19/3/2015

3º ΓΡΑΜΜΙΚΑ ΣΥΣΤΗΜΑΤΑ

Δεδομένα: $A \in \mathbb{R}^{n,n}$ nxn πίνακας
$b \in \mathbb{R}^n$ διάνυσμα με n συνιστώσα
Ζητούμενα: $x \in \mathbb{R}^n$ τ.ω. $Ax = b$

Τέτοια γραμμικά συστήματα προκύπτουν πολύ συχνά στις εφαρμογές, συνήθως ως μέρος συνθετότερων προβλημάτων.

Θα μας απασχολήσουν τα θέματα
α) Αριθμητικές μέθοδοι για την προσέγγιση του x
- κόστος (απαιτούμενες πράξεις και απαιτούμενη μνήμη)
- ευστάθεια μεθόδων

β) Κατάσταση γραμμικών συστημάτων.

Υπάρχουν 2 μεγάλες κατηγορίες αριθμητικών μεθόδων:
α) Άμεσες (παραλλαγές της μεθόδου απαλοιφής του Gauss)
Όταν οι πράξεις γίνονται ακριβώς, δίνουν τη λύση ακριβώς με πεπερασμένο πλήθος.
β) Επαναληπτικές: Δίνουν μια ακολουθία προσεγγίσεων $x^{(0)}, x^{(1)} \ldots$ της x

"Γενικά για γραμμικά συστήματα"

Δεδομένα: $a_{ij} \in \mathbb{R}$ $i, j = 1, 2, \ldots n$ συντελεστές.
$b_i \in \mathbb{R}$ $i = 1, 2, \ldots n$ δεύτερα μέλη.
Ζητούμενα: $x_1, \ldots, x_n$ $n \in \mathbb{R}$

$$\left.\begin{array}{l} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + \quad = b_2 \\ \vdots \\ a_{n1}x_1 + \cdots + a_{nn}x_n = b_n \end{array}\right\} (*)$$

Με τον πίνακα A

$$A = (a_{ij})_{i,j=1,2,\ldots,n} = \begin{pmatrix} a_{11} & a_{12} & a_{\ldots} & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & & \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$ και τα διανύσματα

$$b = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix} \quad \text{και} \quad x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$$

το γραμμικό σύστημα(*) γράφεται στη μορφή $Ax = b$

Θα ασχοληθούμε με γραμμικά συστήματα που έχουν ακριβώς μια λύση.

Κάθε μια από τις ακόλουθες συνθήκες ικανή και αναγκαία για να έχει το $Ax=b$ ακριβώς μια λύση.

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

ii) $\det A \neq 0$

iii) $Ax = 0 \Rightarrow x = 0$! Το αντίστροφο δεν ισχύει.

iv) Οι γραμμές ή οι στήλες του A είναι γραμμικά ανεξάρτητα διανύσματα

Θυμάμαι:

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

α) κανόνας του Crammer: Γράφουμε τον A στη μορφή $A = (a^1, \ldots, a^n)$ με a^i την i-οστή στήλη του A, και θέτουμε:

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

Τότε: $x_i = \frac{\det A_i}{\det A}$ $i = 1, 2, \ldots, n$

β) με τον A^{-1}: Υπολογίζουμε την A^{-1} και έχουμε $x = A^{-1}b$

γ) μέθοδος του Gauss (αναλυτικά πίσω →)

Παρατήρηση:
Οι α, β τρόποι έχουν μόνο θεωρητική σημασία, όχι πρακτική.

ⓐ Grammer
Αναπτύσσοντας ως προς τη στήλη j έχουμε
$\det A = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det A_{ij}$ με $A_{ij} \in \mathbb{R}^{n-1,n-1}$ που προκύπτει από τον A αν διαγράψουμε την i γραμμή και j στήλη.
Συνεχίζοντας με τον τρόπο αυτό γράφουμε την $\det A$ ως άθροισμα $n!$ όρων με n παράγοντες ο καθένας. Συνολικά απαιτούνται τότε $n!(n-1)$ πολλαπλασιασμοί.
Αυτός ο αριθμός αυξάνει πάρα πολύ γρήγορα με το n.
Στον κανόνα του Grammer πρέπει να υπολογίσουμε $n+1$ ορίζουσες $n \times n$ πινάκων. Συνολικά απαιτούνται $(n+1)!(n-1)$ πολλαπλασιασμοί.

ⓑ $x = A^{-1} b$
Σπάνια στην πράξη απαιτείται ο υπολογισμός του A^{-1}. Όταν χρειάζεται, αυτό γίνεται ως εξής. Έστω $e^1, e^2, \dots, e^n$ η κανονική βάση του $\mathbb{R}^n$

$$e^i_j = \delta_{ij} = \begin{cases} 1 & \text{για } j=i \\ 0 & \text{για διαφορετικά} \end{cases}$$

(δέλτα του Krmecker ή σύμβολο Krmecker)
Έστω $u^1, u^2, \dots, u^n \in \mathbb{R}^n$ τ.ω. $Au^i = e^i$ $i = 1, 2, \dots, n$
Ισχυρισμός: $A^{-1} = (u^1 \dots, u^n)$
Πράγματι έχουμε $AA^{-1} = A(u^1 \dots u^n) = (Au^1, Au^2, \dots, Au^n) = (e^1, \dots, e^n)$
$(e^1, e^2, \dots, e^n) = I_n$
Για να υπολογίσουμε τον A^{-1} πρέπει να λύσουμε n γραμμικά συστήματα.
Αν θέλαμε να χρησιμοποιήσουμε τον A^{-1} για να λύσουμε $Ax=b$, απαιτείται επί πλέον ο υπολογισμός του $A^{-1}b$, $x = A^{-1}b$.

"Δύο μεγάλες κατηγορίες γραμμικών συστημάτων (πινάκων)"

α) Πυκνοί (ή αποθηκεύσιμοι) πίνακες: Έχουν στοιχεία a_{ij} γενικά διάφορα του μηδενός
β) Αραιοί (ή σποραδικοί) πίνακες: Έχουν πολλά μηδενικά στοιχεία που αν τα εκμεταλλευτούμε αποκομίζουμε υπολογιστικά οφέλη.
(Wilkinson, Turing)

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

$n < 100$: μικροί πίνακες
$100 \leq n < 1000$: μεσαίοι πίνακες (μετρίου μεγέθους)
$n \geq 1000$: μεγάλοι πίνακες
Στην πράξη λύνουμε συστήματα με πυκνούς πίνακες μέχρι μετρίου μεγέθους, ενώ με αραιούς πίνακες ακόμα και μεγάλου μεγέθους.

Οι άμεσες μέθοδοι χρησιμοποιούνται κυρίως για συστήματα με πυκνούς πίνακες ενώ οι επαναληπτικές για συστήματα με αραιούς πίνακες.

Ⓐ Η μέθοδος απαλοιφής του Gauss
Έστω $U = (u_{ij}) \in \mathbb{R}^{n,n}$ ένας αντιστρέψιμος άνω τριγωνικός (ή κάτω) πίνακας, δηλαδή τ.ω. $u_{ij} = 0$ για $i > j$
Έχουμε: $\det U = u_{11} \cdot u_{22} \cdots u_{nn}$
Συμπέρασμα: Ο U είναι ένας αντιστρέψιμος αν και μόνο αν όλα τα διαγώνια στοιχεία του είναι μη μηδενικά.
Έστω $y \in \mathbb{R}^n$. Θεωρούμε το γραμμικό σύστημα $Ux = y$, δηλαδή

$$\left.\begin{aligned}
u_{11}x_1 + u_{12}x_2 + \cdots + u_{1n}x_n &= y_1 \\
u_{21}x_1 + u_{22}x_2 + \cdots + u_{2n}x_n &= y_2 \\
&\;\;\vdots \\
u_{nn}x_n &= y_n
\end{aligned}\right\}$$

Αυτό λύνεται εύκολα με οπισθοδρόμηση $x_n = \frac{y_n}{u_{nn}}$ για $k = n-1, n-2, \dots, 1$

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

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

24/3/2015

"Η μέθοδος της απαλοιφής του Gauss"

$u \in \mathbb{R}^{n,n}$ αντιστρέψιμος άνω τριγωνικός

$Ux = y \Rightarrow$

$$\left.\begin{array}{r} u_{11}x + u_{12}x + \dots + u_{1n}x = y_1 \\ u_{22}x + \dots + u_{2n}x = y_2 \\ \vdots \\ u_{nn}x = y_n \end{array}\right\}$$

"Αλγόριθμος της οπισθοδρόμησης"

$x_n = \frac{y_n}{u_{nn}}$ για $k = n-1, n-2, \dots, 1$

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

Απαιτούμενες πράξεις

1. διαιρέσεις : n

2. πολλαπλασιασμός και προσθέσεις:

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

3. Θέσεις μνήμης:
$\frac{n^2}{2} + O(n)$ για τα u_{ij} $(j=i)$ και y_i
(Τα x_k αποθηκεύονται στις θέσεις των y_k)

"Ιδέα στη μέθοδο απαλοιφής του Gauss"

Με κατάλληλους μετασχηματισμούς γραμμών (δηλαδή εναλλαγή δύο εξισώσεων ή πρόσθεση κατά μέλη ενός πολλαπλάσιου μιας εξίσωσης σε μια άλλη και αντικατάσταση της τελευταίας με το αποτέλεσμα), μετατρέπουμε ένα γραμμικό σύστημα $Ax=b$ σε ένα ισοδύναμο $Ux=y$ με άνω τριγωνικό πίνακα U (τριγωνοποίηση!). Ύστερα λύνουμε το $Ux=y$ με οπισθοδρόμηση.

Δύο στάδια: τριγωνοποίηση, οπισθοδρόμηση

Γενική περιγραφή: $Ax=b$, A αντιστρέψιμος

Τριγωνοποίηση
1ο Βήμα: Υπόθεση $a_{11}^{(1)} \neq 0$ (μπορεί να επιτευχθεί)
Πολλαπλασιαστές: Πολλαπλασιάζουμε την 1η εξίσωση επί m_{i1}, αφαιρούμε το αποτέλεσμα από την i-οστή γραμμή και αντικαθιστούμε την i-οστή εξίσωση με τη νέα που προκύπτει με αυτόν τον τρόπο για $i=2,\dots,n$. Έτσι μετά το πρώτο βήμα παίρνουμε το σύστημα $A^{(2)}x=b^{(2)}$ με:

$$A^{(2)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ \vdots & \vdots & & \vdots \\ 0 & a_{n2}^{(2)} & \cdots & a_{nn}^{(2)} \end{pmatrix}, \quad b^{(2)} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_n^{(2)} \end{pmatrix}$$

με $a_{ij}^{(2)} = a_{ij}^{(1)} \cdot m_{i1} a_{1j}^{(1)}$, $i,j=2,\dots,n$
$b_i^{(2)} = b_i^{(1)} - m_{i1} b_1^{(1)}$, $i=2,\dots,n$

Βήμα r: Ξεκινάμε από το σύστημα $A^{(r)}x = b^{(r)}$

$$A^{(r)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & & a_{2n}^{(2)} \\ \vdots & 0 & & & \\ \vdots & \vdots & a_{rr}^{(r)} & \cdots & a_{rn}^{(r)} \\ \vdots & \vdots & \vdots & & \vdots \\ 0 & 0 & a_{nr}^{(r)} & \cdots & a_{nn}^{(r)} \end{pmatrix}, \quad b^{(r)} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(2)} \\ \vdots \\ b_r^{(r)} \\ b_{r1}^{(r)} \\ \vdots \\ b_n^{(r)} \end{pmatrix}$$

$\tilde{A}^{(r)} = (a_{ij}^{(r)})_{ij=r,\dots,n}$

Έχουμε $\det A = \det A^{(r)} = a_{11} \cdot a_{22} \cdots a_{k-1,r-1}^{(r-1)} \det \tilde{A}^{(r)} \Rightarrow$

αν δεν έχουν γίνει εναλλαγές γραμμών

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

Πολλαπλασιαστές (υπόθεση: $a_{rr}^{(r)} \neq 0$)

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

Πολλαπλασιάζουμε την r-οστή εξίσωση επί m_{ir}, αφαιρούμε από την i-οστή, και αντικαθιστούμε την i-οστή με το αποτέλεσμα για $i = r+1, \dots, n$

Ύστερα από n-1 βήματα, το σύστημα γίνεται:

(*) $A^{(n)} \cdot x = b^{(n)}$ με άνω τριγωνικό πίνακα $A^{(n)}$

Εδώ τελειώνει το στάδιο της τριγωνοποίησης

Λύνουμε το (*) με οπισθοδρόμηση

Βήμα r της τριγωνοποίησης

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

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

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

"Απαιτούμενες πράξεις και μνήμη"

Μετράμε μόνο πολλαπλασιασμούς και διαιρέσεις!

• Τριγωνοποίηση

1ο Βήμα: • για τον υπολογισμό των πολλαπλασιαστών m_{i1}, $i=2,\dots,n$ απαιτούνται $n-1$ πράξεις.

