $\lambda I - A$,

$$r = \mathrm{rank}(\lambda I - A) \leq n - 1.$$

Αφού το μητρώο είναι τετραγωνικό, οι διαστάσεις του μηδενόχωρου και του αριστερού μηδενόχωρου του $A - \lambda I$ είναι ίσες επομένως

$$1 \leq \dim(\mathrm{null}(\lambda I - A)) = \dim(\mathrm{null}(\lambda I - A)^*) = n - r \leq n - 1$$

Ιδιοδιανύσματα είναι τα μέλη των μηδενοχώρων αυτών:

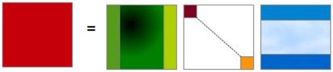
δεξιά ιδιοδιανύσματα για την ιδιοτιμή λ είναι τα διανύσματα $x \in \mathrm{null}(\lambda I - A)$ δηλ. τ.ώ. $Ax = \lambda x$.

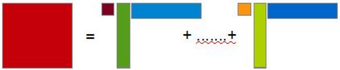
αριστερά ιδιοδιανύσματα για την ιδιοτιμή λ είναι τα διανύσματα $y \in \mathrm{null}(\lambda I - A)^*$, δηλ. τ.ώ. $y^* A = \lambda y^*$.

Ανάπτυγμα μητρώου συναρτήσει ιδιοτιμών και ιδιοδιανυσμάτων

Αν το μητρώο A διαθέτει n γ.α. ιδιοδιοδιανύσματα, τότε $A = X\Lambda X^{-1}$ όπου Λ το διαγώνιο μητρώο των ιδιοτιμών και X το μητρώο των ιδιοδιανυσμάτων.
Θυμηθείτε τη γραφή γινομένου μητρώων ως άθροισμα μητρώων πρώτης τάξης (στήλη του 1ου επί γραμμή του 2ου).
Θέτουμε για ευκολία $Y = (X^{-1})^*$, δηλ. Y είναι το συζυγές αντίστροφο του X, τότε μπορούμε να γράψουμε το A με βάση το φασματικό ανάπτυγμα.

$$\begin{aligned} A &= X\Lambda Y^* \\ &= \lambda_1 x_1 y_1^* + \lambda_2 x_2 y_2^* + \cdots + \lambda_n x_n y_n^* \end{aligned}$$





Ιδιόχωροι και γεωμετρική πολλαπλότητα ιδιοτιμής

Για κάθε διαφορετική ιδιοτιμή λ_j:

- ο μηδενόχωρος $\mathrm{null}(\lambda_j I - A)$ λέγεται ιδιόχωρος της ιδιοτιμής λ_j.
- Η διάσταση του ιδιόχωρου λέγεται γεωμετρική πολλαπλότητα της ιδιοτιμής λ_j.
- Για κάθε ιδιοτιμή $1 \leq$ γεωμ. πολλ/τα$(\lambda_j) \leq$ αλγεβρ. πολλ/τα.(λ_j)
- Κάθε ιδιοτιμή για την οποία

  $$1 < \text{αλγεβρ. πολλ/τα.} = \text{αλγεβρ. πολλ/τα.}$$

  αποκαλείται ημιαπλή.
- Αν κάποια ιδιοτιμή έχει γεωμετρική πολλαπλότητα μικρότερη της αλγεβρικής, δηλ.

  $$1 \leq \text{γεωμ. πολλ/τα}(\lambda_j) < \text{αλγεβρ. πολλ/τα.}(\lambda_j)$$

  χαρακτηρίζεται ως ελλειμματική[3]. Αν υπάρχει έστω και μια ελλειμματική, το μητρώο λέγεται ελλειμματικό.
- Τα ελλειμματικά μητρώα είναι μη διαγωνιοποιήσιμα.

[3] defective

αλγ. πολλ.	γεω. πολλ.	ΑΠ-ΓΠ	ιδιοτιμή	μητρώο
1	1	0	απλή	διαγωνιοποιήσιμο, αν όλες οι ιδιοτιμές όπως εδώ
>1	>1	0	ημιαπλή	
>1	>1	>0	ελλειμματική	ελλειμματικό (μη διαγωνιοποιήσιμο)

Παραδείγματα

Για τα παρακάτω τριγωνικά μητρώα (ιδιοτιμές στη διαγώνιο)

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

A	*διαγωνιοποιήσιμο*	*διακριτές ιδιοτιμές*	$X = \begin{pmatrix} 1 & \frac{2\sqrt{5}}{5} & \frac{7\sqrt{69}}{69} \\ 0 & \frac{\sqrt{5}}{5} & \frac{4\sqrt{69}}{69} \\ 0 & 0 & \frac{2\sqrt{69}}{69} \end{pmatrix}$
B	*διαγωνιοποιήσιμο*	$B = \begin{pmatrix} I_2 & B_{1:2,2} \\ 0 & 1 \end{pmatrix}$	$X = \begin{pmatrix} 1 & 0 & -\frac{\sqrt{2}}{2} \\ 0 & 1 & -\frac{\sqrt{2}}{2} \\ 0 & 0 & 0 \end{pmatrix}$
C	*μη διαγωνιοπ.*	*μορφή Jordan*	
D	*μη διαγωνιοπ.*	$D = \begin{pmatrix} D_{1:2,1:2} & D_{1:2,2} \\ 0 & 1 \end{pmatrix}$	
$A_{1:2,1:2}$	*διαγωνιοποιήσιμο*	*διακριτές ιδιοτιμές*	

Περί διαγωνιοποιήσιμων μητρώων

Κάθε μητρώο είναι διαγωνιοποιήσιμο ή μη διαγωνιοποιήσιμο (ελλειμματικό).

