

22/3/2016

3ο κεφάλαιο

Γραμμικά συστήματα

Δεδομένα: $A \in \mathbb{R}^{n,n}$ $n \times n$ πίνακας
 $b \in \mathbb{R}^n$ διάνυσμα με n διαστάσεις

Ζητούμενο: $x \in \mathbb{R}^n$ τ.ω $Ax=b$

Τέτοια γραμμικά συστήματα προκύπτουν πολύ συχνά στις εφαρμογές, συνήθως ως μέρος σθετερότερου προβλήματος.

Θα μας απασχολήσουν τα θέματα:

α) Αριθμητικές μέθοδοι για την προσέγγιση του x .

- κόστος (απαιτούμενες πράξεις και απαιτούμενα μνήμη)
- ευστάθεια

β) κατάλυση γραμμικών συστημάτων

Κατηγορίες αριθμητικών μεθόδων:

• Οι άμεσες (παράλλαξη της μεθόδου απαλοιφής του Gauss). Δίνουν την λύση ακριβώς με πεπερασμένο πλήθος πράξεων, αν οι πράξεις γίνουν ακριβώς.

• Οι επαναληπτικές. Δίνουν μία ακολουθία προσεγγίσεων $x^{(0)}, x^{(1)}, x^{(2)}, \dots$ της λύσης x .

Γενικά για γραμμικά συστήματα

Δεδομένα: Συντελεστές $a_{ij} \in \mathbb{R}, i=1,2,\dots,n, j=1,2,\dots,n$

Δεύτερα μέλη $b_i \in \mathbb{R}, i=1,2,\dots,n$

Ζητούμενα: $x_1, \dots, x_n \in \mathbb{R}$ τ.ω

$$\left. \begin{aligned} a_{11}x_1 + a_{12}x_2 + \dots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \dots + a_{2n}x_n &= b_2 \\ \vdots & \\ a_{n1}x_1 + a_{n2}x_2 + \dots + a_{nn}x_n &= b_n \end{aligned} \right\}$$

Με τον πίνακα A , $A = (a_{ij})_{i,j=1,\dots,n} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix}$

και τα διανύσματα $b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$, $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$ το γραμμικό σύστημα
γράφεται στη μορφή $A \cdot x = b$.

Θα ασπληθούμε με γραμμικά συστήματα που έχουν ακριβώς μία λύση.

Κάθε μία από τις ακόλουθες συνθήκες είναι κατά & αναγκαία για να έχει το $Ax = b$ ακριβώς μία λύση:

i) Ο A είναι αυτοαπέγγραφος, δηλαδή υπάρχει ο A^{-1} .

ii) $\det A \neq 0$

iii) $A \cdot x = 0 \Rightarrow x = 0$

iv) Οι στήλες (ή οι γραμμές) του A είναι γραμμικά ανεξάρτητα διανύσματα του \mathbb{R}^n .

Τρόποι επίλυσης του $Ax = b$, γνωστοί από την γραμμική άλγεβρα.

(α) Κανόνας του Cramer

$A = (a^1 \ a^2 \ \dots \ a^n)$ $a^i = n$ i -οστή στήλη του A .

$A_i = (a^1 \ a^2 \ \dots \ a^{i-1} \ b \ a^{i+1} \ \dots \ a^n)$ $i = 1, \dots, n$

τότε, $x_i = \frac{\det A_i}{\det A}$, $i = 1, 2, \dots, n$

(β) Με τον \bar{A}^{-1} : υπολογίζουμε τον \bar{A}^{-1} και έχουμε: $x = \bar{A}^{-1} \cdot b$

Οι δύο αυτοί τρόποι έχουν μόνο θεωρητική σημασία και όχι πρακτική.

α) Cramer: Αναπτύσσοντας ως προς τη στήλη j έχουμε $\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{i,j}$ με $A_{i,j} \in \mathbb{R}^{(n+1, n-1)}$ που προκύπτει από τον A αν διαγράψουμε την γραμμή i και την στήλη j . Συνεχίζοντας με αυτό τον τρόπο γράφουμε την $\det A$ ως άθροισμα $n!$ όρων με n πορείες ο καθένας. Για τον υπολογισμό του αθροίσματος απαιτούμε $n!(n-1)$ πολλαπλασιασμούς.

Αυτός ο αριθμός αυξάνει πολύ γρήγορα με το n και καθιστά την μέθοδο ασύμφορη:

Στη μέθοδο του Cramer απαιτούμε $(n+1) \cdot n! \cdot (n-1)$ πορίσματα.
↳ πλήθος ορίσμων
($\frac{\det A_i, i=1, \dots, n}{\det A} \rightarrow 1$ γάρ $n+1$)

(Υπάρχουν και άλλοι τρόποι υπολογισμού ορίσμων που απαιτούν λιγότερες πράξεις).

β) $x = \bar{A}^{-1} \cdot b$:

Σπάνια χρειάζεται στω πράξη ο υπολογισμός του αντιστρόφου (\bar{A}^{-1}). Όταν χρειάζεται αυτό γίνεται με τον εξής τρόπο:

Έστω e^1, e^2, \dots, e^n η κανονική βάση του \mathbb{R}^n , δηλαδή

$e_j^i = \delta_{ij} = \begin{cases} 0 & \text{διαφορετικά} \\ 1 & \text{για } i=j \end{cases}$. Το σύμβολο δ_{ij} λέγεται

δέλτα του Kronecker ή σύμβολο του Kronecker.

Έστω $u^i \in \mathbb{R}^n$ τ.ω $A \cdot u^i = e^i, i=1, \dots, n$. Αυτά είναι n γραμμικά υψύματα με $n \times n$ πίνακα.

Ίσχυρισμός :

$$\bar{A}^{-1} = (u^1 \ u^2 \ \dots \ u^n). \text{ Πράγματι, } A(u^1, \dots, u^n) = (Au^1, Au^2, \dots, Au^n) \\ = (e^1, e^2, \dots, e^n) = I_n.$$

Για να υπολογίσουμε τον \bar{A}^{-1} λύνουμε n γραμμικά συστήματα με πίνακα A !

- Επί πλέον κόστος : υπολογισμός του $\bar{A}^{-1} \cdot b$.

Δύο μεγάλες κατηγορίες γραμμικών συστημάτων (πινάκων)

- Πυκνοί (αποθηκεύσιμοι) πίνακες : Έχουν στοιχεία a_{ij} γενικά διάφορα του μηδενός.
- Αραιοί (αποραθικοί πίνακες) : Έχουν πολλά μηδενικά στοιχεία που αυτα εκμεταλλευτούμε αποκομίζουμε υπολογιστικά οφέλη.

Μέγεθος πινάκων : $A \in \mathbb{R}^{n,n}$

$n < 100$: μικρός

$100 \leq n < 1000$: μεσαίοι

$n \geq 1000$: μεγάλοι

Στην πράξη λύνουμε γραμμικά συστήματα με πυκνούς πίνακες μέχρι μετρίου μεγέθους, ενώ με αραιούς πίνακες και μεγάλου μεγέθους.

κατηγορίες αριθμ. μεθόδων

- Άμεσες : χρησιμοποιούνται κυρίως για πυκνούς πίνακες.
- Επαναληπτικές : χρησιμοποιούνται κυρίως για αραιούς πίνακες.

Η μέθοδος αναλοίφης του Gauss

Έστω $u = (u_{ij}) \in \mathbb{R}^{n,n}$ ένας αντιστρέψιμος, άνω τριγωνικός πίνακας, δηλαδή $u_{ij} = 0$ για $i > j$.

Τότε: $\det u = u_{11} \cdot u_{22} \cdot \dots \cdot u_{nn}$. (u αντιστρέψιμος $\Rightarrow u_{ii} \neq 0$ $i = 1, \dots, n$)

Έστω $y \in \mathbb{R}^n$. Το γραμμικό σύστημα $u \cdot x = y$
δηλαδή
$$\left. \begin{aligned} u_{11}x_1 + u_{12}x_2 + \dots + u_{1n}x_n &= y_1 \\ u_{22}x_2 + \dots + u_{2n}x_n &= y_2 \\ &\vdots \\ u_{nn}x_n &= y_n \end{aligned} \right\}$$

όταν είναι άνω ή κάτω τριγωνικός μόνο ο πίνακας τότε $\det A = a_{11} \cdot \dots \cdot a_{nn}$

Λύνεται εύκολα με μία διαδικασία που λέγεται οπισθοδρόμηση. Λύνουμε την τελευταία εξίσωση ως προς το x_n , αντικαθιστούμε στη προηγούμενη εξίσωση και βρίσκουμε το x_{n-1} , ..., και τέλος αντικαθιστούμε τα x_n μέχρι x_2 στην πρώτη εξίσωση και υπολογίζουμε το x_1 .

Αλγόριθμος της οπισθοδρόμησης

$$x_n = \frac{y_n}{u_{nn}}$$

Για $k = n-1, n-2, \dots, 1$:

$$x_k = \frac{1}{u_{kk}} \left[y_k - \sum_{j=k+1}^n u_{kj} x_j \right]$$

τέλος - για

$$\sum_{j=k+1}^n u_{kj} x_j = y_k \Rightarrow$$

$$(u_{kk}) x_k = y_k - \sum_{j=k+1}^n u_{kj} x_j$$

$$\Rightarrow \text{διαιρούμε με } u_{kk}$$

Απαιτούμενες πράξεις

Διαιρέσεις: n

$$\rightarrow \frac{n(n-1)}{2}$$

Πολλαπλασιασμοί (και προσθέσεις)

$$1 + 2 + \dots + (n-1) = \textcircled{3} \frac{n(n-1)}{2} = \frac{n^2}{2} - \frac{n}{2}$$

Θέσεις μύθας:

$\frac{n^2}{2} + O(n)$ για τα $u_{ij} \neq 0$ και τα y_i . Τα x_i ορισθηκαν εδω και οι θέσεις των y_i .

Ιδέα στη μέθοδο αναλοιφής του Gauss.

Με κατάλληλους μετασχηματισμούς γραμμών (εναλλαγή δύο εξισώσεων ή πρόσθεση κατά μέλη ενός πολλαπλασίου μιας εξίσωσης σε κάποια άλλη) με τέτοιες διαδικασίες μετατρέπουμε ένα γραμμικό σύστημα $A \cdot x = b$ σε ένα ισοδύναμο $U \cdot x = \gamma$ με έναν τριγωνικό πίνακα u . Η διαδικασία αυτή λέγεται τριγωνοποίηση.

• Επίλυση με οπισθοδρόμηση.

Δύο στάδια: Τριγωνοποίηση
Οπισθοδρόμηση.

Γενική περιγραφή:

$A \cdot x = b$. Θέτουμε $A^{(1)} = (a_{ij}^{(1)}) = A$ και $b^{(1)} = \begin{pmatrix} b_1^{(1)} \\ \vdots \\ b_n^{(1)} \end{pmatrix} = b$

και το $A \cdot x = b$ γράφεται ως $A^{(1)} \cdot x = b^{(1)}$.

Τριγωνοποίηση

1^ο βήμα: Υποθέτω ότι $a_{11}^{(1)} \neq 0$ (μπορεί να επιτευχθεί με εναλλαγή γραμμών)

Πολλαπλασιαστές

$m_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}}$, $i = 2, \dots, n$. Πολλαπλασιάζουμε την 1^η εξίσωση

με m_{i1} , την αφαιρούμε από την i -οστή εξίσωση και αντικαθιστούμε την i -οστή με το αποτέλεσμα. $i = 2, \dots, n$.

