

• Γραμμικά συστήματα

Δεδομένα: $A \in \mathbb{R}^{n,n}$ $n \times n$ πίνακας

$b \in \mathbb{R}^n$ διάνυσμα με n συνιστώσες

Ζητούμενο: $x \in \mathbb{R}^n$ 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 των μεθόδων

β) Κατάσταση γραμμικών συστημάτων

- Υπάρχουν δύο μεγάλες κατηγορίες αριθμητικών μεθόδων:

α) Άμεσες (παράλλαγές της μεθόδου απαλοιφής του Gauss):
Όταν οι πράξεις γίνονται ακριβώς δίνουν τη λύση ακριβώς με πεπερασμένο πλήθος πράξεων.

β) Επαναληπτικές: Δίνουν μια ακολουθία προσεγγίσεων $x^{(0)}, x^{(1)}, x^{(2)}, \dots$ τις λύσεις x .

• Γενικά για γραμμικά συστήματα:

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

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

Ζητούμενο: $x_1, x_2, \dots, x_n \in \mathbb{R}$ τ.ω.

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

Με πίνακα,

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

και τα διανύσματα,

$$b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix} \text{ και } x = \begin{pmatrix} x_1 \\ \vdots \\ x_{n-1} \\ x_n \end{pmatrix}$$

το γραμμικό αυτό σύστημα γράφεται στη μορφή,

$$\boxed{Ax = b}$$

→ Θα ασχοληθούμε με γραμμικά συστήματα που έχουν ακριβώς μία λύση.

- Κάθε μια από τις ακόλουθες συνθήκες είναι 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) $Ax=0 \Rightarrow x=0$
- iv) Οι στήλες (ή οι γραμμές) του A είναι γραμμικά ανεξάρτητα διανύσματα.

- Τρόπος επίλυσης του $Ax=b$, γνωστοί από τη Γραμμική Άλγεβρα.

α) Κανόνας του Cramer:

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

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

Τότε,

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

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

! (Όχι $x = \frac{b}{A}$ ούτε $x = bA^{-1}$)!

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

α) Cramer: Αναπτύσσοντας ως προς στήλη j

έχουμε,

$$\det A = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det A_{ij}$$

με $A_{ij} \in \mathbb{R}^{n-1, n-1}$ τον πίνακα που προκύπτει από τον A αν διαγράψουμε τη γραμμή του i και τη στήλη του j .

- Συνεχίζοντας με τον ίδιο τρόπο γράφουμε την $\det A$ ως άθροισμα $n!$ όρων με n παράγοντες ο καθένας. Συνολικά απαιτούνται $n!(n-1)$ πολλαπλασιασμοί.
- Στον κανόνα του Cramer απαιτείται ο υπολογισμός $n+1$ οριζουσών $n \times n$ πινάκων.

Συνολικά απαιτούνται,

$(n+1)(n!(n-1))$ πολλαπλασιασμοί.

- Αυτός ο αριθμός αυξάνεται ταχύτατα με το n και καθιστά αυτή τη διαδικασία μη πρακτική.

(Υπάρχουν άλλοι τρόποι για τον υπολογισμό οριζουσών που απαιτούν λιγότερες πράξεις αλλά και με αυτούς η μέθοδος είναι ασύμφωνη).

• Γραμμικά συστήματα :

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

b) $x = A^{-1}b$

- Σπάνια χρειάζεται να υπολογίσουμε τον A^{-1} . Όταν αυτό χρειάζεται, γίνεται ως εξής. Έστω e^1, \dots, e^n η κανονική βάση του του \mathbb{R}^n ,

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

δ : σύμβολο του Kronecker ή δέλτα του Kronecker

- Έστω $u^i \in \mathbb{R}^n$ τ.ω. $Au^i = e^i, i=1, \dots, n$

• Ισχυρισμός :

$$A^{-1} = (u^1, u^2, \dots, u^n)$$

Πράγματι,

$$\begin{aligned} A(u^1, u^2, \dots, u^n) &= (Au^1, Au^2, \dots, Au^n) = \\ &= (e^1, e^2, \dots, e^n) = I_n \end{aligned}$$

$$\Rightarrow A^{-1} = (u^1, \dots, u^n)$$

→ Για να υπολογίσουμε τον A^{-1} πρέπει να λύσουμε η γραμμικά συστήματα με πίνακα συντελεστών τον A .

Επιπλέον κόστος : $x = \underbrace{A^{-1}}_{\text{πολλ. πίνακα}} b$ επί διάνυσμα

→ Υπερβολικά μεγάλο κόστος!

• Δύο μεγάλες κατηγορίες γραμμικών συστημάτων (πινάκων):

α) Πυκνοί (ή αποθηκεύσιμοι) πίνακες. Έχουν στοιχεία a_{ij} γενικά διαφορετικά του μηδενός.

β) Αραιοί (ή σποραδικοί) πίνακες: Έχουν πολλά μηδενικά στοιχεία που αν τα εκμεταλλευτούμε αποκομίζουμε υπολογιστικά οφέλη.

• Μέγεθος πινάκων (τετραγωνικών) $A \in \mathbb{R}^{n,n}$:

$n < 100$: μικροί

$100 \leq n < 1000$: μεσαίοι (μέτριοι)

$n \geq 1000$: μεγάλοι

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

• Δύο μεγάλες κατηγορίες αριθμητικών μεθόδων:

α) Άμεσες: χρησιμοποιούνται κυρίως για πυκνούς πίνακες.

β) Επαναληπτικές: χρησιμοποιούνται κυρίως για αραιούς πίνακες.

