Γραμμική Άλγεβρα

Ενότητα **6** : Ιδιοτιμές & Ιδιοδιανύσματα

Ευστράτιος Γαλλόπουλος
Τμήμα Μηχανικών Η/Υ & Πληροφορικής

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο
ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ
Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης
ΕΣΠΑ 2007-2013
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Άδειες Χρήσης

Χρηματοδότηση

- Το παρόν εκπαιδευτικό υλικό έχει αναπτυχθεί στα πλαίσια του εκπαιδευτικού έργου του διδάσκοντα.
- Το έργο **«Ανοικτά Ακαδημαϊκά Μαθήματα στο Πανεπιστήμιο Πατρών»** έχει χρηματοδοτήσει μόνο τη αναδιαμόρφωση του εκπαιδευτικού υλικού.
- Το έργο υλοποιείται στο πλαίσιο του Επιχειρησιακού Προγράμματος «Εκπαίδευση και Δια Βίου Μάθηση» και συγχρηματοδοτείται από την Ευρωπαϊκή Ένωση (Ευρωπαϊκό Κοινωνικό Ταμείο) και από εθνικούς πόρους.

Ευρωπαϊκή Ένωση
Ευρωπαϊκό Κοινωνικό Ταμείο

ΕΠΙΧΕΙΡΗΣΙΑΚΟ ΠΡΟΓΡΑΜΜΑ
ΕΚΠΑΙΔΕΥΣΗ ΚΑΙ ΔΙΑ ΒΙΟΥ ΜΑΘΗΣΗ
επένδυση στην κοινωνία της γνώσης
ΥΠΟΥΡΓΕΙΟ ΠΑΙΔΕΙΑΣ & ΘΡΗΣΚΕΥΜΑΤΩΝ, ΠΟΛΙΤΙΣΜΟΥ & ΑΘΛΗΤΙΣΜΟΥ
ΕΙΔΙΚΗ ΥΠΗΡΕΣΙΑ ΔΙΑΧΕΙΡΙΣΗΣ

ΕΣΠΑ 2007-2013
πρόγραμμα για την ανάπτυξη
ΕΥΡΩΠΑΪΚΟ ΚΟΙΝΩΝΙΚΟ ΤΑΜΕΙΟ

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης

Σκοπός Ενότητας

- Εύρεση Ιδιοτιμών & Ιδιοδιανυσμάτων
- Διαγωνοποιήση Μητρώων
- Συμμετρικά, Συμμετρικά και Θετικά Ορισμένα, Όμοια Μητρώα
- Ανάλυση Ιδιαζουσών Τιμών (SVD)

Περιεχόμενα

Υπενθύμιση και πρόγραμμα διάλεξης

Συζητάμε το **Αλγεβρικό Πρόβλημα Ιδιοτιμών (ΑΠΙ)**

1. Συναρτήσεις μητρώων.
2. Διαγωνοποιήσιμα μητρώα.
3. Αναλλοίωτοι υπόχωροι.
4. Μετασχηματισμοί ομοιότητας.
5. Παραγοντοποίηση Schur και αναγωγή μητρώων σε τριγωνική μορφή.
6. Φασματικές ιδιότητες συμμετρικών πραγματικών μητρώων.
7. Ομοιότητα μητρώων.
8. Κανονική μορφή Jordan.

Υπενθύμιση και πρόγραμμα διάλεξης

Συζητάμε το **Αλγεβρικό Πρόβλημα Ιδιοτιμών (ΑΠΙ)**

1. Συναρτήσεις μητρώων.
2. Διαγωνοποιήσιμα μητρώα.
3. Αναλλοίωτοι υπόχωροι.
4. Μετασχηματισμοί ομοιότητας.
5. Παραγοντοποίηση Schur και αναγωγή μητρώων σε τριγωνική μορφή.
6. Φασματικές ιδιότητες συμμετρικών πραγματικών μητρώων.
7. Ομοιότητα μητρώων.
8. Κανονική μορφή Jordan.

Σήμερα θα συζητήσουμε:

1. Μέθοδος δύναμης (απλή μέθοδος υπολογισμού ιδιοζεύγους)
2. Ρίζες πολυωνύμου με γραμμική άλγεβρα: το συνοδευτικό μητρώο.
3. Λύση γραμμικού συστήματος από το φασματικό ανάπτυγμα.
4. Παραγοντοποιήσεις μητρώων.
5. Εισαγωγή στην παραγοντοποίηση ιδιαζουσών τιμών.