Έτσι, ύστερα από το 1ο βήμα περνάμε το γραμμικό σύστημα $A \cdot x = b$

$$\text{με } A^{(2)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{n2}^{(2)} & \dots & a_{nn}^{(2)} \end{pmatrix} \text{ και } b^{(2)} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(2)} \end{pmatrix} \text{ όπου}$$

$$a_{ij}^{(2)} = a_{ij}^{(1)} - m_{i1} a_{j1}^{(1)} \quad i, j = 2, \dots, n \quad \text{και} \quad b_i^{(2)} = b_i^{(1)} - m_{i1} b_1^{(1)} \quad i = 2, \dots, n$$

~~...~~

24/3/2016

$A \cdot x = b$, $A \in \mathbb{R}^{n,n}$ αντιστρέψιμος

Μέθοδος απαλοιφής του Gauss

• Τριγωνοποίηση

Βήμα V: $1 \leq r \leq n-1$. Ζευγαίρε από το σύστημα $A^{(r)} \cdot x = b^{(r)}$

$$A^{(r)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{r-1,2}^{(r-1)} & \dots & a_{r-1,n}^{(r-1)} \\ \boxed{\begin{matrix} a_{rr}^{(r)} & \dots & a_{rn}^{(r)} \\ \vdots & & \vdots \\ a_{nr}^{(r)} & & a_{nn}^{(r)} \end{matrix}} \\ 0 & & & \end{pmatrix}, \quad b^{(r)} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_{r-1}^{(r-1)} \\ b_r^{(r)} \\ b_r^{(r)} \\ \vdots \\ b_n^{(r)} \end{pmatrix}$$

θέτουμε $\tilde{A}^{(r)} = (a_{ij}^{(r)}) \quad i, j = r, \dots, n$

• $A^{(r)}$ είναι αντιστρέψιμος

• $\det A^{(r)} = a_{11}^{(1)} \cdot a_{22}^{(2)} \cdot \dots \cdot a_{r-1,r-1}^{(r-1)} \cdot \det \tilde{A}^{(r)} \neq 0$

$\Rightarrow \det \tilde{A}^{(r)} \neq 0 \Rightarrow \tilde{A}^{(r)}$ αντιστρέψιμος.

Υπόθεση

$a_{rr}^{(r)} \neq 0$ (μπορεί να επιτευχθεί με εναλλαγές γραμμών r με κάποια από τις γραμμές $r+1, \dots, n$)

Πολλαπλασιαστές

$$m_{ir} = \frac{a_{ir}^{(r)}}{a_{rr}^{(r)}}, \quad i = r+1, \dots, n$$

Πολλαπλασιάζουμε τη γραμμή r επί m_{ir} , αφαιρούμε το αποτέλεσμα από τη γραμμή i και αυτακαθιστούμε τη γραμμή i με αυτό που βρήκαμε, για $i = r+1, \dots, n$.

Καταλύουμε στο σύστημα $A \cdot x = b$ με τις πρώτες r γραμμές του A και αυτιστοιχώς του b ίδιες με τις αυτιστοιχες του $A^{(r)}$ και $b^{(r)}$ αυτιστοιχά, για

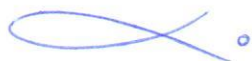
$$a_{ij}^{(r+1)} = a_{ij}^{(r)} - m_{ir} a_{rj}^{(r)}, \quad i, j = r+1, \dots, n$$

$$b_i^{(r+1)} = b_i^{(r)} - m_{ir} \cdot b_r^{(r)}, \quad i = r+1, \dots, n$$

και $a_{ij}^{(r+1)} = 0, \quad i = r+1, \dots, n, j = 1, \dots, r$.

Υπερα από $n-1$ βήματα προκύπτει το σύστημα $\textcircled{*} A \cdot x = b$ με $A^{(n)}$ άνω τριγωνικός πίνακας. Εδώ τελειώνει η τριγωνοποίηση.

Το $\textcircled{*}$ λύεται με οπισθοδρόμηση.



Απαιτούμενες πράξεις και μύμα!

Το πλήθος των προσθαφαιρέσεων είναι περίπου ίσο με το πλήθος των πολλαπλασμών και διαυρέσεων. Έχει επικρατήσει να μετράμε ως αριθμητικές πράξεις μόνο πολλαπλούς και διαυρέσεις.

• Τριγωνοποίηση

1ο βήμα: για τον υπολογισμό των πολλαπλών m_{ij} , $i=2, \dots, n$ απαιτούνται $n-1$ πράξεις. Για ~~την ολοκλήρωση~~ τον υπολογισμό των $a_{ij}^{(2)}$, $i, j=2, \dots, n$ απαιτούνται $(n-1)^2$ πράξεις.

Βήμα r: πολλαπλές: ~~n-r~~ $n-r$. Για τα στοιχεία $a_{ij}^{(r+1)}$, $i, j=r+1, \dots, n$ απαιτούνται $(n-r)^2$ πράξεις.

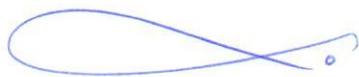
Για τον A απαιτούνται συνολικά στα $n-1$ βήματα:

$$\sum_{i=1}^{n-1} [(n-i)^2 + (n-i)] = \frac{n^3 - n}{3} \text{ πράξεις}$$

$$\left(\sum_{i=1}^k i = \frac{k(k+1)}{2}, \quad \sum_{i=1}^k i^2 = \frac{k(k+1)(2k+1)}{2} \right)$$

Για το b, απαιτούνται $n-r$ πράξεις στο βήμα r, οπότε συνολικά:

$$\sum_{i=1}^{n-1} (n-i) = \frac{n(n-1)}{2} = \frac{n^2}{2} - \frac{n}{2}$$



Για την οπισθοδρόμηση απαιτούνται: $\frac{n^2 + n}{2}$ πράξεις

Παράδειγμα

$n=20$. Υπόθεση, ο υπολογιστής εκτελεί 10^6 πράξεις/second.

Λύση
Για τη μέθοδο απαλοιφής Gauss απαιτούνται $\frac{16}{3} \cdot 10^{-3}$ sec ■

Καώνας του Cramer: Για τους $21 \cdot 20! \cdot 19$ πολλαπλούς μόνον απαιτούμενοι $\approx 3 \cdot 10^5$ αιώσεις. ■ τέλος

Απαιτούμενα μύρια

Για τον A n^2 θέσεις μύρια

Για το b n θέσεις μύρια

Δεν απαιτείται επί πλέον μύρια. Οι πολλαπλές m_{ir} , $i=r+1, \dots, n$ αποθηκεύονται στις θέσεις των στοιχείων ~~a_{ir}~~ a_{ir} , $i=r+1, \dots, n$. Δηλαδή συνολικά οι πολλαπλές αποθηκεύονται στις θέσεις (i,j) με $i>j$. Τα νέα στοιχεία σε κάθε βήμα υπολογίζονται από τους τύπους

$$a_{ij} \leftarrow a_{ij} - m_{ir} \cdot a_{rj}, \quad i,j = r+1, \dots, n$$

Οι στοιχειώδεις του χ αποθηκεύονται στις αντίστοιχες θέσεις των στοιχειωσών του b.

Παρατήρηση υπολογισμός ορίσουςας.

$\det A = (-1)^m \cdot \det A^{(n)}$ με m το πλήθος εναλλαγών γραμμών κατά τη τριγωνοποίηση.

$$\rightarrow = (-1)^m a_{11}^{(1)} \cdot a_{22}^{(2)} \cdot \dots \cdot a_{nn}^{(n)}$$

Απαιτούμενες πράξεις : $\frac{n^3}{3} + o(n)$ ■

Για την επίλυση του $A \cdot x = b$ με τον καώνα του Cramer :

$$\underbrace{(n+1)}_{\text{ορίσεις}} \left(\frac{n^3}{3} + o(n) \right) \approx \frac{n^4}{3}$$

∞.

Οδήγηση

Τα διαγώνια στοιχεία $a_{ii}^{(i)}$ του $A^{(n)}$ λέγονται οδηγοί. Επειδή οι οδηγοί εμφανίζονται ως παρονομαστές είναι αναμενόμενο να εμφανιστούν προβλήματα ευστάθειας στη περίπτωση που έχω μικρή απόλυτη τιμή.

Παράδειγμα

$$\left. \begin{matrix} 10^{-4} \cdot x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{matrix} \right\} \left. \begin{matrix} \text{Ακριβής λύση: } x_1 = 1.0001... \\ x_2 = 0.9998... \end{matrix} \right\}$$

Παίρνω εύσω υπολογιστή με $b=10, t=3, L=-20, U=20$ στρογγύλευση.

Τότε η προσεγγιστική λύση που πάρω είναι $x_1=0$ } πολύ κακή
 $x_2=1$ } προσέγγιση.

τι συνέβα; εφαρμόζοντας

θα πάρω $a_{22}^{(2)} = fl(a_{22}^{(1)} - m_{21} \cdot a_{12}^{(1)}) = fl((1) - 10) \cdot 10^4 \gg 1$

$$m_{21} = \frac{a_{21}^{(1)}}{a_{11}^{(1)}} = \frac{1}{10^{-4}} = 10^4$$

$$= fl(-9.999) = -10^4$$

↑
στρογγύλευση

Ακόμα, $b_2 = b_2 - m_{21} \cdot b_1 = fl(b_2 - m_{21} \cdot b_1) = fl(2 - 10^4) = -10^4$

Χάθηκε εκτελώς η πληροφορία από την 2η γραμμή επίλυση. ■

Μερική οδήγηση (ή οδήγηση κατά γραμμές)

Στο r-οστό βήμα την τριγωνοποίησης εφαρμόζουμε τα στοιχεία που είναι στη στήλη $a_{kr}^{(r)}$ $k=r, r+1, \dots, n$ της r-οστής στήλης του $A^{(r)}$, βρίσκουμε ένα με τη μέγιστη απόλυτη τιμή και με εναλλαχώς γραμμές το φέρουμε στη θέση του οδηγού.

Επιπλέον κόστος:

θέλω $\frac{n^2}{2} - \frac{n}{2}$

Στο παράδειγμα μας: $x_1 + x_2 = 9$ } στον υπολογισμό τα παίρνουμε;
 $10^{-4}x_1 + x_2 = 1$ }

$x_1 = 1$ } πολύ καλή προσέγγιση!
 $x_2 = 1$ } ■

Παρατήρηση

Ο όρος "πολύ μικρός οδηγός" είναι ασαφής. Το αρχικό μας σύστημα (στο παράδειγμα) μπορούμε να το γράψουμε στη μορφή:

$$\left. \begin{array}{l} x_1 + 10^4 x_2 = 10^4 \\ x_1 + x_2 = 9 \end{array} \right\} \text{ στον υπολογισμό: } \left. \begin{array}{l} x_1 = 0 \\ x_2 = 1 \end{array} \right\}$$

Ενώ αν κάνουμε πρώτα εναλλαγές γραμμών παίρνουμε τω καλή λύση $\left. \begin{array}{l} x_1 = 1 \\ x_2 = 1 \end{array} \right\}$. Ο πραγματικός λόγος αποτυχίας είναι η περίπτωση είναι ότι το $|a_{11}|$ είναι πολύ μικρό σε σύγκριση με το $|a_{12}|$,

$$\left| \frac{a_{11}}{a_{12}} \right| = 10^{-4} \text{ (με ότι και να πολλαπλασιάσω αυτό)}$$

$$\text{ενώ } \left| \frac{a_{21}}{a_{22}} \right| = 1. \quad \blacksquare$$

ολική οδύνηση (ή οδύνηση κατά γραμμές και στήλες)

Στο r -οστό βήμα της τριγωνοποίησης με εναλλαγές γραμμών και σελών, φέρνουμε στη θέση του οδηγού $a_{rr}^{(r)}$ ένα από τα στοιχεία $a_{ij}^{(r)}$, $i, j = r, \dots, n$. (προσοχή! Η εναλλαγή των σελών r και s απαιτεί και τω εναλλαγή των στοιχείων x_r και x_s)

Επί πλέον κόστος

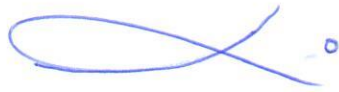
$$\frac{n^3}{3} + O(n^2) \text{ Διπλασιάζει το κόστος!}$$

• Η μέθοδος απαλοιφής του Gauss χωρίς οδήγηση θεωρείται απλά αλγόριθμος. Είναι εύκολη για κάποιες κατηγορίες πινάκων, π.χ για θετικά ορισμένους πίνακες A ,

$\forall x \in \mathbb{R}^n, x \neq 0, x^T A x > 0$, και μόνο σε τέτοιες περιπτώσεις εφαρμόζεται.