- Οι άμεσες μέθοδοι είναι παράλλαξες της μεθόδου απαλοιφής του Gauss.
- Οι άμεσες μέθοδοι δίνουν τη λύση ακριβώς (αν οι πράξεις γίνονται ακριβώς) με πεπερασμένο πλήθος πράξεων.
- Οι επαναληπτικές μέθοδοι δίνουν μια ακολουθία προσεγγίσεων της λύσης.

• Η μέθοδος απαλοιφής του Gauss:

- Ειδική Περίπτωση:

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

Έχουμε,

$$\det u = u_{11} \cdot u_{22} \cdot \dots \cdot u_{nn},$$

δηλαδή, u αντιστρέψιμος (\Rightarrow) $u_{ii} \neq 0$, $i = 1, 2, 3, \dots, n$.

- Έστω $y \in \mathbb{R}^n$. Το γραμμικό σύστημα $u x = y$, δηλαδή

$$\left. \begin{array}{rcl} u_{11}x_1 + u_{12}x_2 + \dots + u_{1n}x_n & = & y_1 \\ u_{22}x_2 + \dots + u_{2n}x_n & = & y_2 \\ \vdots & & \vdots \\ u_{n-1,n-1}x_{n-1} + u_{n-1,n}x_n & = & y_{n-1} \\ + u_{nn}x_n & = & y_n \end{array} \right\} \begin{array}{l} \text{λύνεται} \\ \text{εύκολα με} \\ \text{οπισθοδρόμηση} \end{array}$$

• Οπισθοδρόμηση: Λύνουμε την τελευταία εξίσωση ως προς x_n , αντικαθιστούμε στην προηγούμενη εξίσωση και βρίσκουμε το x_{n-1} , συνεχίζοντας όμοια στο τέλος αντικαθιστούμε τα x_n, \dots, x_2 στην πρώτη εξίσωση και βρίσκουμε το x_1 .

• Αλγόριθμος της οπισθοδρόμησης:

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

Για $k = n-1, n-2, \dots, 1$:

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

Έχουμε, $\sum_{j=1}^n u_{kj} x_j = \sum_{j=k}^n u_{kj} x_j = y_k \Rightarrow$
άνω τριγωνικός

$\Rightarrow u_{kk} \cdot x_k + \sum_{j=k+1}^n u_{kj} x_j = y_k$, πηγαίνοντας το άθροισμα στο δεξιά μέλος της εξίσωσης και διαιρώντας με u_{kk} προκύπτει η (*).

• Απαιτούμενες πράξεις:

Διαιρέσεις: n

Πολλαπλασιασμοί και προσθέσεις:

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

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

• Θέσεις μνήμης:

$$\frac{n^2}{2} + o(n) \quad \text{για τα } u_{ij} \text{ με } j \geq i \text{ και } y_i$$

Τα x_i αποθηκεύονται στις θέσεις των y_i .

• Ιδέα στη μέθοδο απαλοιφής του Gauss. Με κατάλληλους μετασχηματισμούς γραμμών (δηλ. εναλλαγή δύο εξισώσεων ή πρόσθεση κατά μέλη ενός πολλαπλάσιου μιας εξίσωσης σε κάποια άλλη) μετατρέπουμε το $Ax=b$ σε ένα ισοδύναμο $U \cdot x = y$ με άνω τριγωνικό πίνακα U . Αυτή η διαδικασία λέγεται τριγωνοποίηση.

Το $Ux=y$ λύνεται με οπισθοδρόμηση.

• Μέθοδος απαλοιφής του Gauss

Γενική περίπτωση:

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

$$\text{Θέτουμε } A^{(1)} = (a_{ij}^{(1)}) = A$$

$$\text{και } b^{(1)} = \begin{pmatrix} b_1^{(1)} \\ \vdots \\ b_n^{(1)} \end{pmatrix} = b$$

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

• Τριγωνοποίηση:

1^ο Βήμα: Υπόθεση: $a_{11}^{(1)} \neq 0$ (μπορεί να επιτευχθεί με εναλλαγή γραμμών).

Πολλαπλασιαστές :

$$m_{i1} = \frac{a_{i1}^{(1)}}{a_{11}^{(1)}}, i=2, \dots, n$$

↳ λόγω 1ου βήματος

- Πολλαπλασιάζουμε την πρώτη εξίσωση επί m_{i1} , αφαιρούμε από την i -οστή εξίσωση και αντικαθιστούμε την i -οστή εξίσωση με αυτή που προέκυψε, για $i=2, \dots, n$.

Έτσι μετά το πρώτο βήμα τις τριγωνοποιήσεις το γραμμικό σύστημα γίνεται,

$$A^{(2)}x = b^{(2)} \text{ με,}$$

$$A^{(2)} = \begin{pmatrix} a_{11}^{(1)} & a_{12}^{(1)} & \dots & a_{1n}^{(1)} \\ 0 & a_{22}^{(2)} & \dots & a_{2n}^{(2)} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & a_{2n}^{(2)} & \dots & a_{nn}^{(2)} \end{pmatrix}, \quad b^{(2)} = \begin{pmatrix} b_1^{(1)} \\ b_2^{(2)} \\ b_3^{(2)} \\ \vdots \\ b_n^{(2)} \end{pmatrix}$$

με,

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

$$b_i^{(2)} = b_i^{(1)} - m_{i1} b_1^{(1)}, \quad i = 2, \dots, n$$

Βήμα r : $1 \leq r \leq n-1$

Ξεκινάμε από το σύστημα

$$A^{(r)}x = b^{(r)}$$

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

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

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

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

- Υπόθεση : $a_{rr}^{(r)} \neq 0$ (μπορεί να επιτευχθεί με εναλλαγή της εξίσωσης r με μια εξίσωση i με $i > r$).