Υπό συζήτηση ενότητες

Μία (αφελής) μέθοδος υπολογισμού του ΑΠΙ

Σημείωση: Συνήθως εννοούμε την εύρεση των ιδιοτιμών και αντίστοιχων δεξιών, κανονικοποιημένων, ιδιοδιανυσμάτων.
ΜΟΝΟΝ ΓΙΑ ΠΟΛΥ ΜΙΚΡΑ ΠΡΟΒΛΗΜΑΤΑ!

1. Εύρεση του χαρακτηριστικού πολυωνύμου
2. Υπολογισμός των ριζών του (ιδιοτιμές)
3. Για κάθε διακριτή ιδιοτιμή, επίλυση του $(A - \lambda I)x = 0$ (εύρεση μηδενόχωρου, κεφ. 3). Κανονικοποίηση των διανυσμάτων που (βάση) που επελέγη να εκπροσωπεί τον υπόχωρο.

Μέθοδος δύναμης I

Απλή μέθοδος για ... ΓΙΓΑΝΤΙΑΙΑ ΠΡΟΒΛΗΜΑΤΑ

> Αν η μέγιστη σε απόλυτη τιμή ιδιοτιμή, έστω $\lambda_{\max}$, ενός $A \in \mathbb{R}^{n \times n}$ είναι μοναδική (άρα πραγματική!), με αντίστοιχο ιδιοδιάνυσμα q_1, και έστω $x \in \mathbb{R}^n$ τυχαίο, τ.ώ. $x^\top q_1 \neq 0$, τότε η ακολουθία των διανυσματων $\{Ax, A^2x, \ldots, A^kx, \ldots\}$ τείνει σε διάνυσμα του ιδιόχωρου του $\lambda_{\max}$.

Σε περίπτωση που η μέγιστη σε απόλυτη τιμή είναι μοναδική, τότε τείνει σε διάνυσμα συγγραμμικό του q_1.

Σημείωση: Το q_1 αποκαλείται «μέγιστο ιδιοδιάνυσμα» γιατί αντιστοιχεί στη μέγιστη ιδιοτιμή, όχι επειδή συμβαίνει κάτι ιδιαίτερο με τα στοιχεία του σε σχέση με των υπολοίπων.

Μέθοδος δύναμης I

$$A = \begin{pmatrix} 2 & -3 & 7 & 4 \\ -3 & -5 & -1 & -5 \\ -5 & 4 & 1 & 0 \\ 1 & 5 & -6 & 6 \end{pmatrix}, \text{ και κάποιο διάνυσμα, π.χ. } x = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix}$$

παρακολουθούμε τις τιμές $A^k x$ για $k = 1, 2, \ldots$

$$\begin{array}{cccccc} Ax & A^2x & A^3x & A^4x & A^5x & A^6x \end{array}$$
$$\begin{pmatrix} 10 & 86 & -696 & -2106 & 37370 & 294572 \\ -14 & 10 & -82 & -146 & -23798 & -81002 \\ 0 & -106 & -496 & 2656 & 12602 & -269440 \\ 6 & -24 & 628 & 5638 & 15056 & -66896 \end{pmatrix}$$

$$\begin{array}{cccccc} (Ax) \oslash x & (A^2x) \oslash (Ax) & (A^3x) \oslash (A^2x) & (A^4x) \oslash (A^3x) & (A^5)x \oslash (A^4x) & (A^6x) \oslash (A^5x) \end{array}$$
$$\begin{pmatrix} 22.0 & 21.05 & 21.12 & 21.15 & 21.16 & 21.16 \\ 18.0 & 21.22 & 21.17 & 21.17 & 21.17 & 21.17 \\ 20.0 & 21.1 & 21.17 & 21.16 & 21.17 & 21.17 \\ 27.0 & 21.67 & 21.3 & 21.2 & 21.18 & 21.17 \end{pmatrix}$$

Μέθοδος δύναμης II

$$\begin{array}{cccccc} Ax & A^2x & A^3x & A^4x & A^5x & A^6x \end{array}$$
$$\begin{pmatrix} 22 & 463 & 9777 & 206822 & 4376758 & 92633035 \\ 18 & 382 & 8088 & 171240 & 3624578 & 76718462 \\ 20 & 422 & 8935 & 189095 & 4002384 & 84712934 \\ 27 & 585 & 12461 & 264203 & 5594843 & 118435377 \end{pmatrix}$$

$$\begin{array}{cccccc} (Ax) \oslash x & (A^2x) \oslash (Ax) & (A^3x) \oslash (A^2x) & (A^4x) \oslash (A^3x) & (A^5)x \oslash (A^4x) & (A^6x) \oslash (A^5x) \end{array}$$
$$\begin{pmatrix} 22.0 & 21.05 & 21.12 & 21.15 & 21.16 & 21.16 \\ 18.0 & 21.22 & 21.17 & 21.17 & 21.17 & 21.17 \\ 20.0 & 21.1 & 21.17 & 21.16 & 21.17 & 21.17 \\ 27.0 & 21.67 & 21.3 & 21.2 & 21.18 & 21.17 \end{pmatrix}$$