• για τον υπολογισμό των $a_{ij}^{(2)}$, $i,j=2,\dots,n$ απαιτούνται $(n-1)^2$ πράξεις

Βήμα r: • $n-r$ πράξεις για τους m_{ir}; $i=r+1,\dots,n$

• $(n-r)^2$ πράξεις για τα $a_{ij}^{(r+1)}$, $i,j=r+1,\dots,n$

Συνολικά για τον A:

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

$$\left(\sum_{i=1}^{k} i^2=\frac{k(k+1)(2k+1)}{6}\right)$$

Για το διάνυσμα b:

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

• Οπισθοδρόμηση

$\frac{n^2+n}{2}$ πράξεις

Συνολικό κόστος:

$$\frac{n^3+3n^2-n}{3}=\frac{n^3}{3}+O(n^2)$$

Παράδειγμα $n=20$

Υπόθεση: Ο υπολογιστής κάνει 10^6 πράξεις/sec

Απαλοιφή Gauss: $\frac{16}{9}\cdot 10^{-3}$ sec

Cramer: Για τους $21 \cdot 20! \cdot 19$ πολλαπλασιασμούς απαιτούνται $\approx 3 \cdot 10^5$ αιώνες

"Απαιτούμενη μνήμη"

Για τον A: n^2 θέσεις μνήμης
Για τον B: n θέσεις μνήμης

Δεν απαιτείται επί πλέον μνήμη. Οι πολλαπλασιαστές m_{ir}, $i=r+1, \dots, n$, αποθηκεύονται στις θέσεις των στοιχείων a_{ir}, $i=r+1, \dots, n$. Οι πολλαπλασιαστές, δηλαδή, αποθηκεύονται στις θέσεις (i,j) με $i>j$ του πίνακα. Τα νέα στοιχεία υπολογίζονται $a_{ij} \leftarrow a_{ij} - m_{ir} \cdot a_{rj}$, $i,j = r+1, \dots, n$.
Τα x_i αποθηκεύονται στις θέσεις των b_i.

"Υπολογισμός ορίζουσας"

$$\det A = (-1)^m \det A^{(n)} = (-1)^m a_{11}^{(1)} \cdot a_{22}^{(2)} \cdot \ldots \cdot a_{nn}^{(n)}$$

m = πλήθος των εναλλαγών γραμμών.

απαιτούμενες πράξεις: $\frac{n^3}{3} + O(n)$ πράξεις

κανόνας Cramer: $(n+1)\left(\frac{n^3}{3} + O(n)\right) = \frac{n^4}{3} + O(n^3)$ πράξεις

"Οδήγηση"

Ορισμός: Τα στοιχεία $a_{ii}^{(i)}$, $i=1, \dots, n$ λέγονται οδηγοί. Αναμένονται προβλήματα ευστάθειας, αν κάποιος οδηγός έχει μικρή απόλυτη τιμή.

Παράδειγμα

$$\left.\begin{aligned} 10^{-4}x_1 + x_2 &= 1 \\ x_1 + x_2 &= 2 \end{aligned}\right\} \Rightarrow \left.\begin{aligned} x_1 &= 1.0001 \\ x_2 &= 0.9998 \end{aligned}\right\}$$

$b=10$, $t=3$, $L=-20$, $u=20$ στρογγύλευση.

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

$$a_{22}^{(2)} = fl(a_{22}^{(1)} - m_{21}a_{12}^{(1)}) = fl(1 - 10^4) = fl(-999) = -0.100 \cdot 10^5 = -10^4$$

$$b_2^{(2)} = fl(b_2^{(1)} - m_{21}b_1^{(1)}) = fl(2 - 10^4) = -10^4$$

$$\left.\begin{aligned} 10^{-4}x_1 + x_2 &= 1 \\ -10^4 x_2 &= -10^4 \end{aligned}\right\} \Rightarrow \left.\begin{aligned} x_1 &= 0 \\ x_2 &= 1 \end{aligned}\right) \rightarrow \text{πολύ κακή προσέγγιση.}$$

Με εναλλαγή γραμμών γράφουμε το αρχικό σύστημα στη μορφή:

$$\left.\begin{aligned} x_1 + x_2 &= 2 \\ 10^{-4}x_1 + x_2 &= 1 \end{aligned}\right\} \text{υπολογιστής μιας } \left.\begin{aligned} x_1 &= 1 \\ x_2 &= 1 \end{aligned}\right\} \text{πολύ καλή προσέγγιση}$$

Συμπέρασμα: Επιδιώκουμε να έχουμε οδηγούς με "μεγάλη" απόλυτη τιμή.

1η στρατηγική: Μερική οδήγηση ή οδήγηση κατά γραμμές.

Στο βήμα r της τριγωνοποίησης, με εναλλαγή της r-οστής γραμμής με κάποια από τις i-οστές γραμμές, $i=r+1,\dots,n$ φέρνουμε στη θέση του οδηγού ένα στοιχείο με τη μεγαλύτερη δυνατή απόλυτη τιμή.

Επί πλέον κόστος $\frac{n^2}{2} + O(n)$

26/3/2015

Παρατήρηση
Ο όρος "πολύ μικρός οδηγός" είναι ασαφής.
Γράφουμε το στη μορφή

$$\left.\begin{aligned} x_1 + 10^4 x_2 &= 10^4 \\ x_1 + x_2 &= 2 \end{aligned}\right\} \quad (**)$$

Ο "υπολογιστής μας" δίνει για το (**) την προσεγγιστική λύση $\left.\begin{aligned} x_1&=0 \\ x_2&=1 \end{aligned}\right\}$ ενώ αν κάνουμε πρώτα εναλλαγή γραμμών δίνει $\left.\begin{aligned} x_1&=1 \\ x_2&=1 \end{aligned}\right\}$

Ο πραγματικός λόγος της αποτυχίας στην πρώτη περίπτωση είναι ότι το μέγεθος του $|a_{11}|$ είναι πολύ μικρό σε σύγκριση με το $|a_{12}|$. Πράγματι, έχουμε $\left|\frac{a_{11}}{a_{12}}\right| = 10^{-4}$ ενώ $\left|\frac{a_{21}}{a_{22}}\right| = 1$, που μας υποδεικνύει να χρησιμοποιήσουμε το a_{21} ως οδηγό (δηλαδή να κάνουμε εναλλαγή γραμμών)

2η στρατηγική Ολική οδήγηση ή οδήγηση κατά γραμμές + στήλες
Στο i-οστό βήμα της τριγωνοποίησης με εναλλαγές γραμμών και στηλών φέρνουμε στη θέση του οδηγού $a_{ii}^{(i)}$ ένα από τα στοιχεία του υποπίνακα $a_{k\ell}^{(i)}$ $k,\ell = i, \dots, n$ με τη μέγιστη απόλυτη τιμή.

ΠΡΟΣΟΧΗ: Η εναλλαγή δύο στηλών r και s απαιτεί τη μετονομασία των αντίστοιχων αγνώστων x_r σε x_s και x_s σε x_r

Επί πλέον κόστος: $\frac{n^3}{3} + O(n^2)$

υμπεράσματα

Η μέθοδος απαλοιφής του Gauss χωρίς οδήγηση θεωρείται σταθής αλγόριθμος. Για κάποιες κατηγορίες συστημάτων, π.χ. ο πίνακας A είναι θετικά ορισμένος, δηλαδή $\forall x \in \mathbb{R}^n, x \neq 0$ $^TAx > 0$, είναι ευσταθής και μόνο σε τέτοιες περιπτώσεις ρησιμοποιείται η μέθοδος στην πράξη.

Η μέθοδος ολικής οδήγησης θεωρείται ευσταθής στην πράξη. ρησιμοποιείται σπάνια γιατί διπλασιάζει το κόστος.

Η πιθανότητα να είναι ασταθής η μέθοδος απαλοιφής του auss με μερική οδήγηση είναι πολύ μικρή. Αυτή η παραλλαγή ησιμοποιείται πολύ συχνά στην πράξη γιατί αυξάνει μόνο λίγο ο κόστος.

ο αλγόριθμος της απαλοιφής στην πράξη"

$x=b$

στάδιο (1η φάση)

τελούμε τους υπολογισμούς της τριγωνοποίησης που αφορούν μόνο πίνακα A. (DECOMP)

στάδιο

ησιμοποιούμε το αποτέλεσμα του 1ου σταδίου και το b, και ολογίζουμε το x. (SOLVE)

στος 1ου σταδίου $\frac{n^3}{3}$

στος 2ου σταδίου n^2

$x=b$

$x'=b'$

$x''=b''$

λούμε μια φορά τη DECOMP και με αλλεπάλληλες κλήσεις SOLVE υπολογίζουμε τα x, x', x''.

Αν έχουμε m συστήματα:
κόστος $\frac{n^3}{3} + m \cdot n^2 + \dots$

Εφαρμογή
Έστω ότι θέλουμε να υπολογίσουμε τον A^{-1}. Έχουμε $A^{-1} = (u^1, u^2, \dots, u^n)$ με $Au^i = e^i$, $i = 1, \dots, n$
συνολικό κόστος: $n^3 + O(n^2)$

"Ανάλυση LU"

Η φάση της τριγωνοποίησης του πίνακα A κατά την απαλοιφή Gauss (πιθανόν με εναλλαγές γραμμών) μπορεί να ερμηνευθεί ως ανάλυση του A σε γινόμενο $A = P^{-1}LU$ όπου:
P: πίνακας μετάθεσης που καταγράφει τις εναλλαγές γραμμών. Ένας πίνακας μετάθεσης προκύπτει από το μοναδιαίο πίνακα με κατάλληλες εναλλαγές γραμμών.
L: κάτω τριγωνικός πίνακας με μονάδες στη διαγώνιο. Κάτω από τη διαγώνιο περιέχει τους πολλαπλασιαστές m_{ij}, $i > j$
U: άνω τριγωνικός πίνακας, το τελικό προϊόν $A^{(n)}$ της τριγωνοποίησης

Συνέπειες της ανάλυσης

$Ax = b \Rightarrow PAx = b \Rightarrow LUx = Pb$
1ο στάδιο $A \rightsquigarrow P, L, U$
$A = P^{-1}LU$
(απαιτεί ακριβώς τις πράξεις της τριγωνοποίησης του A)
2ο στάδιο
- Pb (με κατάλληλη μετάθεση των συνιστωσών του b)
- $Ly = Pb \rightsquigarrow y$. Βρίσκουμε αναδρομικά τα $y_1, y_2, \dots, y_n$

• $Ux = y \rightsquigarrow$ οπισθοδρόμηση $\rightsquigarrow x$

1η περίπτωση Υποθέτουμε ότι κατά την απαλοιφή του Gauss δεν γίνονται εναλλαγές γραμμών. (Τότε $P = I_n$)
Με τον πίνακα:

$$M_1 = \begin{pmatrix} 1 & & & & 0 \\ -m_{21} & \ddots & & & \\ -m_{31} & & \ddots & & \\ \vdots & & & \ddots & \\ -m_{n1} & & & & 1 \end{pmatrix}$$ έχουμε $A^{(2)} = M_1 A$.

Αντίστοιχα έχουμε $A^{(n)} = M_{n-1} M_{n-2} \cdots M_1 A$ (*), όπου M_r είναι ο κάτω τριγωνικός πίνακας με

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

Θέτουμε $U = A^{(n)}$. Οι πίνακες M_r έχουν ορίζουσα ίση με 1, ιδιαίτερα αν είναι αντιστρέψιμος οπότε η (*) δίνει:
$A = M_1^{-1} \cdots M_{n-2}^{-1} M_{n-1}^{-1} U$ (**)
Τώρα (M_r^{-1}).

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

$$M_1^{-1} \cdots M_{n-1}^{-1} = \begin{pmatrix} 1 & & & & 0 \\ m_{21} & 1 & & & \\ m_{31} & m_{32} & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ m_{n1} & m_{n2} & \cdots & & m_{n,n-1} \end{pmatrix} = L$$

Δηλαδή έχουμε $A = LU$

2η περίπτωση Έστω ότι κατά την απαλοιφή Gauss γίνονται εναλλαγές γραμμών (είτε για να βρούμε οδηγούς $a_{ii}^{(i)} \neq 0$ είτε λόγω μερικής οδήγησης)

Η γραμμή που στο βήμα i φέρναμε τη θέση του οδηγού (δηλαδή αυτή που εναλλάσσεται με τη γραμμή i) ούτε αλλάζει πλέον θέση ούτε τα στοιχεία αλλάζουν. Επομένως, υπάρχει μια μετάθεση των γραμμών του πίνακα A που, αν την κάναμε πριν από το πρώτο βήμα της απαλοιφής θα οδηγούσε σε ένα πίνακα A', η τριγωνοποίηση του οποίου θα μπορούσε να γίνει χωρίς εναλλαγές γραμμών. Άρα, σύμφωνα με την πρώτη περίπτωση $A' = LU$ με L και U πίνακες με τις επιθυμητές ιδιότητες. Ο πίνακας ~~με τη~~ μετάθεσης P που αντιστοιχεί στη μετάθεση $\pi = \begin{pmatrix} 1 & 2 & \cdots & n \\ i_1 & i_2 & \cdots & i_n \end{pmatrix}$ προκύπτει απ' τον I_n αν μεταθέσαμε τις γραμμές του κατά τη μετάθεση π, δηλαδή η i_k γραμμή του P είναι η k γραμμή του I_n.

Παράδειγμα

Στη μετάθεση $\begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \end{pmatrix}$ αντιστοιχεί ο πίνακας:

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

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

$$\begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \begin{pmatrix} a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{11} & a_{12} & a_{13} \end{pmatrix}$$

Επομένως, έχουμε $A' = PA$ με κατάλληλο πίνακα μετάθεσης P.