Πολλαπλασιαστές :

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

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

→ Έτσι προκύπτει το γραμμικό σύστημα $A^{(r+1)}x = b^{(r+1)}$ με τις πρώτες r γραμμές του $A^{(r+1)}$ ίδιες με τις αντίστοιχες του $A^{(r)}$ (και αντίστοιχα για τις συνιστώσες του $b^{(r+1)}$)

και

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

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

και

$$a_{ij}^{(r+1)} = 0 \text{ για } \begin{matrix} i = r+1, \dots, n \\ j = 1, \dots, r \end{matrix}$$

→ Υστερα από $n-1$ βήματα παίρνουμε το γραμμικό σύστημα,

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

→ Εδώ τελειώνει η τριγωνοποίηση. Το \otimes το λύνουμε με οπισθοδρόμηση.

• Απαιτούμενες πράξεις και μνήμη

Το πλήθος των προσθαφαιρέσεων είναι περίπου ίσο με το πλήθος των πολλαπλασιασμών.

Έχει επικρατήσει να μετράμε ως αριθμητικές πράξεις μόνο πολλαπλασιασμούς και διαιρέσεις.

Τριγωνοποίηση :

1^ο 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 $(n-1)^2$

Βήμα r : $(n-r)$ πράξεις για τους πολλαπλασιαστές και $(n-r)^2$ πράξεις για τα νέα στοιχεία.

→ Για τον υπολογισμό του πίνακα A απαιτούνται :

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

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

Για το b απαιτούνται $n-1$ πράξεις στο βήμα i ,
οπότε συνολικά,

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

→ Συνολικά μαζί με την οπισθοδρόμηση απαιτούνται,

$$\frac{n^3 + 3n^2 - n}{3} \approx \frac{n^3}{3} \text{ πράξεις.}$$

18/04/2018

Παράδειγμα

$$Ax=b, A \in \mathbb{R}^{n,n}, n=20$$

Έστω ότι ο υπολογιστής κάνει 10^6 πράξεις/sec.

Με τη μέθοδο απαλοιφής του Gauss απαιτούνται $\frac{16}{3} \cdot 10^3$ sec

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

Απαιτούμενη μνήμη στη μέθοδο απαλοιφής του Gauss

Για τον πίνακα A : n^2 θέσεις μνήμης

Για το b : n θέσεις μνήμης

→ Δεν απαιτείται επιπλέον μνήμη.

Οι πολλαπλασιαστές $m_{ir}, i=r+1, \dots, n$ αποθηκεύονται στις θέσεις των στοιχείων $a_{ir}, i=r+1, \dots, n$. Οι πολλαπλασιαστές αποθηκεύονται δηλαδή τελικά στις θέσεις (i,j) με $i > j$ του πίνακα. Τα νέα στοιχεία κάθε φορά υπολογίζονται από τους τύπους,

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

(Τα m_{ir} παίρνουν τα ονόματα των αντιστοιχών a_{ir})

• Παρατήρηση (Υπολογισμός ορίζουσας)

$\det A = (-1)^m \det A^{(n)}$ όπου m το πλήθος των εναλλαγών γραμμών κατά την αναδιοργάνωση.

Άρα, $\det A = (-1)^m a_{11}^{(1)} a_{22}^{(2)} \dots a_{nn}^{(n)}$

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

- Επίλυση $n \times n$ γραμμικού συστήματος με τον κανόνα του Cramer:

Πράξεις : $(n+1) \left(\frac{n^3}{3} + O(n) \right) \approx \frac{n^4}{3}$ πράξεις

Η μέθοδος είναι ασύμφωνη!

• Οδύνηση :

Παράδειγμα : $\left. \begin{matrix} 10^{-4} x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{matrix} \right\} (*)$ (οδηγός)

Ακριβής λύση : $x_1 = 1.0001 \dots$
 $x_2 = 0.9998 \dots$

Υπολογιστής : $b=10, t=3, L=-20$
 $a=20$, στρογγύλευση

- Με τη μέθοδο αναδιοργάνωσης του Gauss παίρνουμε την προσεγγιστική λύση,

$\left. \begin{matrix} \tilde{x}_1 = 0 \\ \tilde{x}_2 = 1 \end{matrix} \right\}$ πολύ κακή!

Εξήγηση

$$m_{21} = \frac{1}{10^{-4}} = 10^4 \text{ αριθμός μηχανής}$$

$$a_{22}^{(2)} = fl(a_{22}^{(1)} - m_{21}a_{12}^{(1)}) = fl(1 - 10^4) = -10^4$$

→ η μονάδα εξαφανίστηκε λόγω του μεγέθους του 10^4 .

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

Έτσι προκύπτει το γραμμικό σύστημα,

$$\left. \begin{array}{l} 10^{-4}x_1 + x_2 = 1 \\ -10^4x_2 = -10^4 \end{array} \right\} \Rightarrow \left. \begin{array}{l} x_1 = 0 \\ x_2 = 1 \end{array} \right\}$$

Το πρόβλημα οφείλεται στον πολύ μικρό οδικό 10^{-4} που μας έδωσε τον πολύ μεγάλο πολλαπλασιαστή 10^4 .