φαίνεται ότι

$$Ay \approx 21.17y$$

Μέθοδος δύναμης III

κανονικοποιώντας το διάνυσμα και θέτοντας $\mu = 21.17$,

$$\hat{y} = \frac{1}{\|A^5 x\|} A^5 x = \begin{pmatrix} 0.4905 \\ 0.4062 \\ 0.4486 \\ 0.6270 \end{pmatrix} \Rightarrow A\hat{y} - \mu\hat{y} = \begin{pmatrix} -0.0026 \\ -0.0016 \\ -0.0020 \\ -0.0008 \end{pmatrix} \Rightarrow \|A\hat{y} - \mu\hat{y}\|_2 = 0.0037$$

```
>> [V,D]=eig(A)
V =
-0.4905   -0.2586   -0.2925   -0.3602
-0.4062    0.0770    0.6087   -0.5899
-0.4485   -0.0085    0.2586    0.6795
-0.6271    0.9629   -0.6907    0.2462
D =
21.1657         0         0 ...
        0
0      5.9757         0         0
0         0    0.2451         0
0         0         0   -4.3866
>> A*V(:,1)-D(1,1)*V(:,1)
ans =
1.0e-14 *

-0.5329
-0.3553
-0.7105
-0.7105
```

- Η `eig` είναι η συνάρτηση MATLAB για τον υπολογισμό όλων των ιδιοτιμών και κανονικοποιημένων διανυσμάτων.
- Χρησιμοποιεί τον αλγόριθμο QR - περισσότερα σε επόμενο μάθημα.
- Επιστρέφει όλες τις ιδιοτιμές (και τα ιδιοδιανύσματα αν ζητηθούν) και απαιτεί πολλές πράξεις και μνήμη.
- Δεν μπορεί να χρησιμοποιηθεί για πολύ μεγάλα μητρώα.

Μαθηματικός μηχανισμός

Βασική διαδικασία

Εκκίνηση: τυχαίο διάνυσμα $x := x^{(0)}$
for $k = 1, \ldots$
  $x^{(k+1)} \leftarrow Ax^{(k)}$
end

Έστω ότι:

1. το μητρώο είναι διαγωνιοποιήσιμο
2. $(x^{(0)})^\top u_1 \neq 0$.
3. $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$

$$\begin{aligned} x^{(0)} &= \xi_1 u_1 + \cdots + \xi_n u_n, \quad \xi_1 \neq 0 \\ A^k x^{(0)} &= \xi_1 \lambda_1^k u_1 + \cdots + \xi_n \lambda_n^k u_n, \text{ και αφού } |\lambda_1| > |\lambda_j| \\ \frac{1}{\lambda_1^k} A^k x &= \xi_1 u_1 + \cdots + \xi_n \left(\frac{\lambda_n}{\lambda_1}\right)^k u_n \end{aligned}$$

Προσοχή: Το $A^k x$ τείνει προς την κατεύθυνση του u_1.

Βελτιώσεις και παρατηρήσεις

Ζητήματα

- Ενδεχόμενη μεγάλη αύξηση (ή μείωση) των τιμών των $A^k x$ μπορεί να οδηγήσουν σε υπερχείλιση ή υποχείλιση.
- Η χρήση της διαίρεσης (στοιχείο προς στοιχείο) $A^{k+1}x \oslash A^k x$ δεν συνιστάται εδώ, π.χ. αστοχούμε αν κάποιο $A^k x$ περιέχει μηδέν.
- Χρειάζεται καλύτερη κατανόηση της μεθόδου.
- ☺ $\longrightarrow$ *θέμα άλλου μαθήματος (Αριθμητική Ανάλυση και Επιστημονικός Υπολογισμός)*
- *Η μέθοδος δύναμης και παραλλαγές της έχουν χρησιμοποιηθεί για σημαντικές εφαρμογές, π.χ. τον υπολογισμό του PageRank της Google.*