Άρα, $PA = LU \rightarrow A = P^{-1}U$

Παραδείγματα

(α) $A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$

Απαλοιφή Gauss

$m_{21} = \frac{1}{2}$

Άρα $L = \begin{pmatrix} 1 & 0 \\ 1/2 & 1 \end{pmatrix}$ $\quad U = \begin{pmatrix} 2 & 3 \\ 0 & 4-3/2 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 0 & 5/2 \end{pmatrix}$

$A = LU$ χωρίς εναλλαγές γραμμών

(β) $A = \begin{pmatrix} 2 & 4 & 5 \\ 1 & 2 & 3 \\ 2 & 5 & 9 \end{pmatrix}$

Απαλοιφή Gauss

$m_{21} = \frac{1}{2}$, $\quad m_{31} = 1$ $\qquad A^{(2)} = \begin{pmatrix} 2 & 4 & 5 \\ 0 & 0 & 1/2 \\ 0 & 1 & 4 \end{pmatrix}$

$L = \begin{pmatrix} 1 & & 0 \\ 1/2 & 1 & \\ 1 & 2 & 1 \end{pmatrix}$

Για να προχωρήσουμε στο $2^ο$ βήμα της τριγωνοποίησης πρέπει πρώτα να εναλλάξουμε τη $2^η$ με την $3^η$ γραμμή.

$P = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$ $\qquad$ έχουμε $PA = \begin{pmatrix} 2 & 4 & 5 \\ 2 & 5 & 9 \\ 1 & 2 & 3 \end{pmatrix} =: B$

Απαλοιφή Gauss στον B:

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

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

Άρα:

$$L = \begin{pmatrix} 1 & & 0 \\ 1 & 1 & \\ 1/2 & 0 & 1 \end{pmatrix}$$

$A = P^{-1} L \cdot U \Leftrightarrow PA = LU$ (επαλήθευση)

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

31/3/2015

"Κατάσταση γραμμικών συστημάτων"

$Ax = b$

Παράδειγμα

$$\begin{pmatrix} .913 & .659 \\ .0780 & .563 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} .254 \\ .217 \end{pmatrix}$$ Λύση: $x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Απαλοιφή Gauss (χωρίς, με μερική ή με ολική οδήγηση, οι τρεις περιπτώσεις εδώ συμπίπτουν γιατί εδώ το στοιχείο a_{11} έχει τη μεγαλύτερη απόλυτη τιμή, του A)

$\beta = 10$, $t = 3$, αποκοπή

$\tilde{x} = \begin{pmatrix} -.443 \\ 1 \end{pmatrix}$ πολύ κακή προσέγγιση

Πού οφείλεται το γεγονός ότι παίρνουμε τόσο άσχημο αποτέλεσμα, στη μέθοδο ή στο πρόβλημα.

Απάντηση:

$A \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} .253 \\ .218 \end{pmatrix} \rightsquigarrow y = \begin{pmatrix} 1223 \\ -1694 \end{pmatrix}$

μεταβολή της τάξης 10^{-3}

Συμπέρασμα: Το πρόβλημα έχει πολύ κακή κατάσταση.

$\det A = -10^{-6}$ Είναι αυτός ο λόγος που το σύστημα έχει κακή κατάσταση;

Απάντηση: ΟΧΙ

Γιατί;

$\det(\lambda A) = \lambda^{2} A$ → έχω 2×2 πίνακα

Για $\lambda = 10^4$ $\det(10^4 A) = 10^8(-10^{-6}) = -10^2 = -100$

Πολλαπλασιάζοντας το αρχικό σύστημα επί λ η κατάσταση μένει ίδια, ενώ η ορίζουσα πολλαπλασιάζεται με λ^2

Για να μελετήσουμε την κατάσταση γραμμικών συστημάτων χρειαζόμαστε έναν τρόπο να μετράμε μεταβολές διανυσμάτων και μεταβολές πινάκων.

Ένας βολικός τρόπος να γίνει αυτό είναι οι λεγόμενες νόρμες διανυσμάτων και πινάκων που αποτελούν γενίκευση της έννοιας της απόλυτης τιμής.

"Νόρμες διανυσμάτων κ' πινάκων"

1. Νόρμες διανυσμάτων

Ορισμός: Έστω X ένας γραμμικός (=διανυσματικός) χώρος στο $\mathbb{R}$ ή στο $\mathbb{C}$ και $\mathbb{K} = \mathbb{R}$ ή $\mathbb{K} = \mathbb{C}$. Μια απεικόνιση

$\|\ \| : X \to \mathbb{R}$, $x \mapsto \|x\|$

λέγεται νόρμα (νόρμα = στάθμη) αν ισχύουν:

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

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

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

Παρατηρήσεις:

1) <u>Ισχυρισμός:</u> $\forall x \in X \quad \|x\| \geq 0$

$0 = \|x - x\| = \|x + (-x)\|$

$\underset{N3}{\leq} \|x\| + \|-x\| \underset{N2}{=} \|x\| + \|x\| = 2\|x\|$

$\Rightarrow 2\|x\| \geq 0 \Rightarrow \|x\| \geq 0$

2) <u>Ισχυρισμός:</u> $\forall x, y \in X \quad \big| \|x\| - \|y\| \big| \leq \|x-y\|$

τριγωνική ανισότητα προς τα κάτω

$\|x\| = \|(x-y) + y\| \underset{N3}{\leq} \|x-y\| + \|y\| \Rightarrow \|x\| - \|y\| \leq \|x-y\|$

Εντελώς αντίστοιχα, $\|y\| - \|x\| \leq \|y-x\| = \|x-y\|$

Συνολικά, $\big|\|x\| - \|y\|\big| \leq \|x-y\|$

<u>Παραδείγματα</u>

① $(\mathbb{R}, \|\cdot\|)$ με $\|x\| = |x| \quad \forall x \in \mathbb{R}$.
$(\mathbb{C}, \|\cdot\|)$ με $\|z\| = |z| \quad \forall z \in \mathbb{C}$

② $(\mathbb{R}^n, \|\cdot\|_1)$ με $\|x\|_1 = |x_1| + |x_2| + \dots + |x_n| \qquad \sum_{i=1}^{n} |x_i|$

$x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n \qquad (\ell_1$ νόρμα$)$

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

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

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

(N3) $x, y \in \mathbb{R}^n$

$\|x+y\|_1 = \sum_{i=1}^{n} |(x+y)_i| = \sum_{i=1}^{n} |x_i + y_i| \qquad |x_i + y_i| \leq |x_i| + |y_i|$

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

③ $(\mathbb{R}^n, \|\cdot\|_\infty)$ με $\|x\|_\infty = \max\limits_{1\le i\le n}|x_i|$ (ℓ_∞ νόρμα ή νόρμα μεγίστου)

Η απόδειξη είναι απλή (ΣΠΙΤΙ)

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

(N1), (N2) πολύ απλές (ΣΠΙΤΙ)

(N3) Θεωρούμε το εσωτερικό γινόμενο $(\cdot,\cdot)_2 : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$,
$(x,y)_2 = \sum_{i=1}^{n} x_i \cdot y_i$

Τότε $\|x\|_2 = \sqrt{(x,x)_2}$ ○○○

Θυμάμαι:

Ισχυρισμός: $\forall x,y \in \mathbb{R}^n \quad |(x,y)_2| \le \|x\|_2 + \|y\|_2$ ανισότητα των Cauchy-Schwarz

$\left|\sum_{i=1}^{n} x_i y_i\right| \le \left(\sum_{i=1}^{n}|x_i|^2\right)^{1/2}\left(\sum_{i=1}^{n}|y_i|^2\right)^{1/2}$

απόδειξη:

Για $y=0$, τετριμμένο

Για $y\neq 0$ και για $\lambda\in\mathbb{R}$ έχουμε $0 = (x+\lambda y, x+\lambda y)_2 =$
$= (x,x)_2 + \lambda(x,y)_2 + \lambda(x,y)_2 + \lambda^2(y,y)_2 =$
$\lambda^2(y,y)_2 + 2\lambda(x,y)_2 + (x,x)_2 \geq 0$

Επομένως $\Delta \le 0$, δηλαδή $[2(x,y)_2]^2 - 4(y,y)_2(x,x)_2 \le 0 \Leftrightarrow$
$(x,y)_2^2 \le (y,y)_2(x,x)_2$
$\Rightarrow |(x,y)| \le \underbrace{(x,x)_2^{1/2}}_{\|x\|_2} \underbrace{(y,y)_2^{1/2}}_{\|y\|_2}$

○○○ $\|x+y\|_2^2 = (x+y, x+y)_2 = \|x\|_2^2 + 2(x,y)_2 + \|y\|_2^2 \le$
$\underset{C.S.}{\le} \|x\|_2^2 + 2\|x\|_2\|y\|_2 + \|y\|_2^2 = (\|x\|_2 + \|y\|_2)^2$

$\|x+y\|^2 \le (\|x\|_2 + \|y\|_2)^2 \Rightarrow \|x+y\|_2 \le \|x\|_2 + \|y\|_2$

(5) $(\mathbb{C}^n, \|\cdot\|_p)$ με $p = 1, \infty, 2$

$\|z\|_1 = \sum_{i=1}^{n} |z|$, $\|z\|_\infty = \max_{1\le i\le n} |z_i|$

$\|z\|_2 = \left(\sum_{i=1}^{n} |z_i|^2\right)^{1/2}$

Εσωτερικό γινόμενο

$(x,y)_2 = x_1\bar{y}_1 + x_2\bar{y}_2 + \dots + x_n\bar{y}_n = \sum_{i=1}^{n} x_i\bar{y}_i$

Τότε $\|x\|_2 = \sqrt{(x,x)_2}$

2/4/2015

(6) $-\infty < a < b < \infty$

$(C[a,b], \|\cdot\|_\infty)$ με $\|f\|_\infty = \max_{a\le x\le b} |f(x)|$ (νόρμα μεγίστου)

Λύση:

(N1) $f \in C[a,b]$

$\|f\|_\infty = 0 \iff \max_{a\le x\le b} |f(x)| = 0 \iff f(x) = 0 \quad \forall x \in [a,b]$

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

$\|\lambda f\|_\infty = \max_{a\le x\le b} |\lambda f(x)| = \max_{a\le x\le b} (|\lambda| |f(x)|) =$

$= \lambda \max_{a\le x\le b} |f(x)| = |\lambda| \cdot \|f\|_\infty$

(N3) $f, g \in C[a,b]$

$\|f+g\|_\infty = \max_{a\le x\le b} |f(x)+g(x)| \le |f(x^*) + g(x^*)|$ για κατάλληλο x^*

$\le |f(x^*)| + |g(x^*)| \le \max_{a\le x\le b} |f(x)| + \max_{a\le x\le b} |g(x)| =$

$= \|f\|_\infty + \|g\|_\infty$

Ορισμός: (Ισοδύναμες νόρμες)
Έστω X ένας γραμμικός χώρος. Δύο νόρμες $\|\ \|$ και $\|\ \|'$ στον X λέγονται ισοδύναμες (ή συγκρίσιμες), αν υπάρχουν θετικά m, M τ.ω. $\forall x \in X$ $m\|x\| \leq \|x\|' \leq M\|x\|$
(οπότε και $\frac{1}{M}\|x\|' \leq \|x\| \leq \frac{1}{m}\|x\|'$)

Λήμμα: (Ισοδυναμία νόρμας με τη νόρμα μεγίστου)
Κάθε νόρμα $\|\cdot\|$ στον $\mathbb{R}^n$ είναι ισοδύναμη με τη νόρμα μεγίστου.
απόδειξη:
Έστω $x = (x_1, \ldots, x_n)^T \in \mathbb{R}^n$ και $\{e^1, \ldots, e^n\}$ η κανονική βάση του $\mathbb{R}^n$.
Τότε έχουμε: $x = x_1 e^1 + x_2 \cdot e^2 + \ldots + x_n e^n$
οπότε
$$\|x\| = \|x_1 e^1 + x_2 e^2 + \ldots + x_n e^n\|$$
$$\leq \|x_1 e^1\| + \|x_2 e^2\| + \ldots + \|x_n e^n\| =$$
$$|x_1| e^1 + |x_2| e^2 + \ldots + |x_n| e^n$$
$$\leq \left(\max_{1\leq i \leq n} |x_i|\right) \underbrace{\left(\|e^1\| + \ldots + \|e^n\|\right)}_{=M}$$
M ανεξάρτητο του x

$\Rightarrow \|x\| \leq M\|x\|_\infty$

Πρόταση: (Ισοδυναμία νορμών στον $\mathbb{R}^n$)
Έστω $\|\cdot\|$ και $\|\cdot\|'$ νόρμες στον $\mathbb{R}^n$. Τότε αυτές είναι ισοδύναμες μεταξύ τους.
απόδειξη:
Σύμφωνα με το προηγούμενο Λήμμα τόσο η $\|\cdot\|$ όσο και η $\|\cdot\|'$ είναι ισοδύναμες με τη $\|\cdot\|_\infty$, δηλαδή υπάρχουν $m, M, m', M' > 0$ τ.ω. $\forall x \in \mathbb{R}^n$ $m\|x\|_\infty \overset{a}{\leq} \|x\| \overset{b}{\leq} M\|x\|_\infty$
και $\forall x \in \mathbb{R}^n$ $m'\|x\|_\infty \underset{c}{\leq} \|x\| \underset{d}{\leq} M'\|x\|_\infty$

Επομένως,
$\forall x \in \mathbb{R}^n$ $\|x\| \leq M\|x\|_\infty \underset{c}{\leq} \underbrace{M\frac{1}{m'}}_{=\tilde{M}}\|x\|'$
χρησιμοποιώ την ανισότητα b

και $\forall x \in \mathbb{R}^n \quad \|x\|' \overset{d}{\leq} M' \|x\|_\infty \overset{a}{\leq} M' \frac{1}{m} \|x\|$

ή $\frac{m}{M'} \|x\|' \leq \|x\|$, $\frac{m}{M'} = \tilde{m}$

Συνολικά: $\forall x \in \mathbb{R}^n \quad \tilde{m}\|x\|' \leq \|x\| \leq \tilde{M}\|x\|'$