Εναλλακτικός τρόπος:

Εναλλαγή των δύο εξισώσεων,

$$\left. \begin{array}{l} x_1 + x_2 = 2 \\ 10^{-4}x_1 + x_2 = 1 \end{array} \right\}$$

Με τη μέθοδο αναλοιφής του Gauss στον υπολογιστή μας παίρνουμε την προσεγγιστική λύση,

$$\left. \begin{array}{l} x_1 = 1 \\ x_2 = 1 \end{array} \right\} \text{ Πολύ καλά!}$$

• Μέθοδος αναλοίφης του Gauss με μερική οδύχηση (ή οδύχηση κατά γραμμές):

- Στο βήμα r τις τριγωνοποιήσεις εξετάζουμε τα στοιχεία $a_{kr}^{(r)}$, $k=r, r+1, \dots, n$, τις στήλες r του πίνακα $A^{(r)}$, βρίσκουμε ένα από αυτά τα στοιχεία με τη μεγαλύτερη απόλυτη τιμή, και με εναλλαγή γραμμών το φέρνουμε στη θέση του οδύχου (δηλαδή στη θέση (r, r)).

- Επιπλέον κόστος: $\frac{n^2}{2} - \frac{n}{2}$

• Παρατήρηση: Ο όρος "πολύ μικρός οδύχος" είναι ασαφής. Το \otimes γράφεται στη μορφή,

$$\text{ίδιοι αριθμοί} \left\{ \begin{array}{l} 1 \rightarrow x_1 + 10^4 x_2 = 10^4 \\ 1 \rightarrow x_1 + x_2 = 2 \end{array} \right.$$

- Αναλοίφη Gauss:

$$\text{Δίνει } \left. \begin{array}{l} \tilde{x}_1 = 0 \\ \tilde{x}_2 = 1 \end{array} \right\}$$

και (με εναλλαγή)

$$\left. \begin{array}{l} \tilde{x}_1 = 1 \\ \tilde{x}_2 = 1 \end{array} \right\}$$

χωρίς εναλλαγή γραμμών

παρά το γεγονός ότι οι οδύχοι και στις δύο περιπτώσεις είναι ίσοι με ένα!

→ Ο πραγματικός λόγος της αποτυχίας στην πρώτη περίπτωση είναι ότι το μέγεθος του $|a_{11}|$ είναι πολύ μικρό σε σύγκριση με το $|a_{12}|$, $\left(\frac{|a_{11}|}{|a_{12}|} = 10^{-4}\right)$

$$\text{ενώ } \frac{|a_{21}|}{|a_{22}|} = 1$$

Αυτό μας υποδεικνύει την αλλαγή γραμμών.

• Οδύνηση κατά γραμμές και στήλες ή ολική οδύνηση

Στο βήμα r τις τριγωνοποιήσεις με εναλλαγές και στήλων φέρνουμε στη θέση του οδύνη $a_{rr}^{(r)}$ ένα από τα στοιχεία $a_{kr}^{(r)}$, $k, \ell = r, \dots, n$, του υποπίνακα $\tilde{A}^{(r)}$, με τη μέγιστη απόλυτη τιμή.

Προσοχή! Η εναλλαγή δύο στηλών r και s απαιτεί τη μετονομασία των αντίστοιχων αγνώστων x_r σε x_s και x_s σε x_r .

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

- Η μέθοδος απαλοιφής του Gauss χωρίς οδύνηση θεωρείται ασταθής αλγόριθμος.
Για κάποιες κατηγορίες συστημάτων, π.χ. αν ο πίνακας A είναι θετικά ορισμένος, δηλαδή,

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

- Η μέθοδος ολικής οδύνησης θεωρείται στην πράξη ευσταθής αλγόριθμος. Χρησιμοποιείται σπάνια γιατί διπλασιάζει το κόστος της απαλοιφής.

- Η πιθανότητα να είναι ασταθής η μέθοδος απαλοιφής του Gauss με μερική οδήγηση είναι πολύ μικρή. Χρησιμοποιείται πολύ συχνά στην πράξη γιατί αυξάνει μόνο λίγο το κόστος

• Ο αλγόριθμος της απαλοιφής του Gauss στην πράξη

Η μέθοδος απαλοιφής του Gauss για την επίλυση του γραμμικού συστήματος $Ax=b$, γράφεται στην πράξη σε δύο στάδια:

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

- Ένα πολύ γνωστό ζεύγος υποπρογραμμάτων που υλοποιεί αυτά τα δύο στάδια είναι τα `DECOMP` και `SOLVE`.

DECOMP: Υλοποιεί το πρώτο στάδιο. Έχει ως είσοδο τον A τα στοιχεία του οποίου αποθηκεύονται σε n^2 θέσεις. Εκτελεί απαλοιφή Gauss με μερική οδήγηση. Το τελικό προϊόν του A αποθηκεύεται στο πάνω δεξιά "τρίγωνο" του αρχικού πίνακα. Κάτω από τη διαγώνιο αποθηκεύονται οι πολλαπλασιαστές. Καταγράφονται επίσης οι εναλλαγές γραμμών (σε ένα διάνυσμα με n συνιστώσες).

Κόστος : $\frac{n^3}{3} + O(n^2)$ πράξεις

$n^2 + O(n)$ μνήμη

SOLVE : Χρησιμοποιεί ως είσοδο τα αποτελέσματα τις DECOMP και το b . Με τις πληροφορίες για τους πολλαπλασιαστές και για τις εναλλαγές γραμμών εκτελεί πρώτα τις πράξεις τις απαλοιφής που αφορούν το b και στη συνέχεια με οπισθοδρόμηση υπολογίζει το x . Το x αποθηκεύεται στις θέσεις του b .