Πηλίκο Rayleigh

Δοθέντος ενός μητρώου A, το πηλίκο Rayleigh που αντιστοιχεί στο διάνυσμα $x \neq 0$ είναι η τιμή

$$r(A; x) = x^\top Ax / x^\top x.$$

- Αν το x είναι ιδιοδιάνυσμα και $Ax = \lambda x$, τότε $r(A; x) = \lambda$.
- Συνήθως, αν $\hat{x} \approx x$, δηλ. το $\hat{x}$ προσεγγίζει το ιδιοδιάνυσμα x, τότε $r(A; \hat{x}) \approx \lambda$.
- Αν η μέθοδος δύναμης με αρχικό διάνυσμα x_0 προσεγγίζει το ιδιοδιάνυσμα x, τότε μέσω του $(x_0^\top A^k x_0)^{1/k}$ προκύπτει εκτίμηση της ιδιοτιμής. Μπορούμε επίσης να χρησιμοποιήσουμε και το πηλίκο Rayleigh $r(A^k, x_0)^{1/k}$.
- Το πηλίκο Rayleigh είναι ασφαλέστερο από τη χρήση της διαίρεσης που έγινε στο παραπάνω παράδειγμα καθώς αποφεύγονται διαιρέσεις με το 0 ή μικρούς αριθμούς.

Συνοδευτικό μητρώο

Από τα πολυώνυμα στα μητρώα

Για κάθε πολυώνυμο p βαθμού n, υπάρχει μητρώο $A \in \mathbb{C}^{n\times n}$ με χ.π. ίδιο με p.

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

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & 0 & \cdots & -\alpha_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -\alpha_{n-1} \end{pmatrix},$$ το συνοδευτικό μητρώο του p.

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

Παράδειγμα

Πρόβλημα Να υπολογιστεί η μεγαλύτερη ρίζα του
$p(\zeta) = \zeta^5 + 10\zeta^4 - 245\zeta^3 + 430\zeta + 244\zeta - 440.$
Λύση με μέθοδο δύναμης Το συνοδευτικό μητρώο είναι

$$A = \begin{pmatrix} 0 & 0 & 0 & 0 & 440 \\ 1 & 0 & 0 & 0 & -244 \\ 0 & 1 & 0 & 0 & -430 \\ 0 & 0 & 1 & 0 & 245 \\ 0 & 0 & 0 & 1 & -10 \end{pmatrix}$$

Ο κώδικας υλοποιεί τη μέθοδο δύναμης με υπολογισμό της μέγιστης ιδιοτιμής με το πηλίκο Rayleigh. Στο συγκεκριμένο παράδειγμα γνωρίζουμε ότι η μέγιστη ρίζα είναι 10. Οι τιμές του πηλίκου και η απόστασή τους από το 10 παρουσιάζονται στο διπλανό πίνακα.

```
% to r(j) peri'eqei to phl'iko ...
      Rayleigh gia k'aje ...
      pros'eggish x_new
x=ones(length(A),1); j=1;r(1) = ...
      (x'*A*x)/(x'*x);
while ((norm(A*x-r(j)*x)>1e-4) & ...
      (j<imax))
x= x/norm(x);  j=j+1;
x_new = A*x; r(j) = (x'*x_new)/(x'*x);
x = x_new;
end
```

1	1.00000	9.0000e+00
2	1.00000	9.0000e+00
3	14.20183	4.2018e+00
4	10.80301	8.0301e−01
5	10.16367	1.6367e−01
6	10.03069	3.0692e−02
7	10.00525	5.2548e−03
8	10.00083	8.3483e−04
9	10.00013	1.2509e−04
10	10.00002	1.7897e−05
11	10.00000	2.4675e−06

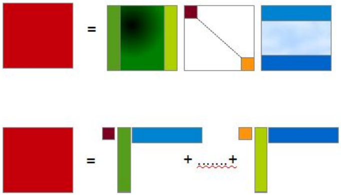
Προσοχή: Προσεγγίστηκε η μέγιστη ρίζα πολυωνύμου σχεδόν αποκλειστικά με πολλαπλασιασμούς του συνοδευτικού μητρώου με διάνυσμα. Περισσότερα στην Αριθμ. Ανάλυση (2ο έτος). ☺

Μητρώο ως ανάπτυγμα ιδιοτιμών και ιδιοδιανυσμάτων (υπενθ.)

Φασματικό ανάπτυγμα

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. 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.

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



Η λύση συστήματος μέσω φασματικού αναπτύγματος

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

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

Η λύση συστήματος μέσω φασματικού αναπτύγματος

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> Συνήθως δεν συνιστάται ως μέθοδος επίλυσης ...