Ορισμός: (Σύγκλιση ακολουθίας)
Λέμε ότι μια ακολουθία $(x^{(n)})_{n\in\mathbb{N}} \subset X$ συγκλίνει ως προς τη νόρμα $\|\cdot\|$ του X, αν υπάρχει $x \in X$ (το όριο της ακολουθίας) τ.ω. $\|x^{(n)} - x\| \to 0 \quad n \to \infty$
Γράφουμε: $x^{(n)} \to x, \ n \to \infty$, $x = \lim_{n\to\infty} x^{(n)}$

Αν έχουμε δύο νόρμες και η ακολουθία συγκλίνει προς τη μια νόρμα τότε συγκλίνει και προς την άλλη στο ίδιο όριο αφού $m\|x^{(n)} - x\|' \leq \|x^{(n)} - x\| \leq M\|x^{(n)} - x\|'$

Ιδιαίτερα στον $\mathbb{R}^n$ σύγκλιση μιας ακολουθίας $(x^{(n)})_{n\in\mathbb{N}}$ στο x σε μια οποιαδήποτε νόρμα, σημαίνει ότι $x_i^{(n)} \to x_i, \ n \to \infty$, για $i = 1, \ldots, n$

Ορισμός (πλήρης χώρος)
Ένας γραμμικός χώρος $(X, \|\cdot\|)$ λέγεται πλήρης, αν κάθε ακολουθία Cauchy (ως προς την $\|\cdot\|$) στον X συγκλίνει (ως προς $\|\cdot\|$), δηλαδή αν για μια ακολουθία $(x^{(n)})_{n\in\mathbb{N}} \subset X$ ισχύει $\forall \varepsilon > 0 \ \exists N \in \mathbb{N}$
$(\forall n, m \geq N \quad \|x^{(n)} - x^{(m)}\| < \varepsilon) \Rightarrow$
$\exists x \in X \quad \|x^{(n)} - x\| \to 0, \quad n \to \infty$

Πρόταση
Ο $(\mathbb{R}^n, \|\cdot\|_\infty)$ πλήρης

Θεώρημα
Αν $\|\cdot\|$ νόρμα στον $\mathbb{R}^n$, τότε ο $(\mathbb{R}^n, \|\cdot\|)$ είναι πλήρης.

2. Νόρμες πινάκων
Ορισμός: Μια απεικόνιση $\|\cdot\|: \mathbb{R}^{n,n} \to \mathbb{R}$, $A \to \|A\|$, $\mathbb{C}^{n,n}$
λέγεται νόρμα πινάκων αν ισχύουν:
(N1)
(N2)
(N3)
(N4) $\forall\, A, B \in \mathbb{R}^{n,n}$ $\|AB\| \leq \|A\| \cdot \|B\|$

"Ειδικές νόρμες" (φυσικές νόρμες θα ασχοληθούμε μόνο μ' αυτές)

Προεργασία για τον ορισμό.
Έστω $\|\cdot\|$ νόρμα στον $\mathbb{R}^n$ και $A \in \mathbb{R}^{n,n}$. Επειδή η $\|\cdot\|$ είναι ισοδύναμη με τη $\|\cdot\|_\infty$ υπάρχουν $m, M > 0$ τω:
$\forall x \in \mathbb{R}^n$ $m\|x\|_\infty \leq \|x\| \leq M\|x\|_\infty$
Επομένως, για κάθε $x \in \mathbb{R}^n$, $x \neq 0$, έχουμε:

$$\frac{\|Ax\|}{\|x\|} \leq \frac{M\|Ax\|_\infty}{m\|x\|_\infty} = \frac{M \max_{1\leq i\leq n} |(Ax)_i|}{\|x\|_\infty} = \frac{M}{m} \frac{\max_{1\leq i\leq n} \left|\sum_{j=1}^n a_{ij} x_j\right|}{\|x\|_\infty}$$

$$\leq \frac{M}{m} \frac{\max_{1\leq i\leq n} \left(\sum_{j=1}^n |a_{ij}| \cdot |x_j|\right)}{\|x\|_\infty} \leq \frac{M}{m} \frac{\max_{1\leq i\leq n} \left(\sum_{j=1}^n |a_{ij}| \cdot \|x\|_\infty\right)}{\|x\|_\infty} =$$

$$\frac{M}{m} \max_{1\leq i\leq n} \left(\sum_{j=1}^n |a_{ij}|\right) = \underset{\text{ανεξ. του } x}{C} < \infty$$

Ορισμός: (φυσική νόρμα)
Έστω $\|\cdot\|$, μια νόρμα στον $\mathbb{R}^n$. Η απεικόνιση:
$\|\cdot\|: \mathbb{R}^{n,n} \longrightarrow \mathbb{R}$, $\|A\| = \overset{\text{ή max}}{\sup_{\substack{x\in\mathbb{R}^n \\ x\neq 0}}} \frac{\|Ax\|}{\|x\|}$ λέγεται ~~φυσική~~

φυσική νόρμα πινάκων ή νόρμα πινάκων παραγόμενη (επαγόμενη) από τη νόρμα $\|\cdot\|$

Παρατηρήσεις:

1. Το sup μπορεί να αντικατασταθεί με max στον αριθμό.
2. $\forall x \in \mathbb{R}^n \quad \|Ax\| \leq \|A\| \cdot \|x\|$

Σημείωση (υπολογισμός φυσικής νόρμας)

- Για τις νόρμες $\|\cdot\|_p$, $p = 1, \infty, 2$ υπάρχουν τύποι.
- Γενικά: Αν $c_1, c_2 > 0$ τ.ω. $\forall x \in \mathbb{R}^n$ ① $\|Ax\| \leq c_1 \|x\| \rightsquigarrow \|A\| \leq c_1$

$\exists \; y \in \mathbb{R}^n$ ② $\|Ay\| \geq c_2 \|y\| \rightsquigarrow \|A\| \geq c_2$

Βήμα 1ο: Αποδεικνύαμε την ① με όσο μικρότερη σταθερά c μπορούμε.

Βήμα 2ο: Επιλέγουμε $y \in \mathbb{R}^n$, $y \neq 0$, τ.ω. στις ανισότητες που μας οδήγησαν στην ① να ισχύει παντού ισότητα.

Παραδείγματα

ⓐ $\|\cdot\|_\infty$ στον $\mathbb{R}^n$. Συμβολίζουμε με $\|\cdot\|_\infty$ την παραγόμενη φυσική νόρμα πινάκων. Τότε ισχύει:

$\|A\|_\infty = \max\limits_{1 \leq i \leq n} \sum\limits_{j=1} |a_{ij}|$ (αθροίζω τις γραμμές)

ⓑ $\|\cdot\|_1$

$\|A\|_1 = \max\limits_{1 \leq j \leq n} \sum\limits_{i=1}^{n} |a_{ij}|$ (αθροίζω τις στήλες) *

ⓒ $\|\cdot\|_2$

$\|A\|_2 = \left[\rho(A^T A)\right]^{1/2}$

Για $A \in \mathbb{C}^{n,n}$

ίδια περίπτωση

* $\|A\|_2 = \left[\rho(A^* A)\right]^{1/2}$ με $A^*_{ij} = \bar{a}_{ij}$

ίδια περίπτωση

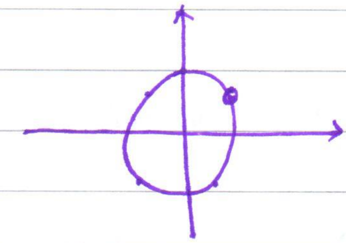
Θυμάμαι:

Φασματική ακτίνα πινάκων

$B \in \mathbb{R}^{n,n}$ $\quad \lambda_1, \lambda_2, \dots, \lambda_n \in \mathbb{C}$

ιδιοτιμές του B

$\rho(B) := \max\limits_{1 \le i \le n} |\lambda_i|$



"Δείκτης (ή αριθμός) κατάστασης συστημάτων"

$A \in \mathbb{R}^{n,n}$, $b \in \mathbb{R}^n$

αντιστρέψιμος

$Ax = b$

$\Delta b \in \mathbb{R}^n$

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

Έχουμε: $A(x+\Delta x) = b + \Delta b \iff \overset{=b}{\cancel{Ax}} + \Delta x = \cancel{b} + \Delta b \iff$

$A\Delta x = \Delta b \iff \Delta x = A^{-1}\Delta b$

Έστω $\|\cdot\|$ νόρμα στο $\mathbb{R}^n$ και η αντίστοιχη νόρμα πινάκων. Τότε

$\|\Delta x\| = \|A^{-1}\Delta b\| \le \|A^{-1}\| \cdot \|\Delta b\|$

Άρα: $\|\Delta x\| \le \|A^{-1}\| \cdot \|\Delta b\|$

Για $b \neq 0$ έχουμε: $\dfrac{\|\Delta x\|}{\|x\|} \le \dfrac{\|A^{-1}\| \cdot \|\Delta b\|}{\|x\|} \le \dfrac{\|A^{-1}\| \cdot \|\Delta b\|}{\frac{\|b\|}{\|a\|}} = \underbrace{\|A\| \cdot \|A^{-1}\|}_{\kappa(A)} \dfrac{\|\Delta b\|}{\|b\|}$

$Ax = b \Rightarrow \|b\| = \|Ax\| \le \|A\| \cdot \|x\| \Rightarrow \|x\| \ge \dfrac{\|b\|}{\|A\|}$

$\kappa(A)$: αριθμός κατάστασης του A.

$\dfrac{\|\Delta x\|}{\|x\|} \le \kappa(A) \dfrac{\|\Delta b\|}{\|b\|}$ $\quad (\kappa(A)$ Turing$)$

21/4/2015

Έστω $\|\cdot\|$ μια νόρμα στον $\mathbb{R}^n$ και η αντίστοιχη νόρμα πινάκων.

Έστω $A \in \mathbb{R}^{n,n}$ αντιστρέψιμος και $b \in \mathbb{R}^n$, $b \neq 0$. Τότε:

$$\left.\begin{array}{l} Ax = b \\ A(x+\Delta x) = b + \Delta b \end{array}\right\} \Rightarrow \frac{\|\Delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\Delta b\|}{\|b\|} \quad (*)$$

$\kappa(A) = \|A\| \cdot \|A^{-1}\|$ δείκτης κατάστασης του A.

<u>Ισχυρισμός:</u> $\kappa(A) \geq 1$

$$\kappa(A) = \|A\| \cdot \|A^{-1}\| \geq \|\underbrace{AA^{-1}}_{= I_n}\| = \|I_n\| = 1$$

$(\|AB\| \leq \|A\| \|B\|)$

<u>Θεώρημα</u> (Εκτίμηση της σχετικής μεταβολής λύσεων γραμμικών συστημάτων)

Έστω $\|\cdot\|$ νόρμα στον $\mathbb{R}^n$ και η παραγόμενη από αυτή φυσική νόρμα στον $\mathbb{R}^{n,n}$.

Έστω $A \in \mathbb{R}^{n,n}$ αντιστρέψιμος, $\Delta A \in \mathbb{R}^{n,n}$ και b, $\Delta b \in \mathbb{R}^n$, $b \neq 0$.

Τότε αν $\kappa(A) = \|A\| \cdot \|A^{-1}\|$ έχουμε:

(α) $$\left.\begin{array}{l} Ax = b \\ A(x+\Delta x) = b + \Delta b \end{array}\right\} \Rightarrow \frac{\|\Delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\Delta b\|}{\|b\|}$$

(β) $$\left.\begin{array}{l} Ax = b \\ (A+\Delta A)(x+\Delta x) = b \\ \|A^{-1}\| \cdot \|\Delta A\| < 1 \end{array}\right\} \Rightarrow \begin{array}{l} A+\Delta A \text{ αντιστρέψιμος} \\ \dfrac{\|\Delta x\|}{\|x\|} \leq \dfrac{\kappa(A)}{1 - \|A^{-1}\| \cdot \|\Delta A\|} \cdot \dfrac{\|\Delta A\|}{\|A\|} \end{array}$$

(γ) $$\left.\begin{array}{l} Ax = b \\ (A+\Delta A)(x+\Delta x) = b + \Delta b \\ \|A^{-1}\| \cdot \|\Delta A\| < 1 \end{array}\right\} \Rightarrow \frac{\|\Delta x\|}{\|x\|} \leq \frac{\kappa(A)}{1 - \|A^{-1}\| \cdot \|\Delta A\|} \left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|} \right)$$

Παρατηρήσεις:

1. Οι εκτιμήσεις αυτές είναι οι καλύτερες δυνατές με την έννοια ότι μπορεί να ισχύουν ως ισότητες για $\Delta x \neq 0$.

2 Τα (α) και (β) είναι ειδικές περιπτώσεις του (γ) για $\Delta A = 0$ και $\Delta b = 0$, αντίστοιχα.