Κόστος : $n^2 + O(n)$ πράξεις

n : θέσεις μνήμης

• Εφαρμογή : Έστω ότι θέλουμε να λύσουμε m γραμμικά συστήματα με τον ίδιο πίνακα A και δεξιά μέλη b, b', b'', \dots

Καλούμε μία φορά τη DECOMP και m φορές τη SOLVE, μία φορά για κάθε δεξιά μέλος.

Συνολικό κόστος :

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

- Παράδειγμα : Υπολογισμός του A^{-1}

Λύουμε τα n συστήματα

$$A u^i = e^i, \quad i = 1, \dots, n$$

με e^1, \dots, e^n την κανονική βάση του \mathbb{R}^n και έχουμε $A^{-1} = (u^1, u^2, \dots, u^n)$

Συνολικό κόστος:

$$\frac{n^3}{3} + o(n^2) + n \left(\frac{2}{3} n^2 + o(n) \right) = n^3 + o(n^2) \text{ πράξεις}$$

↑
λόγω πολλών μηδενικών των e^i

θέσεις μνήμης: $2n^2 + o(n)$

• Η ανάλυση LU

Το στάδιο της τριγωνοποίησης ενός αντιστρέψιμου πίνακα A κατά την αναδοιφή Gauss, πιθανώς και με εναλλαγές γραμμών, μπορεί να εκφραστεί σε γλώσσα πινάκων ως ανάλυση του A σε γινόμενο,

$$A = P^{-1} \cdot L \cdot U, \text{ με } P, L, U \in \mathbb{R}^{n,n} \text{ πίνακες με τις ακόλουθες ιδιότητες.}$$

- Ο P είναι ένας πίνακας μετάθεσης.
Καταγράφει τις εναλλαγές γραμμών κατά την αναδοιφή.
(Ένας πίνακας μετάθεσης προκύπτει από τον μονοδιαίο πίνακα I_n με εναλλαγές γραμμών).
- Ο L είναι κάτω τριγωνικός και έχει μονάδες στη διαγώνιο. Κάτω από τη διαγώνιο έχει τους πολλαπλασιαστές.
- Ο U είναι άνω τριγωνικός, περιέχει το τελικό προϊόν $A^{(n)}$ της τριγωνοποίησης.

• Χρήση της ανάλυσης LU :

$$Ax = b \Leftrightarrow PAx = Pb$$
$$PA = LU$$

α) Ανάλυση του PA σε γινόμενο LU

β) Υπολογισμός του x :

β₁) Κατασκευή του Pb (με εναλλαγή των συνιστωσών του b)

$$LUx = Pb$$

y

Υπολογισμός του y από τη σχέση $Ly = Pb$
(Με τρόπο αντίστοιχο της οπισθοδρόμησης)

β₂) $Ux = y$
Λύνεται με οπισθοδρόμηση.

* Η φάση α) αντιστοιχεί στην DECOMP και η φάση β) στην SOLVE.

• Ανάλυση LU :

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

Απόδειξη ανάλυσης :

$$A = P^{-1}LU$$

1^η Περίπτωση : υποθέτουμε ότι δεν γίνονται εναλλαγές γραμμών κατά την απαλοιφή Gauss ($P = I_n$)

Ορίζουμε τον πίνακα M_1 ,

$$M_1 = \begin{pmatrix} 1 & 0 & \dots & 0 \\ -m_{21} & 1 & & \\ -m_{31} & 0 & \dots & \\ \vdots & \vdots & \ddots & \vdots \\ -m_{n1} & 0 & \dots & 0 & 1 \end{pmatrix}$$

και παρατηρούμε ότι $A^{(2)} = M_1 A$
δηλαδή ότι το πρώτο βήμα στην τριγωνοποίηση του A μπορεί να ερμηνευθεί ως πολλαπλασιασμός του A από αριστερά με τον πίνακα M_1 (m_{i1} , $i=2, \dots, n$ είναι οι πολλαπλασιαστές του πρώτου βήματος της απαλοιφής.)

- Συνεχίζοντας αντίστοιχα βλέπουμε ότι
 $A^{(n)} = \underbrace{M_{n-1} M_{n-2} \dots M_1}_{} A$

με $M_r \in \mathbb{R}^{n,n}$, $r=1, \dots, n-1$ που δίνονται ως εξής,

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

Δηλαδή, ο M_r είναι κάτω τριγωνικός πίνακας με στοιχεία $-m_{ir}$, $i=r+1, \dots, n$ στην υποστήλη r κάτω από τη διαγώνιο και μηδέν όλα τα άλλα στοιχεία. Προφανώς, $\det M_r = 1$, οπότε ο M_r είναι αντιστρέψιμος. Μάλιστα (όπως θα δούμε στην Άσκηση 3.6)

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

Επομένως,

$$A = \underbrace{\left(M_1^{-1} \cdot \dots \cdot M_{n-2}^{-1} \cdot M_{n-1}^{-1} \right)}_{\text{Προσοχή στη σειρά}} A^{(n)} = LU$$

Τώρα (Άσκηση 3.6),

$$M_1^{-1} \cdot M_2^{-1} \cdot \dots \cdot M_{n-1}^{-1} = \begin{pmatrix} 1 & & & & \\ m_{21} & 1 & & & 0 \\ & m_{32} & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ \vdots & \vdots & & \ddots & 1 \\ m_{n1} & m_{n2} & \dots & m_{n,n-1} & 1 \end{pmatrix} = L$$

- Στι σχέση $A=LU$ οδηγείται κανείς κάνοντας ακριβώς τις ίδιες πράξεις που γίνονται στην τριγωνοποίηση του A με την αναλοϊφή Gauss.