• Η μέθοδος ολικής οδήγησης θεωρείται εύκολη αλγόριθμος. Όμως, χρησιμοποιείται σπάνια γιατί διπλασιάζει το κόστος.

• Η πιθανότητα να είναι απλά η μερική οδήγηση είναι πολύ μικρή. Χρησιμοποιείται πολύ συχνά στη πράξη, γιατί αυξάνει λίγο το κόστος.



29/3/2016

Ο αλγόριθμος της απαλοιφής στην πράξη.

Η μέθοδος της απαλοιφής του Gauss για την επίλυση ενός γραμμικού συστήματος $A \cdot x = b$ γράφεται στην πράξη σε 2 στάδια:

- Το πρώτο εκτελεί τους υπολογισμούς της τριγωνοποίησης που αφορούν μόνο τον A .
- Το δεύτερο στάδιο χρησιμοποιεί τα αποτελέσματα του πρώτου και το b για να υπολογίσει το x .

Είναι πολύ γνωστό ζεύγος υποπρογραμμάτων που εκτελεί αυτά τα στάδια είναι τα DECOMP και SOLVE.

• DECOMP: Υλοποιεί το πρώτο στάδιο. Έχει ως είσοδο τα στοιχεία του πίνακα A , που αποθηκεύονται σε n^2 θέσεις μνήμης. Εκτελεί απαλοιφή Gauss με μερική οδήγηση. Το τελικό προϊόν του A αποθηκεύεται στο πάνω "τρίγωνο" του αρχικού πίνακα, δηλαδή στις θέσεις (i, j) με $i \leq j$. Στο κάτω τρίγωνο, δηλαδή στις θέσεις (i, j) με $i > j$, αποθηκεύονται οι πολλαπλασιαστές m_{ij} .

Καταγράφονται επίσης οι πληροφορίες εναλλαγής γραμμών. Σε ένα διάνυσμα του \mathbb{R}^n αποθηκεύουμε στην θέση k τον δείκτη της γραμμής που στο βήμα k της τριγωνοποίησης είναι στην θέση του "οδηγού" δηλαδή εναλλάσσεται με τη γραμμή k .

κόστος : $\frac{n^3}{3} + O(n^2)$ πολλαπλασιασμοί και διααιρέσεις

$n^2 + O(n)$ θέσεις μύμης. ■

SOLVE : Χρησιμοποιεί ως είσοδο τα αποτελέσματα της DECOMP και το διάνυσμα b . Με τους πολλαπλασιασμούς και τις πληροφορίες για τις εναλλαγές γραμμών εκτελεί πρώτα τις πράξεις της τριγωνοποίησης που αφορούν το b . Ύστερα, με οπισθοδρόμηση υπολογίζει τη λύση x και τω αποδιδάει στη θέση του b .

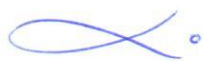
κόστος : $n^2 + O(n)$ πολλαπλασμοί και διααιρέσεις
 n θέσεις μύμης ■

Εφαρμογή

Έστω ότι θέλουμε να λύσουμε m γραμμικό σύστημα με πίνακα A και δεξιά μέλη b, b', b'', \dots

Καλούμε μία φορά την DECOMP και m φορές την SOLVE, μία για κάθε δεξιό μέλος b, b', \dots

Συνολικό κόστος : $\frac{n^3}{3} + O(n^2) + m(n^2 + O(n))$



Παράδειγμα

Υπολογισμός του A^{-1} .

Λύουμε τα συστήματα $A u^i = e^i$ $i=1, \dots, n$ με e^1, \dots, e^n την κανονική βάση του \mathbb{R}^n . Τότε $A^{-1} = (u^1, u^2, \dots, u^n)$

Συνολικό κόστος : $\frac{n^3}{3} + O(n^2) + n \left(\frac{2}{3} n^2 + O(n) \right) = n^3 + O(n^2)$ ■

↑
οφείλεται στα
πολλά μινδικά
του e^i .

Η Ανάλυση LU

Η φάση της τριγωνποίησης ενός αωτιστρέγγου πίνακα A κατά την απαλοιφή Gauss, πιθανώς με εναλλαγές γραμμών, μπορεί να εκφραστεί σε γλώσσα πίνακων ως ανάλυση του A σε γινόμενο, $A = P^{-1}LU$ με $P, L, U \in \mathbb{R}^{n,n}$ με τις εξής ιδιότητες:

- Ο πίνακας P είναι ένας πίνακας μετάθεσης. Αυτός καταγράφει τις εναλλαγές γραμμών που έγιναν κατά την απαλοιφή. (Ένας πίνακας μετάθεσης προκύπτει από τον μοναδιαίο πίνακα I_n με κατάλληλες εναλλαγές γραμμών. Αποθηκεύεται σε ένα διάνυσμα με n στοιχεία)
- Ο L είναι κάτω τριγωντικός πίνακας και στη διαγώνιο έχει μονάδες. Κάτω από τη διαγώνιο περιέχει τους πολλαπλασιαστές της απαλοιφής. (τα διαγώνια στοιχεία δεν αποθηκεύονται).
- Ο U είναι άνω τριγωντικός πίνακας, που είναι το τελικό προϊόν $A^{(n)}$ της απαλοιφής. ■

Αυτός είναι απλώς ένας άλλος τρόπος για να δει κανείς τη φάση της τριγωνποίησης του A . Για την ανάλυση απαιτούνται ακριβώς οι ίδιες πράξεις.

- Απόδειξη της ανάλυσης
- Πώς τη χρησιμοποιούμε?

$$A \cdot x = b \quad \Leftrightarrow \quad \boxed{PA} \cdot x = \cancel{A} \cdot b \quad \Leftrightarrow \quad \boxed{L(Ux)} = Pb$$

↑
Ροωτιστρέγγος

α) Αναλύουμε την PA σε γινόμενο LU , $A = P^{-1}LU$

β) Υπολογισμός της λύσης x :

β1) κατασκευάζουμε το $P \cdot b$ (με εναλλαγή στοιχείων του b) και επιλύουμε το σύστημα $L \cdot y = P \cdot b$ (αυτό επιτυγχάνεται με τρόπο αντίστροφο της οπισθοδρόμησης. Από την πρώτη εξίσωση υπολογίζουμε το y_1 , το οωτικαθιστούμε στα δεξιά και υπολογίζουμε το y_2, \dots)

b2) Επιλύουμε με οπισθοδρόμηση το σύστημα $Ux=y$
 ← γρασσο από το b1. ■

Απόδειξη

1η Περίπτωση

Υποθέτουμε ότι δεν γίνεται εναλλαγές γραμμών (δηλαδή $P=I_n$). Ορίζουμε τον πίνακα

$M_1 = \begin{pmatrix} 1 & & & & 0 \\ -m_{21} & 1 & & & \\ -m_{31} & & 1 & & \\ & & & \ddots & \\ -m_{n1} & & & & 1 \end{pmatrix}$ με $m_{i1}, i=2, \dots, n$ τους πολλαπλασιαστές του 1ου βήματος της τριγωνποίησης του A. Παρατηρούμε ότι

(2)
 $A = M_1 \cdot A$

δηλαδή το πρώτο βήμα της τριγωνποίησης μπορεί να ερμηνευτεί ως πολλαπλασιασμός του A από αριστερά με τον πίνακα M_1 . Συνεχίζοντας αόμοιοιχα βλέπουμε ότι

$A^{(n)} = M_{n-1} \cdot \dots \cdot M_2 \cdot M_1 \cdot A$, με $M_r \in \mathbb{R}^{n,n}$ $r=1, 2, \dots, n-1$ όπου δίδονται ως

$(M_r)_{ij} = \begin{cases} 1 & \text{για } i=j \\ -m_{ir} & \text{για } i=r+1, \dots, n \\ 0 & \text{διαφορετικά.} \end{cases}$

Δηλαδή, ο M_r είναι κάτω τριγωνικός πίνακας με μονάδες στη διαγώνιο, $-m_{ir}, i=r+1, \dots, n$ στην υποδιάγνη r , και παντού αλλού μηδέν.

• Οι M_r είναι αντιστρέψιμοι, $\det M_r = 1$.

Στους κάτω τριγωνικούς πίνακες το $\det A = \prod$ στοιχεία διαγώνιου

Μάλιστα ισχύει $(M_r^{-1})_{ij} = \begin{cases} 1 & \text{για } i=j \\ m_{ir} & \text{για } i=r+1, \dots, n \\ 0 & \text{διαφορετικά} \end{cases}$

(Άσκηση 3.6)

Από την * προκύπτει

$$A = M_1^{-1} \dots M_{n-2}^{-1} M_{n-1}^{-1} \begin{pmatrix} \\ \\ \\ A^{(n)} \end{pmatrix}$$

Τώρα $U = A^{(n)}$ άνω τριγωνικός και $L = M_1^{-1} M_2^{-1} \dots M_{n-1}^{-1} = \begin{pmatrix} 1 & & & & \circ \\ m_{21} & 1 & & & \\ m_{31} & m_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ m_{n1} & m_{n2} & \dots & \dots & 1 \end{pmatrix}$
 κάτω τριγωνικός πίνακας με μονάδες στην διαγώνιο
 και κάτω από τη διαγώνιο έχει τους πολλαπλασιαστές

Άρα, αν έδειξα την επιθυμητή ανάλυση του A να κάνει $A = LU$

■ αποδεικνύει
 τις
 περιπτώσεις

2η περίπτωση (γενική περίπτωση)

Κατά την απαλοιφή Gauss επιτρέπονται και εναλλαγές γραμμών. (είτε για να βρούμε μη μηδενικούς οδηγούς είτε για περικύ αδήγηση)

Η γραμμή που οποία στο βήμα i ^{της τριγωνποίησης} θέτουμε στη θέση του οδηγού (δηλαδή αυτή που εναλλάσσεται από τη γραμμή i) ούτε αλλάζει θέση ούτε αλλοιώνονται πλέον τα στοιχεία της. Επομένως, υπάρχει μία μετάθεση των γραμμών του A που αν την κάνουμε πρώτο από το πρώτο βήμα της τριγωνποίησης θα οδηγούσε σε ένα πίνακα A' , η τριγωνποίηση του οποίου γίνεται απρόσκοπτα χωρίς εναλλαγές γραμμών. Επομένως σύμφωνα με την 1η περίπτωση
 ** $A' = LU$ με L, U πίνακες με τις επιθυμητές ιδιότητες.

Μένει να αποδείξουμε ότι $A' = P \cdot A$ με κατάλληλο πίνακα μετάθεσης P.

Έστω τώρα $\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$ μία μετάθεση των αριθμών $1, 2, \dots, n$. Ο αντίστοιχος πίνακας μετάθεσης P

προκύπτει από τον I_n , αν μεταθέσουμε τις γραμμές του κατά τη μετάθεση P, δηλαδή η i_k γραμμή του P είναι η k γραμμή του I_n $\underline{P\pi}$.

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ 3 & 1 & \dots & 2 \end{pmatrix} \rightsquigarrow P = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

μετάθεση αντίστοιχος
πίνακας
μετάθεσης

και η 3η γραμμή του P είναι η πρώτη γραμμή του I₃.