Ερώτημα Πότε είναι ένα μητρώο διαγωνιοποιήσιμο;

- Αν και μόνον αν υπάρχουν n γ.α. ιδιοδιανύσματα.
- Για παράδειγμα, όταν υπάρχουν n διαφορετικές ιδιοτιμές.

Προσοχή: Αν υπάρχουν επαναλαμβανόμενες ιδιοτιμές δεν εξασφαλίζεται ούτε η διαγωνιοποιησιμότητα ούτε η μη διαγωνιοποιησιμότητα. Εξαρτάται από το μητρώο, επομένως απαιτείται περισσότερη διερεύνηση.

- Προσέξτε ότι αυτή η κατηγοριοποίηση αφορά την ύπαρξη αντιστρέψιμου X ώστε $X^{-1}AX$ να είναι διαγώνιο. Η αντιστρεψιμότητα ή μη του A αφορά μόνον τοA (προφανώς!)
- Έστω $A(\alpha,\beta) = \begin{pmatrix} \alpha & 1 \\ 0 & \beta \end{pmatrix}$. Τότε
  - Αν $\alpha = 1, \beta = 2$, τότε το A είναι αντιστρέψιμο και διαγωνοποιήσιμο.
  - Αν $\alpha = 1, \beta = 1$, τότε το A είναι αντιστρέψιμο και μη διαγωνοποιήσιμο.
  - Αν $\alpha = 1, \beta = 0$, τότε το A είναι μη αντιστρέψιμο και διαγωνοποιήσιμο.
  - Αν $\alpha = 0, \beta = 0$, τότε το A είναι μη αντιστρέψιμο και μη διαγωνοποιήσιμο.
- (Ευτυχώς) στο (τεράστιο) σύνολο των μητρώων, τα μη αντιστρέψιμα μητρώα, όπως και τα μη διαγωνιοποιήσιμα, είναι "λίγα": *Ένα μητρώο με τυχαίες τιμές είναι συνήθως αντιστρέψιμο και διαγωνιοποιήσιμο!*

Πολλαπλές ιδιοτιμές

- Σε κάθε απλή ιδιοτιμή (αλγ. πολλ/τας ίσης με 1) αντιστοιχούν ένα δεξιό και ένα αριστερό ιδιοδιάνυσμα
- Τι γίνεται όταν μια ιδιοτιμή, π.χ. λ_j, έχει αλγ. πολλ/τα, έστω $\mu_j > 1$;
- Το κρίσιμο ερώτημα είναι κατά πόσον υπάρχουν μ_j γραμμικά ανεξάρτητα ιδιοδιανύσματα[4].
- Αν για κάθε διακριτή ιδιοτιμή, η διάσταση του αντίστοιχου μηδενόχωρου είναι ίση με την αλγεβρική πολλαπλότητά της, τότε υπάρχουν συνολικά n γραμμικά ανεξάρτητα ιδιοδιανύσματα και το μητρώο είναι διαγωνιοποιήσιμο.
- Κατά συνέπεια, τα ιδιοδιανύσματα αποτελούν βάση για όλον το διανυσματικό χώρο ($\mathbb{R}^n$ ή $\mathbb{C}^n$).
- Η βάση αυτή (ιδιοβάση) μπορεί να είναι προτιμότερη της τυπικής βάσης $\{e_1, \ldots, e_n\}$ καθώς ως προς κάθε στοιχείο της βάσης αυτής, ο πολλαπλασιασμός με το A και άλλες πράξεις γίνονται πολύ εύκολα.

[4] Στη συνέχεια θα εννοούμε τα δεξιά ιδιοδιανύσματα, αλλά το ίδιο ισχύει για τα αριστερά.

Ιδιοτιμές και ιδιοδιανύσματα δυνάμεων και πολυωνύμων μητρώου

- Για κάθε A, αν $Ax = \lambda x$ τότε $A^k x = \lambda^k x$
- ομοίως $\gamma A^j x + \delta A^k x = \gamma \lambda^j x + \delta \lambda^k x$
- άρα αν οι ιδιοτιμές του A είναι $\{\lambda_1, \dots, \lambda_n\}$ τότε οι ιδιοτιμές του πολυωνύμου $q(A) = \gamma_s A^s + \cdots + \gamma_0 I$, είναι

$$\begin{array}{rcl} q(\lambda_1) & = & \gamma_s \lambda_1^s + \cdots + + \gamma_1 \lambda_1 + \gamma_0 I \\ \vdots & \vdots & \vdots \\ q(\lambda_n) & = & \gamma_s \lambda_n^s + \cdots + \gamma_1 \lambda_n + \gamma_0 I \end{array}$$

δηλ. $\{q(\lambda_1), \dots, q(\lambda_n)\}$.

Συμπεράσματα

- Επομένως αν γνωρίζουμε τις ιδιοτιμές του A γνωρίζουμε και τις ιδιοτιμές οποιουδήποτε πολυωνύμου στο A.
- Τα ιδιοδιανύσματα του $q(A)$ είναι ίδια με τα ιδιοδιανύσματα του A.
- Τα αποτελέσματα επεκτείνονται και σε γενικότερες συναρτήσεις μητρώου, π.χ. Αν $X^{-1}AX = \Lambda$ τότε $X^{-1}\exp(A)X = \exp(\Lambda)$