2^η Περίπτωση: Κατά την απαλοιφή Gauss γίνονται εναλλαγές γραμμών είτε για να βρούμε μη μηδενικό οδύο είτε λόγω μερικής οδύοισης.

Η γραμμή την οποία στο βήμα i φέρνουμε στη θέση του οδύο (δηλαδή αυτή που εναλλάσσεται με τη γραμμή i) ούτε αλλάζει πλέον θέση ούτε τα στοιχεία της αλλοιώνονται στη συνέχεια τις απαλοιφής.

Επομένως, υπάρχει μια μετάθεση των γραμμών του A που αν την κάνουμε πριν από το πρώτο βήμα τις απαλοιφής θα οδύο σε έναν πίνακα A' , η τριγωνοποίηση του οποίου θα μπορούσε να γίνει χωρίς εναλλαγές γραμμών.

- Επομένως, σύμφωνα με την 1^η Περίπτωση θα έχουμε,

$$(*) \quad \underline{A' = LU}$$

με L, U με τις επιθυμητές ιδιότητες.

- Θα δούμε τώρα ότι,

$$(**) \quad \underline{A' = PA} \quad \text{με ένα κατάλληλο πίνακα μετάθεσης.}$$

Υποθέτοντας ότι η **(**)** ισχύει, συμπαιρένουμε ότι

$$PA = LU \Rightarrow A = P^{-1}LU$$

→ Έστω τώρα,

$$P = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$$

μία μετάθεση των αριθμών $1, 2, \dots, n$. Ο αντίστοιχος πίνακας μετάθεσης P προκύπτει από τον μοναδιαίο πίνακα I_n , αν

μεταθέσουμε τις γραμμές του κατά τη μετάθεση π δηλαδή η i -ε γραμμή P είναι k γραμμή του I_n .
 Π.χ.

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

↑
αντίστοιχος πίνακας μετάθεσης

Η 3^η γραμμή του P είναι η 1^η γραμμή του I_3
 Η 1^η γραμμή του P είναι η 2^η γραμμή του I_3
 Η 2^η γραμμή του P είναι η 3^η γραμμή του I_3

- Τώρα, το γινόμενο PA προκύπτει από τον A με τις ίδιες ακριβώς εναλλαγές γραμμών που από τον I_n οδηγούν στον P .

Π.χ. (για τον $P \in \mathbb{R}^{3,3}$ του παραδείγματος)

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

• Συμπέρασμα : Στον πίνακα A' στη δεύτερη περίπτωση οδηγούμαστε πολλαπλασιάζοντας τον A από αριστερά με έναν κατάλληλο πίνακα μετάθεσης P , δηλαδή

$$A' = PA$$

Άρα ισχύει η ****** και έχουμε το επιθυμητό αποτέλεσμα.

Παραδείγματα:

$$1. \quad A = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix}$$

Αναλοισμή Gauss:

$$L = \begin{pmatrix} 1 & 0 \\ 1/2 & 1 \end{pmatrix} \quad U = \begin{pmatrix} 2 & 3 \\ 0 & 5/2 \end{pmatrix} \quad m_{21} = 1/2$$

Επαλήθευση:

$$L \cdot U = \begin{pmatrix} 1 & 0 \\ 1/2 & 1 \end{pmatrix} \cdot \begin{pmatrix} 2 & 3 \\ 0 & 5/2 \end{pmatrix} = \begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} = A$$

$$2. \quad A = \begin{pmatrix} 2 & 4 & 5 \\ 1 & 2 & 3 \\ 2 & 5 & 9 \end{pmatrix}$$

Αναλοισμή Gauss:

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

Άρα,

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

→ οδηγός

→ Δεν μπορούμε να προχωρήσουμε χωρίς εναλλαγές γραμμών λόγω μηδενικού οδηγού!

Άρα,

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

$$PA = \begin{pmatrix} 2 & 4 & 5 \\ 2 & 5 & 9 \\ 1 & 2 & 3 \end{pmatrix} = B$$

Αναλοισφή Gauss στον B:

$$m_{21} = 1, m_{31} = 1/2$$

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

προέκυψε άνω τριγωνικός
πίνακας άρα,

$$B^{(2)} = U$$

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

→ Ισχύει ότι, $PA = LU \Rightarrow A = P^{-1}LU$

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

↑
P ορθογώνιος
πίνακας

• Κατάσταση γραμμικών συστημάτων

Παράδειγμα :

$$\underbrace{\begin{pmatrix} .913 & .659 \\ .780 & .563 \end{pmatrix}}_A \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} .254 \\ .217 \end{pmatrix}$$

Λύση $x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$

Υπολογιστής : $b=10$, $t=3$, αποκοπή

Αναλομφή Gauss (χωρίς οβήχισι, μερική οβήχισι, ολική οβήχισι)

• Προσέγγισι : $\tilde{x} = \begin{pmatrix} -.443 \\ 1 \end{pmatrix} !!$

→ Ερώτημα : Το άσχημο αποτέλεσμα οφείζεται στη μέθοδο ή μήπως στο γραμμικό σύστημα;

$$A \begin{pmatrix} y_1 \\ y_2 \end{pmatrix} = \begin{pmatrix} .253 \\ .218 \end{pmatrix} \rightarrow \text{Λύση} : y = \begin{pmatrix} 1223 \\ -1694 \end{pmatrix}$$

Το σύστημα έχει πολύ κακή κατάσταση!