η 1η -"- -"- -"- 2η -"- -"-

η 2η -"- -"- -"- 3η -"- -"-

Το γινόμενο P.A προκύπτει από τον A με τις ίδιες ακριβώς εναλλαγές γραμμών που από τον I_n οδηγούν στον P. πχ

$$P \cdot A = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \end{pmatrix}$$

~~Συμπέρασμα~~
Συμπέρασμα, στον πίνακα A' οδηγούμε τις ποσότητες του A από αριστερά με έναν πίνακα μεταθέσης P. ■ 2η περίπτωση

Άρα, $A' = P \cdot A$, οπότε $PA = LU$ ή $A = P^{-1}LU$ ■ απόδειξη.

31/3/2016



Η αλάση LU

$A \in \mathbb{R}^{n,n}$ αντιστρέψιμος τότε $A = P^{-1}LU$ (ή $PA = LU$)

Παράδειγμα

1. $A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$ Απολοιφή Gauss $m_{21} = \frac{1}{2}$ $L = \begin{pmatrix} 1 & 0 \\ \frac{1}{2} & 1 \end{pmatrix}$, $A^{(2)} = \begin{pmatrix} 2 & 3 \\ 0 & 5/2 \end{pmatrix} = U$

Επαλήθευση: $P = I_2$, $L \cdot U = \dots = A$

Δεν χρειάζομαι εναλλαγές γραμμών γιατί $P = I_2$. ■ 1ο παράδειγμα

2. $A = \begin{pmatrix} 2 & 4 & 5 \\ 1 & 2 & 3 \\ 1 & 5 & 9 \end{pmatrix}$ Απολοιφή Gauss: $m_{21} = 1/2$, $m_{31} = 1$ $A^{(2)} = \begin{pmatrix} 2 & 4 & 5 \\ 0 & 1 & 1/2 \\ 0 & 1 & 4 \end{pmatrix}$

Για να προκύψει μη μηδενικός οδηγός εναλλάσσουμε την 2η με την 3η γραφή του αρχικού πίνακα. Δηλαδή

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

Όταν σταθεροποιούμε τον P τα L, U είναι μοναδικά.

τώρα, $P \cdot A = \begin{pmatrix} 2 & 4 & 5 \\ 2 & 5 & 9 \\ 1 & 2 & 3 \end{pmatrix} = B$

Απαλοιφή Gauss στον B:

Πολλιστές $m_{21} = 1, m_{31} = \frac{1}{2}$

$$L = \begin{pmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ \frac{1}{2} & 0 & 1 \end{pmatrix}, \quad B^{(2)} = \begin{pmatrix} 2 & 4 & 5 \\ 0 & 1 & 4 \\ 0 & 0 & 1/2 \end{pmatrix} = U$$

$m_{32} = 0$

Επαλήθευση: $L \cdot U = \dots = B = P \cdot A$ ■ 2ο παράδειγμα



κατάσταση γραμμικών συστημάτων

Παράδειγμα

$$\underbrace{\begin{pmatrix} 0.913 & 0.659 \\ 0.780 & 0.563 \end{pmatrix}}_A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.254 \\ 0.217 \end{pmatrix} \rightarrow x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Απαλοιφή Gauss: (χωρίς οδηγία = μερικά οδηγία = ολική οδηγία)

$b=10, t=3$, αποκοπή

Βρίσκω $\tilde{x}_1 = -0.443$
 $\tilde{x}_2 = 1$ } είναι πολύ κακή προσέγγιση!

Τι συμβαίνει?

αλλάξαμε το τελευταίο γινόμενο, 10^{-3} μεταβολή.

$$\begin{pmatrix} 0.913 & 0.659 \\ 0.780 & 0.563 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} 0.253 \\ 0.218 \end{pmatrix} \rightarrow Y = \begin{pmatrix} 1223 \\ -1694 \end{pmatrix} \text{ από } \begin{pmatrix} 1 \\ -1 \end{pmatrix} \text{ που έπαιξε σίω αρχικά.}$$

Άρα, το σύστημα έχει κακή κατάσταση! ■ παράδειγμα 2ο

$$\det A = -10^{-6} \text{ (πολύ μικρή)}$$

Ερώτημα: Μήπως είναι αυτός ο λόγος που έχει κακή κατάσταση το σύστημα?

Απάντηση

Αρρωτικά! ■

πχ

Πολλίμε το αρχικό μας σύστημα με έναν αριθμό, πχ τον 10^4 , οπότε παίρουμε

$$\underbrace{\begin{pmatrix} 0.913 \cdot 10^4 & 0.659 \cdot 10^4 \\ 0.780 \cdot 10^4 & 0.563 \cdot 10^4 \end{pmatrix}}_{\tilde{A}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.254 \cdot 10^4 \\ 0.217 \cdot 10^4 \end{pmatrix} \rightarrow x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$$

Η λύση για το $y = \begin{pmatrix} 1223 \\ -1694 \end{pmatrix}$

$\tilde{A} = 10^4 \cdot A \Rightarrow \det \tilde{A} = (10^4)^2 \det A$ επειδή 2×2 ο πίνακας

$$= 10^8 \cdot (-10^{-6}) = -10^2 = -100 \quad \blacksquare \text{ άρα πορίσμου δευ έχει σχέση.}$$

Πρέπει να βρούμε έναν τρόπο να μετράμε διαταραχές διαυσιμάτων και πινάκων, δηλαδή η απόσταση μεταξύ διαυσιμάτων και μεταξύ πινάκων. Ένας βολικός τρόπος είναι με τις λεγόμενες νόρμες.

Νόρμες Διαυσιμάτων και πινάκων

Η έννοια της νόρμας αποτελεί γενίκευση της έννοιας την απόλυτης τιμής.

• Νόρμες Διαυσιμάτων

Ορισμός (Νόρμα) Έστω X ένας γραμμικός χώρος στο \mathbb{R} είτε στο \mathbb{C} , και θέτουμε $\mathbb{K} = \mathbb{R}$ ή $\mathbb{K} = \mathbb{C}$, αντίστοιχα.

Απεικόνιση

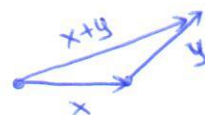
$\|\cdot\| : X \rightarrow \mathbb{R}$ που απεικονίζει ένα στοιχείο x , $x \mapsto \|x\|$,

λέγεται νόρμα (νορμ, σιάθμ). αν ιαχούουν :

(N1) $x \in X \quad \|x\| = 0 \Leftrightarrow x = 0$

(N2) $\forall \lambda \in \mathbb{K}, \forall x \in X \quad \|\lambda x\| = |\lambda| \cdot \|x\|$

(N3) $\forall x, y \in X \quad \|x+y\| \leq \|x\| + \|y\|$ (τριγωνική ανισότητα)



Παρατηρήσεις

$$1) \forall x \in \mathbb{R} \quad \|x\| \geq 0$$

Απόδειξη

$$\begin{aligned}
 0 &= \|x - x\| = \|x + (-x)\| \stackrel{(N3)}{\leq} \|x\| + \|-x\| \stackrel{(N2)}{=} \|x\| + |-1| \cdot \|x\| \\
 &\stackrel{(N1)}{=} \|x\| + \|x\| \\
 &= 2 \cdot \|x\|
 \end{aligned}$$

$$\hookrightarrow 0 \leq 2\|x\| \text{ } \Leftrightarrow \boxed{0 \leq \|x\|} \quad \blacksquare \text{ 1 απόδειξη}$$

$$2) \forall x, y \in \mathbb{R} \quad \|x - y\| \geq \left| \|x\| - \|y\| \right| \quad (\text{τριγωνική ανισότητα προς τα κάτω})$$

Απόδειξη

$$\|x\| = \|(x - y) + y\| \stackrel{(N3)}{\leq} \|x - y\| + \|y\| \quad \Leftrightarrow \boxed{\|x - y\| \geq \|x\| - \|y\|} \quad \textcircled{1}$$

$$\text{Αντίστοιχα προκύπτει } \boxed{\|y\| - \|x\| \leq \|x - y\|} \quad \textcircled{2}$$

$$\text{Από την } \textcircled{1}, \textcircled{2} \text{ προκύπτει } \left| \|x\| - \|y\| \right| \leq \|x - y\|. \quad \blacksquare \text{ 2 απόδειξη}$$

Παραδείγματα

$$1. (\mathbb{R}, \|\cdot\|) \text{ με } \|x\| = |x| \quad \checkmark (\text{εύρα νόρμα})$$

$$(\mathbb{C}, \|\cdot\|) \text{ με } \|z\| = |z| \quad \checkmark$$

$$2. (\mathbb{R}^n, \|\cdot\|_1) \text{ με } \|x\|_1 = \sum_{i=1}^n |x_i| = |x_1| + |x_2| + \dots + |x_n| \quad (\ell_1\text{-νόρμα})$$

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

Απόδειξη

$$(N1): x \in \mathbb{R}^n, \|x\|_1 = 0 \Leftrightarrow \sum_{i=1}^n |x_i| = 0 \Leftrightarrow |x_i| = 0, i=1, 2, \dots, n \Leftrightarrow x = 0 \quad \blacksquare N1$$

$$(N2): \lambda \in \mathbb{R}, x \in \mathbb{R}^n$$

$$\|\lambda x\|_1 = \sum_{i=1}^n |\lambda x_i| = \sum_{i=1}^n |\lambda| \cdot |x_i| = |\lambda| \sum_{i=1}^n |x_i| = |\lambda| \cdot \|x\|_1 \quad \blacksquare N2$$

$$(N3) x, y \in \mathbb{R}^n, \|x + y\|_1 = \sum_{i=1}^n |x_i + y_i| \leq \sum_{i=1}^n (|x_i| + |y_i|) = \sum_{i=1}^n |x_i| + \sum_{i=1}^n |y_i|$$

$$= \|x\|_1 + \|y\|_1 \quad \blacksquare N3$$

■ 2 παράμα

3. $(\mathbb{R}^n, \|\cdot\|_\infty)$ με $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$ (∞ -νόρμη ή νόρμη μεγίστου)

Απόδειξη

πολύ εύκολη ... ■

4. $(\mathbb{R}^n, \|\cdot\|_2)$ με $\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2}$ (l_2 -νόρμη ή ευκλείδεια νόρμη)

Απόδειξη

(N1), (N2) : απλές ■

(N3) : θεωρούμε το εσωτερικό γινόμενο $(\cdot, \cdot)_2 : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$

$$(x, y)_2 = \sum_{i=1}^n x_i y_i$$

Προφανώς, $\|x\|_2 = \sqrt{(x, x)_2} \quad \forall x \in \mathbb{R}^n$.

Ανισότητα των Cauchy-Schwarz

$$\forall x, y \in \mathbb{R}^n \quad |(x, y)_2| \leq \|x\|_2 \cdot \|y\|_2 \quad \left(\left| \sum_{i=1}^n x_i y_i \right| \leq \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2} \left(\sum_{i=1}^n |y_i|^2 \right)^{1/2} \right)$$

Απόδειξη ανισότητας

• Για $y=0$ είναι προφανές.

• Για $y \neq 0$ έχουμε: ~~0~~ $0 \leq (x - \lambda y, x - \lambda y)_2 = (\|x\|_2)^2 + (-\lambda y, x)_2 +$

$$\underbrace{(x, -\lambda y)_2} + \underbrace{(-\lambda y, -\lambda y)_2} =$$

$$= (\|x\|_2)^2 - 2\lambda (x, y)_2 + \lambda^2 (\|y\|_2)^2 = \underbrace{(\|y\|_2)^2}_{\neq 0} \lambda^2 - 2(x, y)_2 \lambda + (\|x\|_2)^2$$

Επειδή το τρίγωνο δεν αλλάζει πρόσημο έχει μη θετικά ~~αριθμο~~ Δ , δηλαδή

$$\Delta = (2(x, y)_2)^2 - 4\|x\|_2^2 \|y\|_2^2 \leq 0$$

Άρα, $4(x, y)_2^2 \leq 4\|x\|_2^2 \|y\|_2^2 \Rightarrow |(x, y)_2| \leq \|x\|_2 \cdot \|y\|_2$ ■ Cauchy-Schwarz.