απόδειξη:

(α) Είναι η (*)

(β) Βοηθητικό αποτέλεσμα:

Αν $B \in \mathbb{R}^{n,n}$ τότε: $(\exists c>0 \ \forall x \in \mathbb{R}^n \ \|Bx\| \geq c\|x\|) \Rightarrow$ ο B είναι αντιστρέψιμος και $\|B^{-1}\| \leq \frac{1}{c}$

• $Bx = 0 \Rightarrow c\|x\| \leq 0 \Rightarrow \|x\| = 0 \Rightarrow x = 0$

Άρα ο B είναι αντιστρέψιμος

• Για $y \in \mathbb{R}^n$ ισχύει: $c\|B^{-1}y\| \leq B \cdot B^{-1}y = \|I_n y\| = \|y\|$

Άρα $\forall y \in \mathbb{R}^n$ $\|B^{-1}y\| \leq \frac{1}{c}\|y\|$ οπότε $\|B^{-1}\| \leq \frac{1}{c}$

Για $y \in \mathbb{R}^n$ έχουμε:

$$\|(A+\Delta A)y\| \geq \|Ay\| - \|\Delta A \cdot y\| \leq \|\Delta A\| \cdot \|y\|$$

$$\geq \|Ay\| - \|\Delta A\| \cdot \|y\|$$

$$\geq \frac{y}{\|A^{-1}\|} - \|\Delta A\| \cdot \|y\|$$

$$\geq \frac{\|y\|}{\|A^{-1}\|} - \|\Delta A\| \cdot \|y\|$$

$$= \frac{1 - \|A^{-1}\| \cdot \|\Delta A\|}{\|A^{-1}\|} \|y\|$$

$= c > 0$

$y = A^{-1}Ay \Rightarrow \|y\| \leq \|A^{-1}\| \cdot \|Ay\| \Rightarrow \|Ay\| \geq \frac{\|y\|}{\|A^{-1}\|}$

Σύμφωνα με το βοηθητικό αποτέλεσμα

• $A + \Delta A$ αντιστρέψιμος

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

(γ)

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

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

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

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

$$\leq \frac{\underbrace{\|A^{-1}\| \, \|A\|}_{=k(A)}}{1-\|A^{-1}\| \cdot \|\Delta A\|} \cdot \frac{\|\Delta A\|}{\|A\|}$$

$\oplus$

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

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

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

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

$b = Ax \Rightarrow \|b\| \leq \|A\| \cdot \|x\|$

"Επαναληπτικές μέθοδοι"

$A \in \mathbb{R}^{n,n}$ αντιστρέψιμος

$b \in \mathbb{R}^n$

$Ax = b$

Για $x^{(0)}$ αρχική προσέγγιση του x οι μέθοδοι αυτές δίνουν μια ακολουθία προσεγγίσεων $x^{(m)}$ $m = 1, 2, \ldots$

Υπό κατάλληλες προϋποθέσεις αυτή η ακολουθία συγκλίνει στο x.

"Ειδικές μέθοδοι"

- Μέθοδος του Jacobi
- Μέθοδος των Gauss-Seidel

Επιπρόσθετη υπόθεση: $a_{ii} \neq 0$, $i=1,\dots,n$

$$Ax=b \Leftrightarrow (Ax)_i = b_i,\ i=1,\dots,n \Leftrightarrow$$

$$\sum_{j=1}^{n} a_{ij}x_j = b_i,\ i=1,\dots,n \Leftrightarrow$$

$$a_{ii}x_i = b_i - \sum_{j=1}^{i-1} a_{ij}x_j - \sum_{j=i+1}^{n} a_{ij}x_j \Leftrightarrow$$

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

- Μέθοδος του Jacobi

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

- Μέθοδος των Gauss-Seidel

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

Παρατηρήσεις:

1. Υπάρχει μια κατηγορία πινάκων για την οποία η μέθοδος των Gauss-Siedel συγκλίνει πολύ ταχύτερα από την μέθοδο Jacobi.
2. Υπάρχουν όμως παραδείγματα πινάκων για τα οποία η μια μέθοδος να συγκλίνει και η άλλη να μην συγκλίνει.
3. Η μέθοδος του Jacobi μπορεί να υλοποιηθεί παράλληλα, ενώ η μέθοδος Gauss-Siedel δεν μπορεί.

Γράφουμε τις μεθόδους αυτές σε μια βολική μορφή για τη θεωρία ως εξής.

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

$$U = \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ & 0 & & \vdots \\ & & \ddots & a_{n-1,n} \\ \text{\Large 0} & & & 0 \end{pmatrix}$$

Δεν έχουν <u>καμία</u> σχέση με την ανάλυση L, U!

Προφανώς $A = L + D + U$

- Μέθοδος Jacobi

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

- Μέθοδος Gauss-Siedel

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

<u>Γενική επαναληπτική μέθοδος</u>

$A = M - N$, M αντιστρέψιμος

- Jacobi: $M_J = D$, $N_J = -(L+U)$

Gauss-Siedel $M_{GS} = L + D$, $N_{GS} = -U$

$$Ax = b \Leftrightarrow (M-N)x = b \Leftrightarrow Mx = Nx + b$$
$$x^{(m)} \rightarrow x^{(m+1)}$$
$$Mx^{(m+1)} = Nx^{(m)} + b$$

<u>Ισχυρισμός:</u> Αν η ακολουθία $(x^{(m)})_{m\in\mathbb{N}}$ συγκλίνει τότε συγκλίνει στη λύση x.

Πράγματι, έχουμε (αν $x^{(m)} \rightarrow y$, $m \rightarrow \infty$)

$$\underset{\downarrow}{Mx^{(m+1)}} = \underset{\downarrow}{Nx^{(m)}} + b$$

$$My = Ny + b \Rightarrow (M-N)y = b \Leftrightarrow Ay = b \Rightarrow y = x$$

Ερώτημα: Πότε συγκλίνει η γενική επαναληπτική μέθοδος (για οποιαδήποτε αρχική προσέγγιση $x^{(0)}$);

Απάντηση:

Έχουμε: $\left.\begin{aligned} Mx^{(m+1)} &= Nx^{(m)} + b \\ Mx &= Nx + b \end{aligned}\right\} \Rightarrow$

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

$\boxed{x^{(m)} - x = G^m (x^{(0)} - x)}$ $m = 1, 2, \ldots$

Άρα: $\|x^{(m)} - x\| \leq \|G^m\| \cdot \|x^{(0)} - x\|$ (*)

Για κατάλληλο $x^{(0)}$ η (*) ισχύει ως ισότητα.

Άρα η μέθοδος συγκλίνει αν και μόνο αν $\|G^m\| \to 0$, για $m \to \infty$

ή ισοδύναμα $\boxed{G^m \to 0 \text{ για } m \to \infty}$

Μάθημα 20ο

23/4/2015

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

"Γενική επαναληπτική μέθοδος"

$A = M - N$, M αντιστρέψιμος

$Ax=b \iff Mx = Nx + b$

(+) $Mx^{(m+1)} = Nx^{(m)} + b \qquad m=0,1,2,\dots$

Τότε: (*) $x^{(m)} - x = G^{(m)}(x^{(0)} - x)$, $m=0,1,2,\dots$ με $G = M^{-1}x$

"Jacobi"

$M_J = D$, $N_J = -(L+U)$

"Gauss-Seidel"

$M_{GS} = L + D$, $N_{GS} = -U$

Έστω $P \in \mathbb{C}^{n,n}$ και $\lambda_i(P) \in \mathbb{C}$, $i=1,\dots,n$ οι ιδιοτιμές του.

"Φασματική ακτίνα"

$\rho(P) := \max_{1 \le i \le n} |\lambda_i(P)|$ αποτελεί τη φασματική ακτίνα.

Λήμμα: (Σχέση φυσικής νόρμας και φασματικής ακτίνας)
Έστω $\|\cdot\|$ μια νόρμα στον $\mathbb{C}^n$ και η παραγόμενη φυσική νόρμα στον $\mathbb{C}^{n,n}$. Τότε για $P \in \mathbb{C}^{n,n}$ ισχύει $\rho(P) \le \|P\|$

Αντίστροφα, για κάθε πίνακα $P \in \mathbb{C}^{n,n}$ και $\varepsilon > 0$ υπάρχει νόρμα $|||\cdot|||$ στον $\mathbb{C}^n$ τ.ω. για την παραγόμενη φυσική νόρμα να έχουμε

$$|||P||| \le \rho(P) + \varepsilon$$

απόδειξη: (μόνο του πρώτου μέρους)
Έστω $\lambda \in \mathbb{C}$ ιδιοτιμή του P. Τότε, υπάρχει $z \in \mathbb{C}^n$, $z \neq 0$, τ.ω. $Pz = \lambda z$.
Άρα, $\|\lambda z\| = \|Pz\|$
οπότε $|\lambda| \cdot \|z\| = \|Pz\|$
Συνεπώς, $|\lambda| \cdot \cancel{\|z\|} \leq \|P\| \cdot \cancel{\|z\|}$
άρα $|\lambda| \leq \|P\|$

Επομένως, αφού αυτή η εκτίμηση ισχύει για όλες τις ιδιοτιμές του P έχουμε: $\rho(P) \leq \|P\|$.

Θεώρημα (ικανές και αναγκαίες συνθήκες σύγκλισης επαναληπτικών μεθόδων)
Έστω x η λύση του $Ax = b$. Τότε, τα παρακάτω είναι ισοδύναμα:
α. Η επαναληπτική μέθοδος (+) συγκλίνει, δηλαδή για κάθε $x^{(0)} \in \mathbb{C}^n$ έχουμε $x^{(m)} \to x$, $m \to \infty$
β. $\rho(G) < 1$, με $G = M^{-1}N$.
γ. Υπάρχει φυσική νόρμα $|||\ |||$ τ.ω. $|||G||| < 1$.
δ. $\lim_{m\to\infty} G^m = 0$

απόδειξη: (α) ⇒ (β) ⇒ (γ) ⇒ (δ) ⇒ (α)
• (α) ⇒ (β)
Έστω $x^{(m)} \to x$, $m \to \infty$. Τότε, σύμφωνα με την (*), έχουμε $G^m(x^{(0)} - x) \to 0$, $m \to \infty$, για οποιοδήποτε $x^{(0)} \in \mathbb{C}^n$.
Έστω $y \in \mathbb{C}^n$. Για $x^{(0)} = x + y$, αυτό σημαίνει ότι $G^m y \to 0$, $m \to \infty$
Έστω λ ιδιοτιμή του G και z αντίστοιχο ιδιοδιάνυσμα. Τότε $Gz = \lambda z$.
οπότε, $G^2 z = \lambda^2 z$, ..., $G^m z = \lambda^m z$, $m \in \mathbb{N}$
Σύμφωνα με την υπόθεση έχουμε ότι $G^m z \to 0$, $m \to \infty$, οπότε $\lambda^m z \to 0$, $m \to \infty$. Αλλά, $\|\lambda^m z\| = |\lambda|^m \|z\|^{\neq 0}$, οπότε για να ισχύει $|\lambda|^m \to 0$, $m \to \infty$, πρέπει και αρκεί $|\lambda| < 1$
Αφού αυτό συμβαίνει για όλες τις ιδιοτιμές του G θα έχουμε $\rho(G) < 1$.

• (β) ⇒ (γ)
Για $0 < \varepsilon < 1 - \rho(G)$ έχουμε (σύμφωνα με το Λήμμα) ότι υπάρχει φυσική νόρμα $|||\cdot|||$ τ.ω. $|||G||| \leq \rho(G) + \varepsilon < 1$

• (γ) ⇒ (δ):
Έχουμε $|||G^m||| \leq \underbrace{|||G|||\cdot|||G|||\cdots|||G|||}_{m \text{ φορές}}$
$= |||G|||^m \to 0, \ m \to \infty$
$\Rightarrow G^m \to 0, \ m \to \infty$

• (δ) ⇒ (α)
$x^{(m)} - x = G^{(m)}(x^{(0)} - x) \Rightarrow \|x^{(m)} - x\| \leq \|G^m\| \cdot \|x^{(0)} - x\| + 0, \ m \to \infty \Rightarrow$
$\Rightarrow x^{(m)} \to x, \ m \to \infty$

Παράδειγμα
$G = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix}, \ \|G\|_\infty = 2, \ \|G\|_1 = 2, \ \|G\|_2 = (\rho(G^TG))^{1/2} = 2$

Έχω:
$G^T = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}, \ G^TG = \begin{pmatrix} 0 & 0 \\ 2 & 0 \end{pmatrix}\begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}$

ιδιοτιμές: $\lambda_1 = 0, \ \lambda_2 = 4$

Δεν βγαίνει συμπέρασμα!

ιδιοτιμές του G: $\lambda_1 = \lambda_2 = 0$
$\rho(G) = 0 < 1$
Η μέθοδος συγκλίνει.

$G^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \Rightarrow G^m \to 0, \ m \to \infty$ σύγκλιση!

Έστω $A \in \mathbb{C}^{n,n}$, με αυστηρά κυριαρχική διαγώνιο, δηλαδή

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

Στόχος: Να αποδείξουμε ότι οι μέθοδοι του Jacobi και των Gauss-Seidel συγκλίνουν για συστήματα $Ax=b$, αν ο A έχει αυστηρά κυριαρχική διαγώνιο.

↳ Ικανή συνθήκη! ΌΧΙ αναγκαία!!

Λήμμα (Ανισότητα του Gerschgorin)

Έστω $A=(a_{ij}) \in \mathbb{C}^{n,n}$ και $\lambda \in \mathbb{C}$ μια ιδιοτιμή του. Τότε, υπάρχει ένα $s \in \{1, \dots, n\}$ τ.ω.

$$|a_{ss}-\lambda| \leq \sum_{\substack{j=1 \\ j\neq s}}^{n} |a_{sj}| \longrightarrow \text{Θέτω } r_s = \sum_{\substack{j=1 \\ j\neq s}}^{n} |a_{sj}|, \ s=1,\dots,n$$

απόδειξη:

Έστω $z \in \mathbb{C}^n$ ιδιοδιάνυσμα του A ως προς την ιδιοτιμή λ, $Az=\lambda z$ $(z\neq 0)$

Τότε, $(Az)_i = \lambda z_i$, $i=1,\dots,n$

δηλαδή

$$\sum_{j=1}^{n} a_{ij} z_j = \lambda z_i, \quad i=1,\dots,n$$

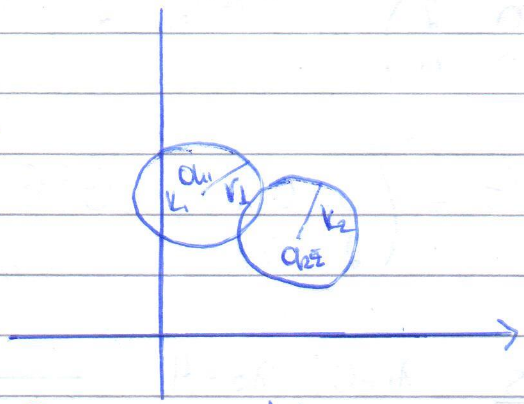
οπότε, $(a_{ii}-\lambda) z_i = -\sum_{\substack{j=1 \\ j\neq i}}^{n} a_{ij} z_j$,

$i=1,\dots,n$

Επομένως, $|a_{ii}-\lambda|\cdot|z_i| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|$.