Παρατήρηση: $\det A = -10^{-6}$

→ Η κατάσταση ενός πίνακα δεν έχει σχέση με την ορίζουσα του!

- Χρειαζόμαστε γενίκευση της έννοιας της απόλυτης τιμής για διανύσματα και πίνακες.

• Νόρμες διανυσμάτων και πινάκων

Οι νόρμες αποτελούν γενίκευση της απόλυτης τιμής.

• Νόρμες διανυσμάτων

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

Μια απεικόνιση,

$$\|\cdot\| : X \rightarrow \mathbb{R}, \quad x \mapsto \|x\|,$$

λέγεται νόρμα (στάθμη, norm), αν ισχύουν:

$$(N_1) \quad x \in X, \quad \|x\| = 0 \Leftrightarrow x = 0$$

$$(N_2) \quad \forall \lambda \in \mathbb{K} \quad \forall x \in X, \quad \|\lambda x\| = |\lambda| \cdot \|x\|$$

$$(N_3) \quad \forall x, y \in X, \quad \|x+y\| \leq \|x\| + \|y\|$$

(τριγωνική ανισότητα)

Παρατηρήσεις:

$$i. \forall x \in X, \|x\| \geq 0$$

Απόδειξη: $\|x-x\| = 0$ (=)

$$\|x-x\| = \|x + (-x)\| \stackrel{(N_3)}{\leq} \|x\| + \|-x\| =$$

$$= \|x\| + |-1| \cdot \|x\| \stackrel{(N_2)}{=} \|x\| + \|x\| =$$

$$= 2\|x\|$$

$$ii. \forall x, y \in X \quad | \|x\| - \|y\| | \leq \|x-y\|$$

τριγωνική ανισότητα προς τα κάτω.

Απόδειξη: $\|x\| = \|(x-y) + y\| \leq \|x-y\| + \|y\| \Rightarrow$

$$\textcircled{1} \quad \|x\| - \|y\| \leq \|x-y\|$$

$$\textcircled{2} \quad \|y\| - \|x\| \leq \|y-x\| = \|x-y\|$$

$$\textcircled{1}, \textcircled{2} \Rightarrow | \|x\| - \|y\| | \leq \|x-y\|$$

Παραδείγματα:

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

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

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

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

Πράγματι, ισχύουν

$$\rightarrow H(N_1) \text{ γιατί: } x \in \mathbb{R}^n \quad \|x\|_1 = 0 \Leftrightarrow |x_1| + |x_2| + \dots + |x_n| = 0 \Leftrightarrow$$

$$\Leftrightarrow x_i = 0 \quad \forall i = 1, \dots, n \Leftrightarrow x = 0$$

$$\rightarrow H(N_2) \text{ γιατί: } \lambda \in \mathbb{R}, x \in \mathbb{R}^n: \|\lambda x\|_1 = \sum_{i=1}^n |\lambda x_i| =$$

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

$$\rightarrow H(N_3) \text{ γιατί: } x, y \in \mathbb{R}^n$$

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

Νόρμες διανυσμάτων

- (N₁) $x \in X, \|x\| = 0 \Leftrightarrow x = 0$
 (N₂) $\forall \lambda \in \mathbb{K}, \forall x \in X, \|\lambda x\| = |\lambda| \|x\|$
 (N₃) $\forall x, y \in X, \|x+y\| \leq \|x\| + \|y\|$

Παραδείγματα:

2. $(\mathbb{R}^n, \|\cdot\|_1), \|x\|_1 = \sum_{i=1}^n |x_i|$ (l_1 -νόρμα)

3. $(\mathbb{R}^n, \|\cdot\|_\infty), \|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$ (l_∞ -νόρμα ή νόρμα
 μέγιστου)
 - Η απόδειξη ότι η $\|\cdot\|_\infty$
 είναι νόρμα είναι εύκολη.

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

- Η συγκεκριμένη νόρμα έχει
 την ιδιότητα να συνδέεται
 με ένα εσωτερικό γινόμενο
 το Ευκλείδειο εσωτερικό
 γινόμενο το οποίο την παρά-
 γει και την κάνει εύχρηστη.

Γενικά αν $1 \leq p < \infty$

$$\|x\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p}$$

$$\forall x \in \mathbb{R}^n, \|x\|_p \rightarrow \|x\|_\infty, p \rightarrow \infty$$

- Ευκλείδειο εσωτερικό γινόμενο:

$$(x, y)_2 = x_1 y_1 + x_2 y_2 + \dots + x_n y_n$$

$$(x, x)_2 = |x_1|^2 + \dots + |x_n|^2 = \|x\|_2^2$$

$$\Rightarrow \|x\|_2 = \sqrt{(x, x)}$$

(N₁), (N₂) πολύ εύκολα

(N₃): Ανισότητα των Cauchy-Schwarz

(*) $\forall x, y \in \mathbb{R}^n, |(x, y)_2| \leq \|x\|_2 \cdot \|y\|_2$

Απόδειξη:

• Για $y=0$ η (*) είναι προφανής

• Έστω $y \neq 0$. Για $\lambda \in \mathbb{R}$ έχουμε

$(a, a) \geq 0$

$$0 \leq (x + \lambda y, x + \lambda y)_2 = \|x\|_2^2 + 2\lambda(x, y)_2 + \lambda^2 \|y\|_2^2$$