Τριγωνική ανισότητα

καίτουμε τις πράξεις αναλυτικά και βρίσκουμε...

$$x, y \in \mathbb{R}^n \quad \|x+y\|_2^2 = (x+y, x+y)_2 = \|x\|_2^2 + 2(x, y)_2 + \|y\|_2^2$$

$$\leq \|x\|_2^2 + 2\|x\|_2 \|y\|_2 + \|y\|_2^2 = (\|x\|_2 + \|y\|_2)^2$$

↑
CS

$$\Rightarrow \boxed{\|x+y\|_2 \leq \|x\|_2 + \|y\|_2} \quad \blacksquare \text{ N3.}$$

5. $(\mathbb{R}^n, \|\cdot\|_p)$ με $p=1, \infty, 2$
 Η $\|\cdot\|_2$ παράγεται από το εσωτερικό γινόμενο $(x, y)_2 = \sum_{i=1}^n x_i \bar{y}_i$.

$$(x, y)_2 = \overline{(y, x)}_2 \blacksquare$$

6. $-\infty < a < b < +\infty$ $(C[a, b], \|\cdot\|_\infty)$

$$\|f\|_\infty = \max_{a \leq x \leq b} |f(x)| \quad (\text{νόρμα μεγίστου})$$

Απόδειξη

(N1) $f \in C[a, b]$, $\|f\|_\infty = 0 \Leftrightarrow \max_{a \leq x \leq b} |f(x)| = 0$

$$\Leftrightarrow f(x) = 0 \quad \forall x \in [a, b]$$

$$\Leftrightarrow f = 0 \quad \blacksquare_{N1}$$

(N2) $\lambda \in \mathbb{R}, f \in C[a, b]$

$$\|\lambda f\|_\infty = \max_{a \leq x \leq b} |\lambda f(x)| = |\lambda| \cdot \max_{a \leq x \leq b} |f(x)| = |\lambda| \cdot \|f\|_\infty \quad \blacksquare_{N2}$$

(N3) $f, g \in C[a, b]$, $\|f+g\|_\infty = \max_{a \leq x \leq b} |f(x)+g(x)|$

$$\stackrel{\substack{\text{για} \\ \text{κατάλληλο} \\ x^*}}{\leftarrow} = |f(x^*) + g(x^*)| \leq |f(x^*)| + |g(x^*)|$$

$$\leq \max_{a \leq x \leq b} |f(x)| + \max_{a \leq x \leq b} |g(x)| = \|f\|_\infty + \|g\|_\infty \quad \blacksquare_{N3}$$

Ορισμός (Ισοδυναμία νόρμών) ∞

Δύο νόρμες $\|\cdot\|$ και $\|\cdot\|'$ σε έναν γραμμικό χώρο \mathcal{X} λέγονται ισοδύναμες, αν υπάρχουν θετικές ~~αριθμοί~~ σταθερές m και M τ.ω. ανεξαρτήτως του x .

$$\forall x \in \mathcal{X}, m\|x\|' \leq \|x\| \leq M\|x\|' \quad \left(\text{οπότε, } \frac{1}{M}\|x\| \leq \|x\|' \leq \frac{1}{m}\|x\| \right)$$

Πρόταση

Όλες οι νόρμες στον \mathbb{R}^n είναι ισοδύναμες μεταξύ τους. (αυτό ισχύει σε κάθε χώρο πεπερασμένης διάστασης αλλά δεν ισχύει σε απειροδιάστατους χώρους).

Ορισμός (σύγκλιση ακολουθιών)

Έστω $(x^{(n)})_{n \in \mathbb{N}}$ μια ακολουθία σε έναν γραμμικό χώρο \mathcal{X} . Έστω μία $\|\cdot\|$ νόρμα στον \mathcal{X} . Λέμε ότι η ακολουθία $(x^{(n)})_{n \in \mathbb{N}}$ συγκλίνει στο x ως προς τη νόρμα $\|\cdot\|$, αν υπάρχει $\kappa \in \mathcal{X}$ (το όριο της ακολουθίας) τ.ω. $\|x^{(n)} - \kappa\| \rightarrow 0, n \rightarrow \infty \quad \blacksquare$

Παρατήρηση: Σέον \mathbb{R}^n αν μία ακολουθία συγκλίνει ως προς μία νόρμα, τότε θα συγκλίνει και ως οποιαδήποτε άλλη και μάλιστα στο ίδιο όριο, λόγω της ισοδυναμίας των νορμών. ■

Ιδιαίτεροι: Μία ακολουθία $(x^{(m)})_{m \in \mathbb{N}}$ του \mathbb{R}^n συγκλίνει σε έσοι $x \in \mathbb{R}^n$, ως προς μία οποιαδήποτε νόρμα, αν συγκλίνει ως προς τη νόρμα $\|\cdot\|_\infty$, δηλαδή $\max_{1 \leq i \leq n} |x_i^{(m)} - x_i| \rightarrow 0, m \rightarrow \infty$

$$\hat{\eta} \lim_{m \rightarrow \infty} x_i^{(m)} = x_i, i=1, \dots, n$$

Με άλλα λόγια, αν οι συνιστώσες $x_i^{(m)}$ συγκλίνουν στο x_i , για κάθε $i=1, \dots, n$. ■

Ορισμός (πλήρης χώρος)

Ένας χώρος X με νόρμα $\|\cdot\|$ λέγεται πλήρης, αν κάθε ακολουθία Cauchy ως προς τη νόρμα $\|\cdot\|$ συγκλίνει ως προς αυτή τη νόρμα, δηλαδή αν:

$$\forall \epsilon > 0 \exists k \in \mathbb{N} \forall k, m \geq N$$

$$\|x^{(k)} - x^{(m)}\| < \epsilon$$

$$\Leftrightarrow \exists x \in X \text{ τω } \|x^{(m)} - x\| \rightarrow 0 \text{ για } m \rightarrow \infty \quad \blacksquare$$

Θεώρημα

ο \mathbb{R}^n είναι πλήρης ως προς οποιαδήποτε νόρμα. (αυτό ισχύει για όλους τους χώρους πεπερασμένης διάστασης) ■



Νόρμες πινάκων

Μια απεικόνιση $\|\cdot\|: \mathbb{R}^{n \times n} \rightarrow \mathbb{R}, A \mapsto \|A\|$, αυτή λέγεται νόρμα πινάκων, αν:

$$(N1) \|A\| = 0 \Leftrightarrow A = 0$$

$$(N2) \forall \lambda \in \mathbb{R}, \forall A \in \mathbb{R}^{n \times n}, \|\lambda A\| = |\lambda| \cdot \|A\|$$

$$(N3) \forall A, B \in \mathbb{R}^{n \times n}, \|A+B\| \leq \|A\| + \|B\|$$

$$(N4) \forall A, B \in \mathbb{R}^{n \times n}, \|A \cdot B\| \leq \|A\| \cdot \|B\|$$

Υπάρχουν πολλές νόρμες πινάκων, εμείς θα ασχοληθούμε μόνο με μία κατηγορία, τις λεγόμενες φυσικές νόρμες πινάκων.

Έστω $\|\cdot\|$ στον \mathbb{R}^n . Για $A \in \mathbb{R}^{n,n}$ ορίζουμε τη νόρμα $\|A\|$ ως εξής:

$$\|A\| = \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|}{\|x\|}$$

Αυτή η ποσότητα είναι καλά ορισμένη και αποδεικνύεται ότι το sup μπορεί να αντικατασταθεί από το μέγιστο.

Αποδεικνύεται εύκολα ότι με αυτόν τον τρόπο ορίζεται μία νόρμα πινάκων.

Λέμε ότι αυτή η νόρμα πινάκων παράγεται (επαγωγικά) από τη νόρμα $\|\cdot\|$ του

\mathbb{R}^n . Συμβατικά ιδιότητες:

$$\forall x \in \mathbb{R}^n \quad \|Ax\| \leq \|A\| \cdot \|x\|$$

• για $x=0$ δίνει $0 \leq 0$ (σωστό)

• για $x \neq 0$ η ανισότητα γράφεται στη μορφή $\frac{\|Ax\|}{\|x\|} \leq \|A\|$

το οποίο λαμβάνει σύμφωνα με τον ορισμό της $\|A\|$.

Επιβαλητικές μέθοδοι

$$A \cdot x = b \quad , \quad A \in \mathbb{R}^{n,n} \text{ αψιτηρέγυφος και } b \in \mathbb{R}^n.$$

Οι επιβαλητικές μέθοδοι ξεκινούν με μία αρχική προσέγγιση $x^{(0)}$, συνήθως αυθαίρετη και δίνουν μία ακολουθία προσεγγίσεων $x^{(1)}, x^{(2)}, \dots$ της x .

Αυτή η ακολουθία συγχλίνει υπό κατάλληλες αυθαίρετες προϋποθέσεις.

• Μέθοδος Jacobi

• Μέθοδος Gauss-Seidel

Υπόθεση: τα διαγώνια στοιχεία του A είναι διάφορα του μηδενός.

$$a_{ii} \neq 0, \quad i=1, \dots, n$$

$$A \cdot x = b \Rightarrow (Ax)_i = b_i, \quad i=1, \dots, n \Rightarrow \sum_{j=1}^n a_{ij} x_j = b_i$$

$$\Rightarrow \sum_{j=1}^{i-1} a_{ij} x_j + \underbrace{a_{ii} x_i}_{\neq 0} + \sum_{j=i+1}^n a_{ij} x_j = b_i$$

$$\Rightarrow \underbrace{a_{ii}}_{\neq 0} x_i = b_i - \sum_{j=1}^{i-1} a_{ij} x_j - \sum_{j=i+1}^n a_{ij} x_j$$

$$\Rightarrow x_i = \frac{1}{a_{ii}} \left[b_i - \sum_{j=1}^{i-1} a_{ij} x_j - \sum_{j=i+1}^n a_{ij} x_j \right], \quad i=1, \dots, n$$

• Jacobi

$$x_i^{(m+1)} = \frac{1}{a_{ii}} \left[b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(m)} - \sum_{j=i+1}^n a_{ij} x_j^{(m)} \right]$$

• Gauss-Seidel

$$x_i^{(m+1)} = \frac{1}{a_{ii}} \left[b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(m+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(m)} \right]$$

Άρα, σύμφωνα με τω (*)

$$\|\Delta x\| \leq \frac{\|\bar{A}^{-1}\|}{1 - \|\bar{A}^{-1}\| \|\Delta A\|} (\|\Delta b\| + \|\Delta A\| \|x\|)$$

απαγκιστρώνω το $\|x\|$

$$\Rightarrow \frac{\|\Delta x\|}{\|x\|} \leq \frac{\|\bar{A}^{-1}\| \cdot \|\Delta A\|}{1 - \|\bar{A}^{-1}\| \|\Delta A\|} \left(\frac{\|\Delta b\|}{\|\Delta A\| \|x\|} + \|\bar{A}^{-1}\| \right)$$

τα πρόσθετα $\|\Delta A\|$.