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

Θεωρούμε στο μιγαδικό επίπεδο τους κύκλους του Gerschgorin με κέντρο a_{ii} και ακτίνα r_i, $K_i(a_{ii}, r_i)$ $i=1,\dots,n$



Συμπέρασμα: Όλες οι ιδιοτιμές του A ανήκουν στην ένωση των κύκλων του Gerschgorin

Έστω $s \in \{1,\dots,n\}$ τ.ω. $|z_s| = \max_{1\leq i \leq n} |z_i| = \|z\|_\infty$

Τότε έχουμε (για $i=s$)

$$|a_{ss}-\lambda|\cdot|z_s| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{sj}| \cdot |z_j| \leq |z_s|$$

$$\Rightarrow |a_{ss}-\lambda|\cdot\cancel{|z_s|} \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{sj}|\cdot\cancel{|z_s|} \overset{z_s\neq 0}{\Rightarrow} |a_{ss}-\lambda| \leq \sum_{\substack{j=1 \\ j\neq s}}^{n} |a_{sj}|$$

Ισχυρισμός Ένας πίνακας $A \in \mathbb{C}^{n,n}$ είναι αντιστρέψιμος, αν και μόνο αν το μηδέν δεν είναι ιδιοτιμή του.

απόδειξη:

A αντιστρέψιμος $\Leftrightarrow (Az = 0 \Rightarrow z = 0) \Leftrightarrow (Az = 0 \cdot z \Leftrightarrow z = 0) \Leftrightarrow$ το μηδέν δεν είναι ιδιοτιμή του A.

Ισχυρισμός Αν ο A έχει αυστηρά κυριαρχική διαγώνιο, τότε είναι αντιστρέψιμος.

απόδειξη:

Έστω ότι ο A δεν είναι αντιστρέψιμος. Τότε, το $\lambda = 0$ είναι ιδιοτιμή του. Επομένως, σύμφωνα με την συνθήκη του Gerschgorin.

Ισχύει, $|a_{ss} - \lambda| \leq \sum_{\substack{j=1 \\ j \neq s}}^{n} |a_{sj}|$ για κάποιο $s \in \{1, \dots, n\}$. Άτοπο!

Πρόταση (σύγκλιση των μεθόδων του Jacobi και των Gauss-Saidel για συστήματα με πίνακα με αυστηρά κυριαρχική διαγώνιο)

Έστω ότι ο πίνακας $A \in \mathbb{C}^{n,n}$ έχει αυστηρά κυριαρχική διαγώνιο.

Τότε:

(α) Οι πίνακες επανάληψης $G_J = -D^{-1}(L+U)$ και $G_{GS} = -(L+D)^{-1}U$ των μεθόδων του Jacobi και Gauss-Saidel, αντίστοιχα, ικανοποιούν τις συνθήκες $\|G_J\|_\infty < 1$, $\|G_{GS}\| < 1$.

(β) Οι μέθοδοι του Jacobi και των Gauss-Saidel συγκλίνουν

απόδειξη:

Το (β) έπεται αμέσως από το (α) σύμφωνα με το θεώρημα για τη σύγκλιση της γενικής επαναληπτικής μεθόδου.

(α) 1ο Μέθοδος Jacobi

$$G_J = -D^{-1}(L+U) = \begin{pmatrix} \frac{1}{a_{11}} & & 0 \\ 0 & \frac{1}{a_{22}} & \ddots \\ 0 & 0 \cdots & \frac{1}{a_{nn}} \end{pmatrix} \begin{pmatrix} 0 & a_{12} & \cdots & a_{1n} \\ a_{21} & 0 & a_{23} & a_{3n} \\ \vdots & & \ddots & \\ a_{n1} & \cdots & a_{n,n-1} & 0 \end{pmatrix} =$$

$$= - \begin{pmatrix} 0 & \frac{a_{12}}{a_{11}} & \cdots & \cdots & \frac{a_{1n}}{a_{11}} \\ \frac{a_{21}}{a_{22}} & 0 & \frac{a_{23}}{a_{22}} & \cdots & \frac{a_{2n}}{a_{22}} \\ \vdots & & \ddots & & \vdots \\ \frac{a_{n1}}{a_{nn}} & \cdots & \cdots & \frac{a_{n,n-1}}{a_{nn}} & 0 \end{pmatrix} \Rightarrow \|G_J\|_\infty = \max_{1\le i\le n} \sum_{\substack{j=1 \\ j\neq i}}^{n} \left| \frac{a_{ij}}{a_{ii}} \right| \Rightarrow$$

$$\|G_J\|_\infty = \max_{1\le i\le n} \left(\frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \right) < 1$$ (*) αυστηρά κυριαρχική διαγώνιος

2ο Μέθοδος Gauss-Seidel

Ο $(L+D)^{-1}$ δεν υπολογίζεται εύκολα, οπότε δεν μπορούμε να βρούμε τον G_{GS}, ανεπώς ούτε την $\|G_{GS}\|_\infty$

Έστω $C := \max_{1\le i\le n} \left(\frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}| \right)$

Προφανώς, $C < 1$ (προκύπτει από την (*))

Ισχυρισμός: $\|G_{GS}\|_\infty < C$

Έστω $y \in \mathbb{C}^n$. Θέτουμε $u = G_{GS} \cdot y$. Τότε, έχουμε $u = -(L+D)^{-1} U y \Leftrightarrow (L+D)u = -Uy$, δηλαδή

$$\sum_{j=1}^{i} a_{ij} u_j = -\sum_{j=i+1}^{n} a_{ij} y_j \quad , \; i=1,\dots,n$$

Δηλαδή $u_i = \frac{1}{a_{ii}} \left[-\sum_{j=1}^{i-1} - \sum_{j=i+1}^{n} a_{ij} y_j \right] \quad , \; i=1,\dots,n$

Ισχυρισμός: $|u_i| \le C \|y\|_\infty \quad i=1,\dots,n$ **

απόδειξη:

Για $i=1$: $u_1 = -\frac{1}{a_{11}} \sum_{j=2}^{n} a_{ij} y_j \Rightarrow |u_1| \le \frac{1}{|a_{11}|} \sum_{j=2}^{n} |a_{ij}| \cdot |y_j|$ $\le \|y\|_\infty$

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

Επαγωγικά: Έστω ότι η (*) ισχύει για $i=1,\dots,i-1$

Τότε;

$$|u_i| \le \frac{1}{|a_{ii}|}\left(\sum_{j=1}^{i-1} |a_{ij}| \cdot \underbrace{|u_j|}_{\le C\|y\|_\infty \le \|y\|_\infty} + \sum_{j=i+1}^{n} |a_{ij}| \cdot \underbrace{|y_j|}_{\le \|y\|_\infty} \right) \Rightarrow$$

$$\Rightarrow |u_i| \le \underbrace{\left[\frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ j\ne i}}^{n} |a_{ij}| \right]}_{\le C} \cdot \|y\|_\infty \Rightarrow |u_i| \le C\,\|y\|_\infty$$

Συμπέρασμα: $u = G_{GS} \cdot y \Rightarrow \|u\|_\infty \le C\|y\|_\infty$

Δηλαδή $\forall y \in \mathbb{C}^n \quad \|G_{GS}\,y\|_\infty \le C\|y\|_\infty \Rightarrow \|G_{GS}\|_\infty \le C$

Ασκήσεις $3^{ου}$ κεφαλαίου

3/4/2015

Άσκηση 3.1 (βιβλίου)

Αποδείξτε ότι το γινόμενο δύο άνω τριγωνικών πινάκων είναι άνω τριγωνικός πίνακας. Επίσης στην περίπτωση που ένας άνω τριγωνικός πίνακας $U \in \mathbb{R}^{n,n}$ είναι αντιστρέψιμος, τότε για τη λύση $x \in \mathbb{R}^n$ του συστήματος $Ux = e^k$ με $e^k \in \mathbb{R}^n$ το διάνυσμα με συνιστώσα ένα στη θέση k και όλες τις άλλες συνιστώσες μηδέν, ισχύει $x_{k+1} = \dots = x_n = 0$, συνεπώς και ο αντίστροφος του U είναι άνω τριγωνικός. Ανάλογα για κάτω τριγωνικούς πίνακες.

Λύση:

$$(UW)_{ij} = \sum_{k=1}^{n} u_{ik} w_{kj} = \sum_{k=i}^{n} u_{ik} w_{kj}$$

$u_{i1} = u_{i2} = \dots = u_{i,i-1} = 0$

$$\Rightarrow (UW)_{ij} = \sum_{k=i}^{n} u_{ik} w_{kj} = u_{ij} w_{ij} + u_{i,i+1} w_{i+1,j} + \dots + u_{in} w_{nj}$$

Για $i > j$ έχουμε $w_{ij} = w_{i+1,j} = \dots = w_{nj} = 0$ οπότε $(UW)_{ij} = 0$

$U^{-1} = (u^1, u^2, \dots, u^n)$ με $u^1, u^2, \dots, u^n \in \mathbb{R}^n$ τ.ω.

$U u^2 = e^k$, $k = 1, \dots, n$ $(*)$

U^{-1} άνω τριγωνικός $\Leftrightarrow$ $u_i^k = 0$, $i = k+1, \dots, n$

Λύνοντας το σύστημα με οπισθοδρόμηση παίρνουμε:

$u_n^k = 0$, $u_{n-1}^k = 0$, ..., $u_{k+1}^k = 0$

$$\begin{pmatrix} u_{11} & \dots & a_{1n} \\ & u_{22} & a_{2n} \\ 0 & \ddots & \vdots \\ & & u_{nn} \end{pmatrix} \cdot \begin{pmatrix} u_1^k \\ u_2^k \\ u_k^k \\ u_n^k \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow e^k$$

Άσκηση 3.3 (βιβλίου)

Έστω $A, B \in \mathbb{R}^{n,n}$, A αντιστρέψιμος και $b \in \mathbb{R}^n$. Πώς υπολογίζουμε, κατά το δυνατόν οικονομικότερα από άποψης πλήθους πράξεων και μνήμης, τα διανύσματα $A^{-4}b$ και $A^{-1}BA^{-1}b$;

Λύση:

$$x = A^{-4}b \Rightarrow A^4 x = b \Rightarrow A(\underbrace{A^3 x}_{y}) = b$$

$$Ay = b \rightsquigarrow A^3 x = y$$

$$A(A^2 x) = y \leftarrow \text{γνωστό} \rightsquigarrow v$$

$$A^2 x = v, \quad A(\underbrace{Ax}_{u}) = u \rightsquigarrow Ax = u \leftarrow \text{γνωστό}$$

$Ax = u$

$$\begin{aligned} Ay &= b \to y \\ Av &= y \to v \\ Au &= v \to u \\ Ax &= u \to x \end{aligned}$$

$$\rightsquigarrow x = A^{-1}BA^{-1}b \Rightarrow Ax = BA^{-1}b$$

$$y = A^{-1}b \Leftrightarrow Ay = b \rightsquigarrow y$$

Τότε γνωστό $Ax = By \leftarrow$ γνωστό

Άσκηση 3.6 (βιβλίου)

Θεωρούμε κάτω τριγωνικούς πίνακες A_i, $i = 1, \dots, n-1$, με μονάδες στη διαγώνιο και μηδενικά στις άλλες θέσεις που δεν δίνονται στοιχεία της μορφής

$$A_i = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & a_{i+1,i} & \ddots & \\ & & a_{ni} & & 1 \end{pmatrix}$$

Βεβαιωθείτε ότι:

$$A_i^{-1} = \begin{pmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & -a_{i+1,i} & \ddots & \\ & & -a_{ni} & & 1 \end{pmatrix}$$

και για $i < j$

$$A_iA_j = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \\ & & 1 & & & \\ & & a_{i+1,i} & \ddots & & \\ & & \vdots & a_{j+1,j} & \ddots & \\ & & a_{ni} & a_{nj} & & 1 \end{pmatrix}$$

Παρόμοια για γινόμενα με τρεις παράγοντες, $A_iA_jA_k$ με $i<j$ κλπ.

Λύση:

Για $i>j$

$$A_iA_j = \begin{pmatrix} 1 & & & & & \\ & \ddots & & & & \mathbf{0} \\ & & 1 & & & \\ & & a_{i+1,i} & \ddots & & \\ & \mathbf{0} & \vdots & a_{j+1,j} & \ddots & \\ & & a_{ni} & a_{nj} & & 1 \end{pmatrix} \qquad A,B \in \mathbb{R}^{n,n}$$

Τότε $AB = A + B - I_n + \underbrace{(A-I_n)(B-I_n)}_{AB-A-B-I_n}$

Για $i \le j$ έχουμε

$$(A_i - I_n)(A_j - I_n) = \begin{pmatrix} 0 & & & \\ & 0 & & \mathbf{0} \\ & \vdots & \ddots & \\ & a_{i+1,i} & & 0 \\ & \vdots & & \ddots \\ & a_{ni} & & & 0 \end{pmatrix} \begin{pmatrix} 0 & & & \\ & \ddots & & \\ & & 0 & \\ & & a_{j+1,j} & \ddots \\ & & \vdots & & 0 \\ & & a_{ni} & & \end{pmatrix} = 0$$

Η γραμμή i του δεύτερου πίνακα είναι 0.
Οπότε προκύπτουν οι ζητούμενες σχέσεις.