Η λύση συστήματος μέσω φασματικού αναπτύγματος

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> Συνήθως δεν συνιστάται ως μέθοδος επίλυσης ...
αναδεικνύει χρήσιμα χαρακτηριστικά της λύσης!

- Πως γράφεται ως γραμμικός συνδυασμός ιδιοδιανυσμάτων
- Πως μεγενθύνονται οι συνιστώσες για όρους $y_j^* b \neq 0$ με μικρά $|\lambda_j|$.
- Ταξινόμηση συνιστωσών;

Κανονική μορφή Jordan I

Τι κάνουμε όταν το μητρώο δεν είναι διαγωνιοποιήσιμο;

Τότε 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

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

Κανονική μορφή Jordan II

Κάθε J_i αποκαλείται απλό (υπο)μητρώο Jordan και έχει τη μορφή

$$J_i = \cdot \; 0$$

- 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.

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; Χρησιμοποιούμε εναλλακτικές παραγοντοποιήσεις και μορφές.

Εναλλακτικές παραγοντοποιήσεις

Schur υπάρχει πάντα ορθομοναδιαίο Q τέτοιο ώστε $Q^*AQ = R$ είναι άνω τριγωνικό.

SVD υπάρχουν πάντα ορθομοναδιαία U, V τέτοια ώστε το $U^*AV = \Sigma$ να είναι διαγώνιο (με μη αρνητικά στοιχεία!)

ΠΑΡΑΓΟΝΤΟΠΟΙΗΣΕΙΣ ΜΗΤΡΩΟΥ

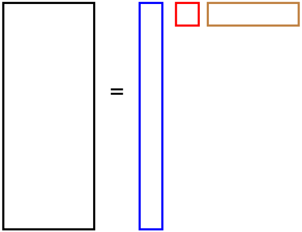
Αναφερόμαστε στη δυνατότητα να εκφράσουμε δοθέν μητρώο $A \in \mathbb{C}^{m\times n}$ ως γινόμενο

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

για κάποια κατάλληλη τιμή k.

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

- 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 $A \approx WKH$.



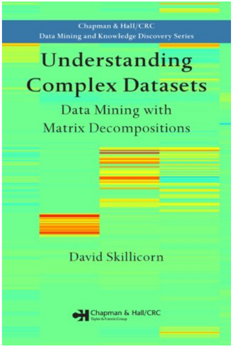
ΠΑΡΑΓΟΝΤΟΠΟΙΗΣΕΙΣ ΜΗΤΡΩΟΥ

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.

- 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 $A \approx WKH$.
- Τέτοιου τύπου προβλήματα εμφανίζονται σε πολλές εφαρμογές: π.χ. Ανάκτηση Πληροφορίας, Επεξεργασία Σημάτων και Εικόνας, Ανάλυση Δεδομένων.



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. $A = BC$

[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

- $U_1^\top A V_2 = 0_{r,n-r}$ γιατί V_2 είναι βάση του $\mathrm{null}(A)$.
- $U_2^\top A V_1 = 0_{m-r,r}$ γιατί U_2 είναι βάση του $\mathrm{null}(A^\top)$.
- Συνεπάγεται ότι $U_2^\top A V_2 = 0_{m-r,n-r}$.
- Το $C \in \mathbb{R}^{r\times r}$ είναι $r \times r$ και έχει τάξη r γιατί $\mathrm{rank}(A) = r = \mathrm{rank}(U^\top A V) = \mathrm{rank}(C)$. Επομένως θα είναι αντιστρέψιμο.
- Αν ονομάσουμε το $U^\top A V = R$ 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 $C = U_1^\top A V_1$ να είναι διαγώνιο με θετικά στοιχεία στη διαγώνιο. Η παραγοντοποίηση που προκύπτει λέγεται παραγοντοποίηση ιδιαζουσών τιμών (SVD.)

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

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

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

Βιβλιογραφία I

G. Strang.
Εισαγωγή στη Γραμμική Άλγεβρα.
Εκδόσεις Πανεπιστημιου Πατρών, 2006.

Χρήση Έργου Τρίτων I

1. https://www.crcpress.com/product/isbn/9781584888321 (βλ. σελ 22)

Σημείωμα Αναφοράς

Copyright Πανεπιστήμιο Πατρών - Ευστράτιος Γαλλόπουλος 2015

``Γραμμική Άλγεβρα'', Έκδοση: 1.0, Πάτρα 2014-2015.
Διαθέσιμο από τη δικτυακή διεύθυνση:

https://eclass.upatras.gr/courses/CEID1097/

Τέλος Ενότητας

Με τη συγχρηματοδότηση της Ελλάδας και της Ευρωπαϊκής Ένωσης