\Rightarrow όπως λαμβάνει ότι $\|b\| \leq \|A\| \|x\|$ οπότε παίρουμε

$$\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \|\bar{A}^{-1}\| \|\Delta A\|} \left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right) \quad \blacksquare \text{ γερώτερο}$$

Απόδειξη βαθμιαίου αποτελέσματος:

Έχω $A + \Delta A = A(\underbrace{I + \bar{A}^{-1} \Delta A})$
 \uparrow αντιστρέψιμος

Θέτουμε $B = \bar{A}^{-1} \Delta A$. Έχουμε $\|B\| = \|\bar{A}^{-1} \Delta A\| \leq \|\bar{A}^{-1}\| \|\Delta A\| < 1$ (Σύμφωνα με τω διάσημη 3.36)

ο $(I_n + B)$ είναι αντιστρέψιμος και $\|(I_n + B)^{-1}\| \leq \frac{1}{1 - \|B\|}$ (**)

$A + \Delta A = A(I_n + B)$
 \uparrow το αποδ. ότι είναι αντιστρέψιμος αντιστρέψιμος

$\Rightarrow A + \Delta A$ αντιστρέψιμος και $(A + \Delta A)^{-1} = (I_n + B)^{-1} \cdot A^{-1}$

ενίσης, $\|A + \Delta A\| \leq \|(I_n + B)^{-1}\| \cdot \|A\|$ (**)

$\Rightarrow \|A + \Delta A\| \leq \frac{1}{1 - \|B\|} \cdot \|A\|$ αλλά

$\|B\| = \|\bar{A}^{-1} \Delta A\| \leq \|\bar{A}^{-1}\| \|\Delta A\| < 1 \rightarrow \leq \frac{1}{1 - \|\bar{A}^{-1}\| \|\Delta A\|} \|A\|$ \blacksquare

$(AB)^{-1} = B^{-1} A^{-1}$
 $(AB)(B^{-1} A^{-1}) =$
 $A(B B^{-1}) A^{-1} = A \cdot I_n \cdot A^{-1} = I_n$

• τέλος ύψους

Ποδείξη

α) το έχουμε κάνει πρῶτ.

Βοηθητικό αποτέλεσμα για τα (β) και (γ).

Ο $(A+\Delta A)$ είναι αντιστρέψιμος και ισχύει $\|(A+\Delta A)^{-1}\| \leq \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|}$. (θα το αποδείξω με μέτ.)

Λύση

Υποθέτουμε το βοηθητικό αποτέλεσμα και έχουμε:

β) $(A+\Delta A)(x+\Delta x) = b \Rightarrow (A+\Delta A)x + (A+\Delta A)\Delta x = b$

$\Rightarrow \cancel{Ax} + \Delta Ax + \cancel{(A+\Delta A)\Delta x} = b$
 $= b$

$\Rightarrow (A+\Delta A)\Delta x = -\Delta A \cdot x$

$\Rightarrow \Delta x = (A+\Delta A)^{-1} \cdot (-\Delta A \cdot x)$

$\Rightarrow \|\Delta x\| = \|(A+\Delta A)^{-1} \cdot (\Delta Ax)\| \leq \|(A+\Delta A)^{-1}\| \cdot \|\Delta A\| \cdot \|x\|$

$\Rightarrow \frac{\|\Delta x\|}{\|x\|} \leq \frac{\|(A+\Delta A)^{-1}\| \cdot \|\Delta A\|}{1}$
 $\leq \frac{\|A^{-1}\| \cdot \|\Delta A\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|}$

το βάζουμε εμπρός γιατί θέλουμε να φτιάξουμε το κ(A).

$\Rightarrow \frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \|A^{-1}\| \cdot \|\Delta A\|} \cdot \frac{\|\Delta A\|}{\|A\|}$ ■ βερτάσμα.

γ) $(A+\Delta A)(x+\Delta x) = b+\Delta b \Rightarrow (A+\Delta A)x + (A+\Delta A)\Delta x = b+\Delta b$

$\Rightarrow \cancel{Ax} + \Delta Ax + (A+\Delta A)\Delta x = \cancel{b} + \Delta b$
 $= b$

$\Rightarrow (A+\Delta A)\Delta x = \Delta b - \Delta Ax$

$\Rightarrow \Delta x = (A+\Delta A)^{-1} (\Delta b - \Delta Ax)$

$\Rightarrow \|\Delta x\| \leq \|(A+\Delta A)^{-1}\| \cdot \|\Delta b - \Delta Ax\|$
 $\leq \|\Delta b\| + \|\Delta A\| \cdot \|x\|$

Ο $\kappa(A) = \frac{\|Ax\|}{\|x\|} = \|A\| \cdot \|A^{-1}\|$ λέγεται δείκτης κατάστασης του A .

$$\frac{\|\Delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\Delta b\|}{\|b\|}$$

Η κατάσταση αυτή είναι βέλτεια (υπάρχει $\Delta b \neq 0$

τ.ω να ιαχθεί ως ακριβότητα)

Έπαιρε, $\kappa(A) = \|A\| \cdot \|A^{-1}\| \geq \|A^{-1}\| = \|I_n\| = 1 \Rightarrow \kappa(A) \geq 1$ ■

Αν ο $\kappa(A)$ είναι μικρός (όχι πολύ μεγαλύτερος του 1) το σύστημα έχει καλή κατάσταση, αν είναι πολύ μεγάλος έχει κακή κατάσταση.

$\kappa(A)$ μεγάλος
τότε έχει κακή
κατάσταση, αν
 $\kappa(A)$ μικρός, καλή
κατάσταση.

Θεώρημα (Εκτίμηση της σχετικής μεταβολής λύσεων γραμμικών συστημάτων)

Έστω $\|\cdot\|$ μία νόρμα στον \mathbb{R}^n και $\|\cdot\|$ αυτίστοιχη φυσική νόρμα στον $\mathbb{R}^{n,n}$. Έστω $A \in \mathbb{R}^{n,n}$ ένας αντιστρέψιμος πίνακας, $\Delta A \in \mathbb{R}^{n,n}$ και $b, \Delta b \in \mathbb{R}^n, b \neq 0$. Τότε, αν $\kappa(A) := \|A\| \cdot \|A^{-1}\|$,

έχουμε:

α) Αν $Ax = b$ και $A(x + \Delta x) = b + \Delta b$, τότε $\frac{\|\Delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\Delta b\|}{\|b\|}$
πρέπει να παραμείνει αντιστρέψιμος

β) Αν $Ax = b$ και $(A + \Delta A)(x + \Delta x) = b$, καθώς και $\|A^{-1}\| \cdot \|\Delta A\| < 1$, τότε ο $(A + \Delta A)$ είναι αντιστρέψιμος και $\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \|A^{-1}\| \cdot \|\Delta A\|} \frac{\|\Delta A\|}{\|A\|}$

γ) Αν $Ax = b$ και $(A + \Delta A)(x + \Delta x) = b + \Delta b$, και $\|A^{-1}\| \cdot \|\Delta A\| < 1$, τότε ο $(A + \Delta A)$ είναι αντιστρέψιμος και $\frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \|A^{-1}\| \cdot \|\Delta A\|} \left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right)$

τα (α), (β) είναι ειδικές περιπτώσεις του (γ). και οι 3 εκφράσεις είναι βέλτιστες.

ο $A^T A$ ως συμμετρικός πίνακας έχει πραγματικές ιδιοτιμές.

Δείκτης κατάστασης πίνακα (Turing)

θεωρούμε ένα σύστημα γραμμικό $Ax=b$ με $A \in \mathbb{R}^{n,n}$ αυτιστρέφσιμος για να έχει μοναδική λύση και $\|\cdot\|$ μία νόρμα στον \mathbb{R}^n .

Ερώτημα: τι μπορούμε να πούμε για την κατάσταση του $Ax=b$ (ως προς τη νόρμα $\|\cdot\|$), ιδιαίτερά για τη μεταβολή της λύσης x σε διαταραχές των δεδομένων A και b ;

* $Ax=b$ $\xrightarrow{b \neq 0}$ μεταβολή $\in \mathbb{R}^n$ } $\|\Delta x\|$ απόλυτο σφάλμα, $\frac{\|\Delta x\|}{\|x\|}$; σχετικό σφάλμα
 * $A(x+\Delta x) = b + \Delta b$

Λύση

$A(x+\Delta x) = b + \Delta b \Rightarrow \underbrace{Ax}_{=b} + A\Delta x = b + \Delta b \Rightarrow A\Delta x = \Delta b$
αυτό μας ενδιαφέρει

$\Rightarrow \boxed{\Delta x = A^{-1} \cdot \Delta b}$

$\|\Delta x\| = \|A^{-1} \cdot \Delta b\| \leq \|A^{-1}\| \cdot \|\Delta b\| \Rightarrow \boxed{\|\Delta x\| \leq \|A^{-1}\| \cdot \|\Delta b\|}$

$\Rightarrow \frac{\|\Delta x\|}{\|x\|} \leq \frac{\|A^{-1}\| \cdot \|\Delta b\|}{\|x\|}$

θεωτώ φέρω οπότε πρέπει να το εκταρύσω.

$\Rightarrow \boxed{\|x\| \geq \frac{\|b\|}{\|A\|}}$

μεταβολή στο διάστημα x

Άρα, $\frac{\|\Delta x\|}{\|x\|} \leq \frac{\|A^{-1}\| \cdot \|\Delta b\|}{\frac{\|b\|}{\|A\|}} \Leftrightarrow \frac{\|\Delta x\|}{\|x\|} \leq \underbrace{\|A^{-1}\| \cdot \|A\|}_{\text{μεταβολή στο διάστημα } b} \cdot \frac{\|\Delta b\|}{\|b\|}$

Νόρμες πινάκων

Έστω $\|\cdot\|$ μία νόρμα στον \mathbb{R}^n . Η αντίστοιχη παραγόμενη νόρμα πινάκων $\|\cdot\|: \mathbb{R}^{n,n} \rightarrow \mathbb{R}$,

$A \mapsto \|A\|$, ορίζεται ως εξής:

$$\|A\| = \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|}{\|x\|}$$

Σημαντικά ιδιότητες: $\forall x \in \mathbb{R}^n \quad \|Ax\| \leq \|A\| \cdot \|x\|$. Τέτοιες νόρμες πινάκων λέγονται φυσικές νόρμες.

Ερώτημα: Πώς υπολογίζουμε μία φυσική νόρμα πίνακα;

Το θέμα αυτό είναι γενικά δύσκολο. Αν $\forall x \in \mathbb{R}^n$ ισχύει $\|Ax\| \leq C_1 \|x\|$, τότε

$$\|A\| \leq C_1.$$

• Αν $\exists y \in \mathbb{R}^n, y \neq 0, \|Ay\| \geq C_2 \|y\| \Rightarrow \|A\| \geq C_2$.

Άρα, προσπαθούμε να αποδείξουμε τέτοιες ανισότητες με όσο μπορούμε μικρότερο C_1 και μεγαλύτερο C_2 . Αν επιτύχουμε $C_1 = C_2$, τότε $\|A\| = C_1 (= C_2)$. ■

Για τις νόρμες $\|\cdot\|_\infty, \|\cdot\|_1$ και $\|\cdot\|_2$ του \mathbb{R}^n αυτή η διαδικασία οδηγεί στα εξής αποτελέσματα:

• $\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$ (αθροίζουμε τις γραμμές) Νόρμα του αθροίσματος γραμμών.

• $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$ (αθροίζω τις στήλες) Νόρμα του αθροίσματος στηλών.

• $\|A\|_2 = (\rho(A^T \cdot A))^{1/2}$ Ευκλείδεια νόρμα ή φασματική νόρμα.

Έστω $B \in \mathbb{R}^{n,n}$ και $\lambda_1, \dots, \lambda_n$ οι ιδιοτιμές του. Τότε θέτουμε $\rho(B) = \max_{1 \leq i \leq n} |\lambda_i|$.

Το $\rho(B)$ λέγεται φασματική ακτίνα του B .

↓
είναι η ακτίνα
του μικρότερου
κύκλου που περιέχει
όλες τις ιδιοτιμές.