24/4/2015

Άσκηση 3.7 (βιβλίου)

$A \in \mathbb{R}^{n,n}$, αντιστρέψιμος

$A = LU$ και $A = \tilde{L}\tilde{U}$, $\tilde{L}, L$ κάτω τριγωνικοί με μονάδες στη διαγώνιο
$\tilde{U}, U$ άνω τριγωνικοί

Λύση:

$LU = \tilde{L}\tilde{U} \Rightarrow \tilde{L}^{-1}LU = \tilde{U} \Rightarrow \tilde{L}^{-1}L = \tilde{U}U^{-1}$

Σύμφωνα με την άσκηση 3.1 στο αριστερό μέλος έχουμε ένα κάτω τριγωνικό πίνακα στο δεξί έναν άνω.

Άρα οι πίνακες $\tilde{L}^{-1}L$ και $\tilde{U}U^{-1}$ είναι διαγώνιοι.

$\tilde{L}^{-1}L = \tilde{U}U^{-1} = D = \text{diag}(d_{11}, \dots, d_{nn})$

$\tilde{L}^{-1}L = D \Rightarrow L = \tilde{L}D \Rightarrow \underset{\overset{\shortparallel}{1}}{L_{ii}} = \underset{\overset{\shortparallel}{1}}{\tilde{L}_{ii}}\, d_{ii}$, $i = 1, \dots, n$

$\Rightarrow d_{ii} = 1$, $i = 1, \dots, n$

Συμπέρασμα: $D = I_n$

$\tilde{L}^{-1}L = \tilde{U}U^{-1} = I_n$

$\tilde{L}^{-1}L = I_n \Rightarrow L = \tilde{L}I_n \Rightarrow L = \tilde{L}$

$\tilde{U}U^{-1} = I_n \Rightarrow \tilde{U} = I_n U \Rightarrow \tilde{U} = U$.

Άσκηση 3.10 (βιβλίου)

$A = (a_{ij}) \in \mathbb{R}^{n,n}$

κύριες ορίζουσες του A:

$\delta_1 = a_{11}$, $\delta_2 = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}$, ..., $\delta_n = \begin{vmatrix} a_{11} & \cdots & a_{1,n-1} \\ & & \\ a_{n-1,1} & \cdots & a_{n-1,n-1} \end{vmatrix}$

Αν $\delta_i \neq 0$, $i = 1, 2, \dots, n-1$

Νδο $A = LU$ (Αν ο A είναι αντιστρέψιμος, τότε η συνθήκη αυτή είναι και αναγκαία)

Απόδειξη:

Αφού $a_{11} \neq 0$, το πρώτο βήμα της απαλοιφής του Gauss γίνεται χωρίς πρόβλημα (χωρίς εναλλαγές γραμμών)

Έστω ότι γίνεται χωρίς εναλλαγές γραμμών $i-1$ βήματα. Τότε

$$\det\begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \cdots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \cdots & a_{2n}^{(2)} \\ & & \ddots & \vdots \\ & 0 & & a_{ii}^{(i)} \end{pmatrix} = \det\begin{pmatrix} a_{11} & \cdots & a_{1i} \\ a_{21} & \cdots & a_{2i} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{ii} \end{pmatrix} = \delta_i \neq 0$$

με τις πράξεις που έγιναν δεν άλλαξε η τιμή της ορίζουσας

$\Rightarrow a_{11}^{(1)} a_{22}^{(2)} \cdots a_{ii}^{(i)} = \delta_i \neq 0 \Rightarrow a_{ii}^{(i)} \neq 0$

Άρα και το i-οστό βήμα γίνεται στην απαλοιφή Gauss χωρίς εναλλαγές γραμμών.

Άσκηση 3.11 (βιβλίου)

$A = (a_{ij}) \in \mathbb{R}^{n,n}$ με αυστηρά κυριαρχική διαγώνιο

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

<u>Ν.δ.ο.</u> A αντιστρέψιμος (χωρίς θεώρημα Gershgorin) $A = LU$

<u>Απόδειξη:</u>

- $Ax = 0 \Rightarrow (Ax)_i = 0, \quad i=1,\ldots,n$

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

$|a_{ii}x_i| \leq \sum_{\substack{j=1 \\ j\neq i}}^{n} |a_{ij}|\cdot|x_j|, \quad i=1,\ldots,n$

Έστω $s \in \{1,\ldots,n\}$ τ.ω. $|x_s| = \|x\|_\infty = \max_{1\leq i\leq n} |x_i|$

Τότε για $i=s$ έχουμε:

$|a_{ss}|\cdot|x_s| \leq \sum_{\substack{j=1 \\ j\neq s}}^{n} |a_{sj}|\,|x_j|$ $\leq |x_s|$

Έστω $x \neq 0$. Τότε $x_s \neq 0$ και έχουμε

$|a_{ss}|\cdot\cancel{|x_s|} \leq \sum_{\substack{j=1 \\ j\neq s}}^{n} |a_{sj}|\,\cancel{|x_s|}$

$\Rightarrow |a_{ss}| = \sum_{\substack{j=1 \\ j\neq s}}^{n} |a_{sj}|$ ↯

Άρα $Ax=0 \Rightarrow x=0 \Rightarrow A$ αντιστρέψιμος

$$A_i = \begin{pmatrix} a_{11} & \cdots & a_{1i} \\ a_{21} & \cdots & a_{2i} \\ \vdots & & \vdots \\ a_{i1} & \cdots & a_{ii} \end{pmatrix}$$ έχει αυστηρά κυριαρχική διαγώνιο (αφού ο A έχει αυτή την ιδιότητα)

$\Rightarrow \delta_i = \det A_i \neq 0 \quad , \quad i=1,2,\ldots,n$

Σύμφωνα με την προηγούμενη Άσκηση ο A αναλύεται σε γινόμενο LU χωρίς εναλλαγές γραμμών.

Άσκηση 3.23

α) $1 < p, q < \infty$ τω $\frac{1}{p} + \frac{1}{q} = 1 \qquad a, b > 0$

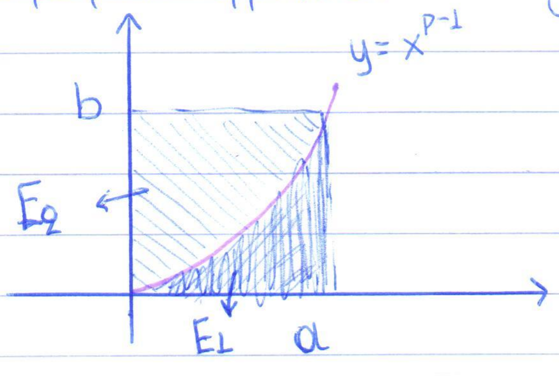
<u>Ν.δ.ο.</u> $ab \leq \frac{a^p}{p} + \frac{b^q}{q}$ (ανισότητα Young)

<u>Απόδειξη:</u>

Ειδική περίπτωση: $p=q=2$

$ab \leq \frac{a^2}{2} + \frac{b^2}{2} \Leftrightarrow a^2+b^2-2ab \geq 0 \Leftrightarrow (a-b)^2 \geq 0$ σωστό

Γεωμετρική ερμηνεία: (στη γενική περίπτωση)



$$E_1 = \int_0^a x^{p-1}dx = \left[\frac{x^p}{p}\right]_{x=0}^{x=a} = \frac{a^p}{p}$$

$$E_2 = \int_0^b y^{1/p-1}\,dy = \int_0^b y^{q-1}dy = \left[\frac{y^q}{q}\right]_{y=0}^{y=b}$$

$$E_2 = \frac{b^q}{q}$$

Ισχυρισμός: $\frac{1}{p-1} = q-1$

$\Leftrightarrow \cancel{1} = (p-1)(q-1) = -p-q+pq+\cancel{1} \Leftrightarrow p+q = pq \Leftrightarrow \frac{p}{pq} + \frac{q}{pq} = 1 \Rightarrow$

$\frac{1}{p} + \frac{1}{q} = 1$

Το ορθογώνιο με α, β πλευρές έχει $E \leq E_1 + E_2$

ooo

συνεχίζεται

<u>Σημείωση:</u> Τις ανισότητες αυτής της άσκησης τις μαθαίνω απ' έξω.

30/4/2015

Άσκηση 3.23 (βιβλίου) συνέχεια

β) Για $1 \le p < \infty$, ορίζουμε $\|\cdot\|_p : \mathbb{C}^n \to \mathbb{R}$, $\|x\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{1/p}$

Για $x, y \in \mathbb{C}^n$ ΝΔΟ $\sum_{i=1}^{n} |x_i \cdot y_i| \le \|x\|_p \cdot \|y\|_p$ με $\frac{1}{p} + \frac{1}{q} = 1$

(ανισότητα του Hölder)

απόδειξη:

→ Για $p=2$ προκύπτει ως ειδική περίπτωση η ανισότητα των Caushy-Schwartz

→ για $p=1$, η ανισότητα παίρνει τη μορφή

$$\sum_{i=1}^{n} |x_i y_i| \le \left(\sum_{i=1}^{n} |x_i|\right) \max_{1 \le j \le n} |y_j|$$

που ισχύει προφανώς

→ για $x=0$ ή $y=0$ είναι επίσης προφανής η ανισότητα.

→ Έστω λοιπόν $p>1$, $x \neq 0$, $y \neq 0$

1η περίπτωση $\|x\|_p = \|y\|_p = 1$

Σύμφωνα με την ανισότητα του Young (α ερώτημα) έχουμε:

$$|x_i - y_i| \le \frac{|x_i|^p}{p} + \frac{|y_i|^q}{q}$$

Άρα:

$$\sum_{i=1}^{n} |x_i y_i| \le \frac{1}{p} \cdot \underbrace{\sum_{i=1}^{n} |x_i|^p}_{= \|x\|_p^p = 1} + \frac{1}{q} \underbrace{\sum_{i=1}^{n} |y_i|^q}_{= \|y\|_q^q = 1} = \frac{1}{p} + \frac{1}{q} = 1 =$$

$\|x\|_p \cdot \|y\|_q$

αποδείχτηκε

2^{η} περίπτωση (γενική)

Θέτουμε: $\tilde{x} = \dfrac{1}{\|x\|_p} x$, $\tilde{y} = \dfrac{1}{\|y\|_q} y$

προφανώς $\|\tilde{x}\|_p = 1$, $\|\tilde{y}\|_q = 1$

Σύμφωνα με την 1^{η} περίπτωση έχουμε $\sum_{i=1}^{n} |\tilde{x}_i \tilde{y}_i| \leq 1$

Επομένως, $\sum_{i=1}^{n} \dfrac{1}{\|x\|_p \|y\|_q} |x_i y_i| \leq 1 \Rightarrow \sum_{i=1}^{n} |x_i y_i| \leq \|x\|_p \cdot \|y\|_q$

γ) N.δ.ο η $\|\cdot\|_p$ είναι νόρμα.

απόδειξη:

⇝ Για $p=1$ είναι εύκολο και το είδαμε στη θεωρία.

⇝ Έστω λοιπόν $1 < p < \infty$

Τότε (N1) $\|x\|_p = 0 \iff x = 0$ τετριμμένο

(N2) $\forall \lambda \in \mathbb{C}\ \forall x \in \mathbb{C}^n \quad \|\lambda x\|_p = |\lambda| \cdot \|x\|_p$ τετριμμένο.

Μένει να αποδείξουμε την τριγωνική ανισότητα,

$\forall x, y \in \mathbb{C}^n \quad \|x+y\|_p \leq \|x\|_p + \|y\|_p$ (ανισότητα Minkowski)

$$\|x+y\|_p^p = \sum_{i=1}^{n} |x_i + y_i|^p = \sum_{i=1}^{n} \underbrace{|x_i+y_i|}_{\leq |x_i| + |y_i|} \cdot |x_i+y_i|^{p-1} \leq$$

$$\leq \sum_{i=1}^{n} |x_i| \cdot |x_i + y_i|^{p-1} + \sum_{i=1}^{n} |y_i| \cdot |x_i+y_i|^{p-1} \qquad \text{(Hölder)}$$

$$\leq \|x\|_p \underline{\left(\sum_{i=1}^{n} \left[|x_i+y_i|^{p-1} \right]^q \right)^{1/q}} + \|y\|_p \underline{\left(\sum_{i=1}^{n} \left[|x_i+y_i|^{p-1} \right]^q \right)^{1/q}}$$

$$\leq (\|x\|_p + \|y\|_p) \left(\sum_{i=1}^{n} |x_i + y_i|^{\overbrace{(p-1)q}^{=p}} \right)^{1/q} \qquad 1/q = \frac{p-1}{p}$$

$$= (\|x\|_p + \|y\|_p) \Bigg[\underbrace{\left(\sum_{i=1}^{n} |x_i+y_i|^p \right)^{1/p}}_{= \|x+y\|_p} \Bigg]^{p-1}$$

Ανακεφαλαιώνοντας:

$\|x+y\|_p^p \le (\|x\|_p + \|y\|_p)\|x+y\|_p^{p-1} \Rightarrow \|x+y\|_p \le \|x\|_p + \|y\|_p$

δ) $1 \le p < q \le \infty$ Νδο: $\forall x \in \mathbb{C}^n$ $\|x\|_q \le \|x\|_p$ (ανισότητα Jensen)

απόδειξη:

Για $x = e^i$ ισχύει ως ισότητα.