Διαισθητικά θα έλεγε κανείς ότι η δεύτερη μέθοδος είναι καλύτερη από την πρώτη.

Υπάρχει όμως μία κατηγορία γραμμικών συστημάτων, για τα οποία η δεύτερη μέθοδος συρρίνιει πολύ καλύτερα από την πρώτη.

Γενικά, όμως, υπάρχουν συστήματα για τα οποία η μία μέθοδος συρρίνιει ενώ η άλλη όχι.

Παράλληλα

Η μέθοδος του Jacobi μπορεί να υλοποιηθεί παράλληλα: χρησιμοποιώντας τα $b, A, x^{(m)}$ μπορεί σε κάθε επεξεργασία να υπολογιστούν μία συνιστώσα $x_i^{(m+1)}$ του $x^{(m+1)}$. Η μέθοδος των Gauss-Seidel δεν μπορεί να υλοποιηθεί παράλληλα.

Οι μέθοδοι υλοποιούνται στην μορφή που αναφέραμε. Για θεωρητικούς σκοπούς θα τις γράψουμε σε διαφορετική μορφή.

Θέτουμε,

$$L = \begin{pmatrix} 0 & & & & 0 \\ a_{21} & 0 & & & \\ a_{31} & a_{32} & 0 & & \\ \vdots & \vdots & & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & 0 \end{pmatrix}, \quad D = \begin{pmatrix} a_{11} & & & & \\ & a_{22} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & a_{nn} \end{pmatrix}$$

$$U = \begin{pmatrix} 0 & a_{12} & \dots & a_{1n} \\ & 0 & a_{23} & \dots & a_{2n} \\ & & & \ddots & \\ & 0 & & & a_{n-1,n} \\ & & & & 0 \end{pmatrix}$$

Τότε: $A = L + D + U$

Τότε:

Μέθοδος Jacobi:

$$D \cdot x^{(m+1)} = b - L \cdot x^{(m)} - U \cdot x^{(m)} \quad (\Leftrightarrow) \quad D \cdot x^{(m+1)} = -(L+U) \cdot x^{(m)} + b \quad (\Leftrightarrow) \quad x^{(m+1)} = -D^{-1} \cdot (L+U) \cdot x^{(m)} + b$$

Μέθοδος Gauss-Seidel:

το πάμε αριστερά γιατί έχει το $(m+1)$.

$$a_{ii} \cdot x_i^{(m+1)} = b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(m+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(m)}$$

$$\Rightarrow \sum_{j=1}^{i-1} a_{ij}^{(m+1)} x_j + a_{ii}^{(m+1)} x_i = b_i - \sum_{j=i+1}^n a_{ij}^{(m+1)} x_j^{(m)}$$

$$\Rightarrow \sum_{j=1}^i a_{ij}^{(m+1)} x_j = b_i - \sum_{j=i+1}^n a_{ij}^{(m+1)} x_j^{(m)} \quad i = 1, \dots, n$$

$$\Rightarrow (L+D)x^{(m+1)} = b - Ux^{(m)}$$

$$\Rightarrow x^{(m+1)} = (L+D)^{-1} b - (L+D)^{-1} U x^{(m)}$$

$$\Rightarrow x^{(m+1)} = - (L+D)^{-1} U x^{(m)} + (L+D)^{-1} b$$



Γενικά επαναληπτικά μέθοδοι:

$A = M - N$, M αντιστρέψιμος

Έχουμε, $A \cdot x = b \Leftrightarrow (M - N) \cdot x = b \Leftrightarrow M \cdot x - N \cdot x = b \Leftrightarrow \boxed{M \cdot x = N \cdot x + b}$

$$x^{(m)} \rightarrow x^{(m+1)}$$

$$M x^{(m+1)} = \underbrace{N x^{(m)} + b}_{\text{πράξεις}}$$

Για να έχουμε ένα γραμμικό σύστημα με πίνακα M . Για να μπορούμε να υλοποιήσουμε εύκολα την μέθοδο πρέπει τέτοια συστήματα να μπορούν να λυθούν εύκολα.

Στην περίπτωση Jacobi $M_J = D, N_J = -(L+U)$ ενώ στην περίπτωση G-S το $M_{GS} = L+D$ και $N_{GS} = -U$.

Ερώτημα: \Rightarrow τι μπορούμε να πούμε σχετικά με την σύγκλιση επαναληπτικών μεθόδων.

(Θέλω τα $x^{(m)}$ να συγκλίνουν στο x για αυθαίρετο $x^{(0)}$).

• Έστω $x^{(m)} \rightarrow \gamma, m \rightarrow \infty$

Ερώτημα: $\gamma = x$;

Λύση

$$\text{Έχουμε } M x^{(m+1)} = N x^{(m)} + b, m \rightarrow \infty$$

$$M \cdot \gamma = N \gamma + b$$

Άρα, έχουμε ότι $M\gamma = N\gamma + b \Leftrightarrow M\gamma - N\gamma = b \Leftrightarrow (M-N)\gamma = b \Leftrightarrow A\gamma = b$

$$\Leftrightarrow \boxed{\gamma = x}$$

Απάντηση: ΝΑΙ ■

• Μένει να δούμε τώρα πότε τέτοιες ακολουθίες $(x^{(m)})_{m \in \mathbb{N}_0}$ συγκλίνουν.

$$\left. \begin{aligned} \text{Έχουμε } M \cdot x^{(m+1)} &= N \cdot x^{(m)} + b \\ M x &= N \cdot x + b \end{aligned} \right\}$$

Άρα, (αφααιρούμε κατά μέλη)

$$M(x^{(m+1)} - x) = N(x^{(m)} - x) \Leftrightarrow x^{(m+1)} - x = \underbrace{\begin{pmatrix} -1 \\ M \cdot N \end{pmatrix}}_{\text{⊛}} (x^{(m)} - x) \quad \text{⊛}$$

$\text{⊛} = G$. Ο G λέγεται πίνακας επανάληψης της μεθόδου.

Γράφουμε την ⊛ στη μορφή: $x^{(m+1)} - x = G(x^{(m)} - x)$ άρα επαγωγικά συμπεραίνουμε ότι

$$\boxed{x^{(m)} - x = G^{(m)}(x^{(0)} - x), m \in \mathbb{N}_0}$$

• $\|x^{(m)} - x\| \leq \|G^{(m)}\| \cdot \|x^{(0)} - x\|$ (για κατάλληλο $x^{(0)}$ αυτή η εκτίμηση ισχύει ως ποσοτήτα). Συμπέρασμα, Η μέθοδος συγκλίνει, δηλαδή για οποιαδήποτε αρχικά τμή $x^{(0)} \in \mathbb{R}^n$, έχουμε $x^{(m)} \rightarrow x, m \rightarrow \infty$, αν και μόνο αν $\|G^{(m)}\| \rightarrow 0, m \rightarrow \infty$, δηλαδή

$$\boxed{\|G^{(m)}\| \rightarrow 0, m \rightarrow \infty} \quad \text{■ ερωτήματος}$$

$Ax = b, A \in \mathbb{R}^{n \times n}$ αυστηρά γινόμενο

$A = M - N, M$ αυστηρά γινόμενο

Γενικό επαναληπτικό μέθοδος: $M \cdot x^{(m+1)} = N \cdot x^{(m)} + b, m \in \mathbb{N}_0$ ④

$G = M^{-1} \cdot N$ πίνακας επανάληψης της μεθόδου.

$$x^{(m+1)} - x^{(m)} = G(x^{(m)} - x^{(m-1)})$$

και

$$x^{(m)} - x^{(0)} = G^m (x^{(0)} - x^{(0)})$$

$$\boxed{x^{(0)} \text{ αυθαίρετο}}$$

$P \in \mathbb{C}^{n \times n}, \lambda_1(P), \dots, \lambda_n(P) \in \mathbb{C}$ ιδιοτιμές του P .

$\rho(P) = \max_{1 \leq i \leq n} |\lambda_i(P)|$ φασματική ακτίνα του P .

Λήμμα (Σχέση φυσικής νόρμας και φασματικής ακτίνας πίνακα)

Έστω $\|\cdot\|$ μία νόρμα στον \mathbb{C}^n και u επαγόμενου νόρμα στον $\mathbb{C}^{n \times n}$. Τότε, για $P \in \mathbb{C}^{n \times n}$ ισχύει $\rho(P) \leq \|P\|$. Αντίστροφα, για κάθε $P \in \mathbb{C}^{n \times n}$ και $\epsilon > 0$ υπάρχει φυσική νόρμα $\|\cdot\|$ τ.ω $\|P\| \leq \rho(P) + \epsilon$.

Απόδειξη

(Μόνο για το μέρος *).

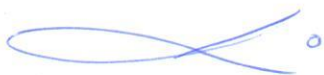
Έστω λ ιδιοτιμή του P και z αντίστοιχο ιδιοδιάνυσμα.

Τότε $P \cdot z = \lambda \cdot z$ ($z \neq 0$) άρα $\|P \cdot z\| = \|\lambda \cdot z\| \Rightarrow |\lambda| \cdot \|z\| \leq \|P \cdot z\| \leq \|P\| \cdot \|z\|$

$$\Rightarrow \boxed{|\lambda| \leq \|P\|}$$

Αφού αυτό u αυθόρμητα ισχύει για τις ιδιοτιμές του P , συμπεραίνουμε ότι

$\max_{1 \leq i \leq n} |\lambda_i(P)| \leq \|P\|$ δηλαδή ισχύει * ④. ■ Λήμμα



Θεώρημα (Ικανή & αναγκαίες συνθήκες σύγκλισης επαναληπτικών μεθόδων)

Έστω x η λύση του συστήματος $Ax = b$. Τα παρακάτω είναι ισοδύναμα:

α) Η επαναληπτική μέθοδος ④ συγκλίνει, δηλαδή για κάθε $x^{(0)} \in \mathbb{C}^n$ έχουμε

$$x^{(m)} \rightarrow x, m \rightarrow \infty$$

β) Η φασματική ακτίνα $\rho(G) < 1$ όπου $G = M^{-1} \cdot N$ ο πίνακας επανάληψης της μεθόδου.

γ) Υπάρχει φυσικό νόρμα πινάκων $\|\cdot\|$ τ.ω $\|\Gamma\| < 1$.

$$\delta) \lim_{m \rightarrow \infty} \Gamma^m = 0$$

Απόδειξη α) \Rightarrow β) \Rightarrow γ) \Rightarrow δ) \Rightarrow α)

(α \Rightarrow β):

Γνωρίζουμε ότι $x^{(m)} - x = \Gamma^m (x^{(0)} - x)$ και ότι $\|x^{(m)} - x\| \rightarrow 0, m \rightarrow \infty$ για οποιοδήποτε $x^{(0)}$. Άρα, ισχύει ότι

$$\|\Gamma^m (x^{(0)} - x)\| \rightarrow 0, m \rightarrow \infty.$$

Έστω λ ιδιοτιμή του πίνακα Γ και z αντίστοιχο ιδιοδιάνυσμα. Για $x^{(0)} = x + z$,

η προηγούμενη σχέση δίνει $\|\Gamma^m z\| \rightarrow 0, m \rightarrow \infty$.

$$\Gamma z = \lambda z \Rightarrow \Gamma^2 z = \lambda \cdot \Gamma z = \lambda^2 z \Rightarrow \dots \Gamma^m z = \lambda^m z = \lambda z$$

Άρα, $\|\lambda^m z\| \rightarrow 0, m \rightarrow \infty$ δηλαδή $|\lambda|^m \cdot \|z\| \rightarrow 0, m \rightarrow \infty \Rightarrow |\lambda|^m (\|z\|) \rightarrow 0, m \rightarrow \infty$

$\Rightarrow |\lambda|^m \rightarrow 0, m \rightarrow \infty$. Επομένως, $|\lambda| < 1$. Επειδή, το λ ήταν αυθαίρετη ιδιοτιμή

του Γ , συμπεραίνουμε ότι $\rho(\Gamma) < 1$. ■ α \Rightarrow β.

(β \Rightarrow γ):

Αφού $\rho(\Gamma) < 1$, \exists θετικό ε τ.ω $\rho(\Gamma) + \varepsilon < 1$. Σύμφωνα με το προηγούμενο ημίμα,

υπάρχει φυσικό νόρμα $\|\cdot\|$ τ.ω $\|\Gamma\| \leq \rho(\Gamma) + \varepsilon$. Από τις δύο τελευταίες

αυτιότητες έπεται ότι $\|\Gamma\| < 1$. ■ β \Rightarrow γ.

(γ \Rightarrow δ):

$$\|\Gamma^m\| = \|\underbrace{\Gamma \cdot \Gamma \cdot \dots \cdot \Gamma}_{m \text{ φορές}}\| \leq \underbrace{\|\Gamma\| \cdot \|\Gamma\| \cdot \dots \cdot \|\Gamma\|}_{m \text{ φορές}} = \|\Gamma\|^m \rightarrow 0, \text{ για } m \rightarrow \infty \text{ αφού}$$

$$\|\Gamma\| < 1. \quad \text{■ } \gamma \Rightarrow \delta$$

(δ \Rightarrow α):

$$x^{(m)} - x = \Gamma^m (x^{(0)} - x) \Rightarrow \|x^{(m)} - x\| \leq \|\Gamma^m\| \cdot \|x^{(0)} - x\| \rightarrow 0, m \rightarrow \infty$$

Άρα, $x^{(m)} - x \rightarrow 0, m \rightarrow \infty$ δηλαδή $x^{(m)} \rightarrow x, m \rightarrow \infty$ ■ δ \Rightarrow α

Παράδειγμα

(συνήθες)

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

$$\|G\|_1 = 2$$

$$\|G\|_\infty = 2 \text{ (γραμμική max)}$$

$$\|G\|_2 = 2$$

Σημείωση ότι συγχλίπει? Δεν μπορούμε να το απαντήσουμε

Το $G^2 = 0$ άρα συγχλίπει. ■

$$A = (a_{ij})_{i,j=1,\dots,n} \in \mathbb{C}^{n,n}$$

Λέμε ότι ο A έχει αυστηρά κυρίαρχα διαγώνια, αν $|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, i=1,\dots,n$.

Τέτοιοι πίνακες είναι ασυμπλέγσιμοι και οι μέθοδοι του Jacobi και του Gauss-Seidel συγχλίσουν για συστήματα $Ax=b$ με τέτοιους πίνακες A .

Λήμμα (αυστότητα του Gershgorin)

Έστω $A = (a_{ij}) \in \mathbb{C}^{n,n}$ ένας πίνακας και $\lambda \in \mathbb{C}$ μία ιδιοτιμή του. Τότε υπάρχει $s \in \{1, 2, \dots, n\}$ τέτοια ώστε $|\lambda - a_{ss}| \leq \sum_{\substack{j=1 \\ j \neq s}}^n |a_{sj}| =: r_s$

Απόδειξη

Έστω $z \in \mathbb{C}^n$ αυθεντικό ιδιοδιάνοσμα, $Az = \lambda z$.

$$\text{Τότε } (Az)_i = \lambda z_i \quad i=1,\dots,n \quad \Leftrightarrow \quad \sum_{j=1}^n a_{ij} z_j = \lambda z_i$$

$$\Leftrightarrow \underbrace{a_{ii} z_i - \lambda z_i}_{(a_{ii} - \lambda) z_i} = - \sum_{\substack{j=1 \\ j \neq i}}^n a_{ij} z_j \quad \Rightarrow \quad |a_{ii} - \lambda| \cdot |z_i| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| |z_j| \quad i=1,\dots,n \quad (*)$$

Έστω $s \in \{1, \dots, n\}$ τέτοια ώστε $|z_s| = \max_{1 \leq j \leq n} |z_j| \neq 0$

$$\text{Για } i = s, (*) \text{ δίνει } |a_{ss} - \lambda| \cdot |z_s| \leq \sum_{\substack{j=1 \\ j \neq s}}^n |a_{sj}| \cdot |z_j| \leq |z_s|$$

$$\Rightarrow |a_{ss} - \lambda| \cdot |z_s| \leq \left(\sum_{\substack{j=1 \\ j \neq s}}^n |a_{sj}| \right) \cdot |z_s| \quad \Rightarrow \quad |a_{ss} - \lambda| \leq \sum_{\substack{j=1 \\ j \neq s}}^n |a_{sj}| \quad \blacksquare \text{ λήμμα}$$

Έστω $A \in \mathbb{C}^{n \times n}$ πίνακας με αυστηρά κυριαρχικά διαγώνιο,

$$|a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \quad i=1, \dots, n$$

Απόδειξη (θέλω να δείξω ότι ο A είναι αυστηρώς)

Έστω ότι ο A δεν είναι αυστηρώς. Τότε υπάρχει $z \in \mathbb{C}, z \neq 0$ τω $Az=0$ δαδαδι $A \cdot z = 0 \cdot z$, δαδαδι το $\lambda=0$ είναι ιδιοτιμή του.

Τότε, με $\lambda=0$, σύμφωνα με το προηγούμενο λήμμα υπάρχει $s \in \{1, \dots, n\}$ τω

$$|a_{ss}-0| \leq \sum_{\substack{j=1 \\ j \neq i}}^n |a_{sj}| \quad (\text{ΑΤΟΝΟ}) \blacksquare$$

Θεώρημα (Σύγκριση των μεθόδων του Jacobi και των Gauss-Seidel για πίνακες με αυστηρά κυριαρχικά διαγώνιο)

Έστω A ένας πίνακας με αυστηρά κυριαρχικά διαγώνιο. Τότε ισχύει:

α) Για τους πίνακες επαναλήψας $G_J = -D^{-1}(L+U)$ και $G_{GS} = -(L+D)^{-1} \cdot U$ τω μεθόδω του Jacobi και τω Gauss-Seidel αντίστοιχα, ισχύει

$$\|G_J\|_\infty < 1, \quad \|G_{GS}\|_\infty < 1.$$

β) Οι μέθοδοι συγκλίνουν για γραμμικά συστήματα με πίνακα συντελεστών τον A.

Απόδειξη

ο πίνακας είναι αυστηρώς και $a_{ii} \neq 0, i=1, \dots, n$ οπότε οι μέθοδοι μπορούν να εφαρμοστούν.

α) θέτουμε $G = \max_{1 \leq i \leq n} \left(\frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right)$. προφανώς $G < 1$.

θα αποδείξουμε ότι ισχύει τα εξής: $\|G_J\|_\infty = G$ και $\|G_{GS}\|_\infty \leq G$.

$$D^{-1} = \begin{pmatrix} 1/a_{11} & & & & \\ & 1/a_{22} & & & \\ & & \ddots & & \\ & & & \ddots & \\ & & & & 1/a_{nn} \end{pmatrix}, \quad L+U = \begin{pmatrix} 0 & a_{12} & \dots & \dots & a_{1n} \\ a_{21} & 0 & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & \dots & 0 \end{pmatrix}$$

$$G_J = - \begin{pmatrix} 0 & \frac{a_{12}}{a_{11}} & \dots & \frac{a_{1n}}{a_{11}} \\ \frac{a_{21}}{a_{22}} & 0 & \frac{a_{23}}{a_{22}} & \frac{a_{2n}}{a_{22}} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{a_{n1}}{a_{nn}} & \dots & \dots & 0 \end{pmatrix}$$

$$\|G_J\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1, j \neq i}^n \frac{|a_{ij}|}{|a_{ii}|}$$

$$= \max_{1 \leq i \leq n} \left(\frac{1}{|a_{ii}|} \sum_{j=1, j \neq i}^n |a_{ij}| \right) = G \quad \text{Jacobi}$$

Gauss-Seidel

$$G_{GS} = -(L+D)^{-1} \cdot U$$

Έστω $y \in \mathbb{C}^n$, ~~αντικείμενο~~ θέτουμε $u := G_{GS} y$ οπότε έχουμε $u = -(L+D)^{-1} U y$

$$\Leftrightarrow (L+D)u = -Uy$$

$$(L+D)u = -Uy \Rightarrow ((L+D)u)_i = -(Uy)_i \quad i=1, \dots, n$$

$$\Rightarrow \sum_{j=1}^i a_{ij} u_j = - \sum_{j=i+1}^n a_{ij} y_j$$

$$\Rightarrow a_{ii} u_i = - \sum_{j=1}^{i-1} a_{ij} u_j - \sum_{j=i+1}^n a_{ij} y_j \Rightarrow u_i = - \frac{1}{a_{ii}} \left[\sum_{j=1}^{i-1} a_{ij} u_j + \sum_{j=i+1}^n a_{ij} y_j \right]$$

$$i=1, \dots, n$$

Παροπισμός

$$|u_i| \leq G \|y\|_\infty, \quad i=1, \dots, n \quad \textcircled{*}$$

Απόδειξη

Επαγωγικά: Για $i=1$:

$$u_1 = - \frac{1}{a_{11}} \sum_{j=2}^n a_{1j} y_j \Rightarrow$$

$$|u_1| \leq \frac{1}{|a_{11}|} \sum_{j=2}^n |a_{1j}| |y_j| \stackrel{\leq \|y\|_\infty}{\Rightarrow}$$

$$|u_1| \leq \left(\frac{1}{|a_{11}|} \sum_{j=2}^n |a_{1j}| \right) \|y\|_\infty \Rightarrow$$

$$|u_1| \leq G \|y\|_\infty \quad \checkmark$$

Εστω ότι η \oplus λαμβεί για $i=1, \dots, k-1$. Θα αποδείξουμε ότι λαμβεί για $i=k$.

Πράγματι,

$$|u_k| \leq \frac{1}{|a_{kk}|} \left[\sum_{j=1}^{k-1} |a_{kj}| |u_j| + \sum_{j=k+1}^n |a_{kj}| |y_j| \right]$$

$\leq d \|y\|_\infty$ $\leq \|y\|_\infty$

$$\Rightarrow |u_k| \leq \frac{1}{|a_{kk}|} \left[\sum_{j=1}^{k-1} |a_{kj}| \cdot d + \sum_{j=k+1}^n |a_{kj}| \right] \|y\|_\infty$$

$\downarrow < 1$

$$\Rightarrow |u_k| \leq \frac{1}{|a_{kk}|} \left[\sum_{j=1}^{k-1} |a_{kj}| + \sum_{j=k+1}^n |a_{kj}| \right] \|y\|_\infty$$

$\underbrace{\sum_{j=1, j \neq k}^n |a_{kj}|}$

$$\Rightarrow |u_k| \leq \underbrace{\left(\frac{1}{|a_{kk}|} \sum_{j=1, j \neq k}^n |a_{kj}| \right)}_{\leq d} \|y\|_\infty$$

$\Rightarrow |u_k| \leq d \cdot \|y\|_\infty$ δηλαδή η \oplus λαμβεί και για $i=k$. ■ επαγωγικά
 Η \oplus δίνει $\max_{1 \leq i \leq n} |u_i| \leq d \cdot \|y\|_\infty$ δηλαδή $\|G_S y\|_\infty \leq d \|y\|_\infty \quad \forall y \in \mathbb{C}^n$
 $\|u\|_\infty$

$$\Rightarrow \|G_S\|_\infty \leq d \quad \blacksquare$$

β) Έπεται από τω α) χρησιμοποιώντας το θεώρημα για τη σύγκλιση επαναληπτικών μεθόδων. ■