Προφανώς ισχύει $|x_i| \le \|x\|_p = \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p} \Rightarrow$

$\Rightarrow |x_i| \le \|x\|_p \Rightarrow \max_{1\le i\le n} |x_i| \le \|x\|_p \Rightarrow \|x\|_\infty \le \|x\|_p$

Συμπέρασμα: Η ανισότητα ισχύει για $q = \infty$.

Έστω τώρα $q < \infty$. Τότε

$|x_i|^q = |x_i|^p \cdot \underbrace{|x_i|}_{\le \|x\|_p}{}^{q-p} \Rightarrow$

$|x_i|^q \le |x_i|^p \cdot \|x\|_p^{q-p} \Rightarrow \sum_{i=1}^{n} |x_i|^q \le \left(\sum_{i=1}^{n} |x_i|^q\right) \|x\|_p^{q-p} \Rightarrow$

$\Rightarrow \|x\|_q^q \le \|x\|_p^p \cdot \|x\|_p^{q-p} \Rightarrow \|x\|_q^q \le \|x\|_p^q \Rightarrow \|x\|_q \le \|x\|_p$

Άσκηση 3.24 (Βιβλίου)

$\mathbb{R}^n$, $\|\cdot\|_1$, $\|\cdot\|_2$, $\|\cdot\|_\infty$

Να βρεθούν οι βέλτιστες σταθερές σύγκρισης.

απόδειξη:

Σύμφωνα με την ανισότητα του Jensen $\forall x \in \mathbb{R}^n$ $\|x\|_\infty \le \|x\|_2 \le \|x\|_1$ και αυτές οι σταθερές (οι μονάδες) δεν μπορούν να βελτιωθούν (Για $x = e^i$ ισχύουν για ισότητες)

a) $\forall x \in \mathbb{R}^n$ $\|x\|_1 \le n\|x\|_\infty$

Πράγματι: $\|x\|_1 = \sum_{i=1}^{n} \underbrace{|x_i|}_{\le \|x\|_\infty} \le n\|x\|_\infty$

(Για $x = (1, 1, \dots, 1)$ ισχύει ως ισότητα) β) $\forall x \in \mathbb{R}^n \quad \|x\|_2 \leq \sqrt{n}\, \|x\|_\infty$

Πράγματι: $\|x\|_2^2 = \sum_{i=1}^{n} |x_i|^2 \leq n \|x\|_\infty^2 \Rightarrow \|x\|_2 \leq \sqrt{n}\, \|x\|_\infty$

(Για $x = (1, 1, \dots, 1)$ ισχύει ως ισότητα)

γ) $\forall x \in \mathbb{R}^n \quad \|x\|_1 \leq \sqrt{n}\, \|x\|_2$

Πράγματι: $\|x\|_1 = \sum_{i=1}^{n} |x_i| = \sum_{i=1}^{n} 1 \cdot |x_i| \underset{CS}{\leq} \left(\sum_{i=1}^{n} |x_i|^2\right)^{1/2} \cdot \left(\sum_{i=1}^{n} 1^2\right)^{1/2} = \sqrt{n}\, \|x\|_2$

(Για $x = (1, 1, \dots, 1)^T$ ισχύει ως ισότητα)

5/5/2015

Άσκηση 3.31 (βιβλίου)

Υπάρχει φυσική νόρμα $\|\cdot\|$ τω για τον $A = \begin{pmatrix} -1 & 2 \\ 2 & -1 \end{pmatrix}$ να ισχύει

$\|A\| = 2.5$

Φασματική ακτίνα του A;

Χαρακτηριστικό πολυώνυμο:

$p(\lambda) = \det(A - \lambda I_2) = \begin{vmatrix} -1-\lambda & +2 \\ 2 & -1-\lambda \end{vmatrix} = (-1-\lambda)^2 - 4 = (\lambda+1)^2 - 4$

Ιδιοτιμές: $p(\lambda) = 0 \Leftrightarrow (\lambda+1)^2 = 4 \Leftrightarrow \lambda + 1 = \pm 2 \Leftrightarrow \begin{cases} \lambda_1 = -3 \\ \lambda_2 = 1 \end{cases}$

$\rho(A) = \max(|\lambda_1|, |\lambda_2|) = 3$

Άρα $\rho(A) > \|A\|$ οπότε <u>δεν</u> υπάρχει τέτοια φυσική νόρμα

Άσκηση 3.32

a) $\|A\|_{max} = \max_{1 \leq i,j \leq n} |a_{ij}|$

$\|A\|_F = \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right)^{1/2}$ Είναι φυσικές νόρμες;

$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix} \longrightarrow \|A\|_{max} = 2$

$p(\lambda) = \begin{vmatrix} 1-\lambda & 2 \\ 2 & 1-\lambda \end{vmatrix} = (\lambda-1)^2 - 4$

$p(\lambda) = 0 \Leftrightarrow (\lambda-1)^2 = 4 \Leftrightarrow \lambda - 1 = \pm 2 \Leftrightarrow \begin{cases} \lambda_1 = 3 \\ \lambda_2 = -1 \end{cases}$

Φασματική ακτίνα του A : $\rho(A) = 3$

Επειδή $\rho(A) > \|A\|_{max}$ η $\|\cdot\|_{max}$ δεν είναι φυσική

$\|I_n\|_E = \sqrt{n} > 1$ για $n > 1$

Συμπέρασμα: Η $\|\cdot\|_E$ δεν είναι φυσική νόρμα.

Σημείωση

Για φυσικές νόρμες $\|I_n\| = \max\limits_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \dfrac{\|I_n x\|}{\|x\|} = \max\limits_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \dfrac{\|x\|}{\|x\|} = 1$

б) $\forall x \in \mathbb{R}^n \ \forall A \in \mathbb{R}^{n,n} \quad \|Ax\|_2 \leq \|A\|_E \cdot \|x\|_2$ (δηλαδή $\|A\|_E \geq \|A\|_2$)

Λύση:

$\|Ax\|_2^2 = \sum\limits_{i=1}^{n} ((Ax)_i)^2 = \sum\limits_{i=1}^{n} \Big(\underbrace{\sum\limits_{j=1}^{n} a_{ij} x_j}_{(Ax)_i} \Big)^2$

$\underset{CS}{\leq} \sum\limits_{i=1}^{n} \sum\limits_{j=1}^{n} |a_{ij}|^2 \sum\limits_{j=1}^{n} |x_j|^2 = \|x\|_2^2$

$= \Big(\sum\limits_{i,j=1}^{n} |a_{ij}|^2 \Big) \|x\|_2^2 = \|A\|_E^2 \cdot \|x\|_2^2 \Rightarrow \|Ax\|_2 \leq \|A\|_E \|x\|_2$

Άσκηση 3.36 (βιβλίου)

$\|\cdot\|$ νόρμα στον $\mathbb{R}^n$

φυσική νόρμα στον $\mathbb{R}^{n,n}$

$A \in \mathbb{R}^{n,n}$ $\|A\| < 1$.

Νδο α) $I_n - A$ αντιστρέψιμος

β) $\frac{1}{1+\|A\|} \leq \|(I_n-A)^{-1}\| \leq \frac{1}{1-\|A\|}$

Λύση:

α) $(I_n - A)x = 0 \Rightarrow x - Ax = 0 \Rightarrow x = Ax \Rightarrow \|x\| = \|Ax\| \leq \|A\| \cdot \|x\|$

$\|x\| \leq \|A\| \cdot \|x\|$

Για $x \neq 0$ η ανισότητα δίνει $1 \leq \|A\|$ ↯

Άρα $x = 0$, οπότε ο $I_n - A$ είναι αντιστρέψιμος

β) $\|(I_n-A)^{-1}(I_n-A)\| = 1$

• $1 \leq \|(I_n-A)^{-1}\| \cdot \underbrace{\|(I_n-A)\|}_{\leq \underbrace{\|I_n\|}_{1} + \|A\|}$

Άρα $1 \leq \|(I_n-A)^{-1}\| \leq 1 + \|A\| \Rightarrow \frac{1}{1+\|A\|} \leq \|(I_n-A)^{-1}\|$

• $\|(I_n-A)^{-1}\cdot(I_n-A)\| = 1 \Rightarrow \|(I_n-A)^{-1} - (I_n-A)^{-1}A\| \geq$

$\|(I_n-A)^{-1}\| - \underbrace{\|(I_n-A)^{-1}A\|}_{\leq \|(I_n-A)^{-1}\| \cdot \|A\|}$

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

$\|(I_n-A)^{-1}\| \leq \frac{1}{1-\|A\|}$

Άσκηση 3.64 (βιβλίου)

$Ax=b$

$$A=\begin{pmatrix}1 & 2 & -7/2\\ 1 & 1 & 7/4\\ 1 & 1 & 1\end{pmatrix}$$

- Jacobi: συγκλίνει
- Gauss-Seidel: γενικά αποκλίνει.

απόδειξη:

Jacobi

$\rightarrow D=I_3 \rightarrow D^{-1}=I_3$

$$G_J=-D^{-1}(L+U)=-(L+U)=-\begin{pmatrix}0 & 2 & -7/2\\ 1 & 0 & 7/4\\ 1 & 1 & 0\end{pmatrix}$$

χαρακτηριστικό πολυώνυμο: $p(\lambda)=\det(G_J-\lambda I_3)=-\lambda\left(\lambda^2-\frac{1}{4}\right)$

ιδιοτιμές: $p(\lambda)=0 \Rightarrow \begin{cases}\lambda_1=0\\ \lambda_2=1/2\\ \lambda_3=-1/2\end{cases}$

Φασματική ακτίνα: $\rho(G_J)=\frac{1}{2}<1$

Άρα η μέθοδος συγκλίνει.

Gauss-Siedel

$$G_{GS}=-(L+D)^{-1}U=\begin{pmatrix}1 & 0 & 0\\ 1 & 1 & 0\\ 1 & 1 & 1\end{pmatrix}^{-1}\begin{pmatrix}0 & 2 & -7/2\\ 0 & 0 & 7/4\\ 0 & 0 & 0\end{pmatrix}=$$

$$=\begin{pmatrix}1 & 0 & 0\\ -1 & 1 & 0\\ 0 & -1 & 1\end{pmatrix}\cdot\begin{pmatrix}0 & 2 & -7/2\\ 0 & 0 & 7/4\\ 0 & 0 & 0\end{pmatrix}=\begin{pmatrix}0 & -2 & 7/2\\ 0 & 2 & -21/4\\ 0 & 0 & 7/4\end{pmatrix}$$

Οι αριθμοί με το χρώμα πρέπει να υπολογιστούν

Ιδιοτιμές: $\lambda_1=0,\ \lambda_2=2,\ \lambda_3=7/4$

Φασματική ακτίνα: $\rho(G_{GS})=2\geq 1$

Η μέθοδος αποκλίνει γενικά.

Άσκηση 3.65 (βιβλίου)

$Ax=b$

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

<u>Ν.δ.ο.</u> • Gauss-Siedel: συγκλίνει
• Jacobi: γενικά αποκλίνει

απόδειξη:

<u>Gauss-Siedel</u>

$$G_{GS} = -(L+D)^{-1}U = \begin{pmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 0 & -1 & 1 \end{pmatrix} \cdot \begin{pmatrix} 0 & 1/2 & 1 \\ 0 & 0 & 1/3 \\ 0 & 0 & 0 \end{pmatrix} = \begin{pmatrix} 0 & -1/2 & -1 \\ 0 & 1/2 & 2/3 \\ 0 & 0 & 1/3 \end{pmatrix}$$

ιδιοτιμές: $\lambda_1=0$, $\lambda_2=1/2$, $\lambda_3=1/3$

Φασματική ακτίνα: $\rho(G_{GS}) = \frac{1}{2} < 1$

Η μέθοδος συγκλίνει.

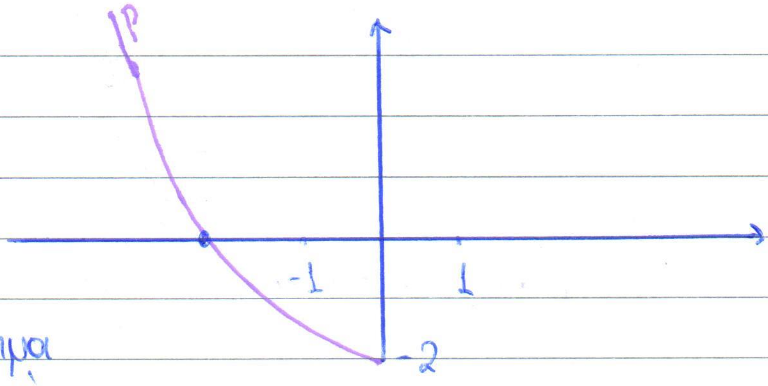
<u>Jacobi</u>

$$G_J = -D^{-1}(L+U) = -(L+U) = \begin{pmatrix} 0 & -1/2 & -1 \\ -1 & 0 & -1/3 \\ -1 & -1 & 0 \end{pmatrix}$$

χαρακτηριστικό πολυώνυμο: $p(\lambda) = \det(G_J - \lambda I_3) \Rightarrow p(\lambda) = -\lambda^3 + \frac{11}{6}\lambda - \frac{7}{6}$

Τώρα $p(-1) = -2$ και

$\lim_{\lambda\to-\infty} p(\lambda) = \infty$



Συμπέρασμα: Υπάρχει ιδιοτιμή του G_J στο διάστημα $(-\infty,-1)$. Άρα $\rho(G_J)>1$. Επομένως η μέθοδος αποκλίνει.