$\phi(\lambda)$

$\neq 0$

$$ax^2 + bx + \gamma$$
$$\Delta = b^2 - 4a\gamma$$

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

$$\Rightarrow 4(x, y)_2^2 \leq 4\|y\|_2^2 \cdot \|x\|_2^2$$

$$\Rightarrow |(x, y)_2| \leq \|x\|_2 \cdot \|y\|_2$$

Τώρα, για $x, y \in \mathbb{R}^n$

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

$$\leq \|x\|_2^2 + 2\|x\|_2 \cdot \|y\|_2 + \|y\|_2^2 =$$

CS \nearrow

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

$$\Rightarrow \|x+y\|_2 \leq \|x\|_2 + \|y\|_2$$

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

- Οι νόρμες δίνονται από τους ίδιους τύπους όπως προηγουμένως

$$(\text{Εσωτερικό γινόμενο: } (x, y)_2 = \sum_{i=1}^n x_i \bar{y}_i)$$

6. $(C[a, b], \|\cdot\|_\infty)$, $\|f\|_\infty = \max_{a \leq x \leq b} |f(x)|$

$$(N_1): f \in C[a, b], \|f\|_\infty = 0 \Leftrightarrow$$

$$\max_{a \leq x \leq b} |f(x)| = 0 \Leftrightarrow f(x) = 0 \quad \forall x \in [a, b]$$

$$\Leftrightarrow f = 0$$

$$(N_2): \lambda \in \mathbb{R}, f \in C[a, b]$$

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

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

$$(N_3): f, g \in C[a, b]$$

$$\begin{aligned} \|f+g\|_\infty &= \max_{a \leq x \leq b} |f(x)+g(x)| = |f(\bar{x})+g(\bar{x})| \\ &\leq |f(\bar{x})| + |g(\bar{x})| \\ &\leq \|f\|_\infty + \|g\|_\infty \end{aligned}$$

• Ορισμός: Ισοδύναμες (ή συγκρίσιμες) νόρμες :

- Έστω X ένας γραμμικός χώρος. Δύο νόρμες $\|\cdot\|$ και $\|\cdot\|'$ στον X λέγονται ισοδύναμες, αν υπάρχουν $M, m > 0$ τ.ω.

$$\forall x \in X \quad m\|x\| \leq \|x\|' \leq M\|x\|$$

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

$\swarrow m' \qquad \searrow m'$

- Πρόταση: Ισοδυναμία νορμών στον \mathbb{R}^n

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

• Ορισμός: Σύγκλιση ακολουθίας:

Λέμε ότι η ακολουθία $(x^{(n)})_{n \in \mathbb{N}} \subset X$ συγκλίνει ως προς μία συγκεκριμένη νόρμα $\|\cdot\|$ του X αν υπάρχει $x \in X$ (το όριο της ακολουθίας ως προς την νόρμα $\|\cdot\|$), τέτοιο ώστε,

$$\|x^{(n)} - x\| \rightarrow 0, n \rightarrow \infty$$

• Συμβολισμός:

$$x^{(n)} \rightarrow x, n \rightarrow \infty \text{ (ως προς τη νόρμα } \|\cdot\| \text{)}$$

- Σύγκλιση μιας ακολουθίας ως προς μία νόρμα συνεπάγεται σύγκλιση της ίδιας ακολουθίας στο ίδιο όριο ως προς οποιαδήποτε άλλη νόρμα ισοδύναμη της αρχικής.

- Ιδιαίτερα, στον \mathbb{R}^n , η σύγκλιση μιας ακολουθίας $(x^{(m)})_{m \in \mathbb{N}}$ σε ένα $x \in \mathbb{R}^n$ είναι ισοδύναμη με σύγκλιση ως προς τη νόρμα μέγιστου, δηλαδή

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

• Ορισμός: Πλήρης χώρος :

Ένας γραμμικός χώρος με νόρμα $(X, \|\cdot\|)$ λέγεται πλήρης αν κάθε ακολουθία Cauchy (ως προς τη $\|\cdot\|$) στον X συγκλίνει (ως προς $\|\cdot\|$), δηλαδή αν για $(x^{(n)})_{n \in \mathbb{N}} \subset X$

$$\forall \varepsilon > 0 \exists N \in \mathbb{N}, \forall m, n \geq N \quad \|x^{(m)} - x^{(n)}\| < \varepsilon$$

$$\Leftrightarrow \exists x \in X \quad \lim_{n \rightarrow \infty} \|x^{(n)} - x\| = 0$$

• Θεώρημα: Πληρότητα του \mathbb{R}^n :

- Έστω $\|\cdot\|$ μια νόρμα στον \mathbb{R}^n . Τότε ο $(\mathbb{R}^n, \|\cdot\|)$ είναι πλήρης.

Γενικά: $(a_n)_{n \in \mathbb{N}} \subset \mathbb{R}$ ακολουθία Cauchy
 $\forall \varepsilon > 0 \exists N \in \mathbb{N} \forall n, m \geq N, |a_n - a_m| < \varepsilon$

• Νόρμες Πινάκων:

Οι $n \times n$ πίνακες μας ενδιαφέρουν κυρίως ως απεικονίσεις από τον \mathbb{R}^n στον \mathbb{R}^n . Για αυτό τροποποιούμε λίγο τον ορισμό της νόρμας ως εξής:

• Ορισμός: Νόρμα Πινάκων:

- Μια απεικόνιση $\|\cdot\|: \mathbb{R}^{n,n} \rightarrow \mathbb{R}$ λέγεται νόρμα πινάκων, αν ισχύουν:

$$(N_1) \quad A \in \mathbb{R}^{n,n}, \quad \|A\| = 0 \quad (\Leftrightarrow) \quad A = 0$$

$$(N_2) \quad \forall \lambda \in \mathbb{R} \quad \forall A \in \mathbb{R}^{n,n}, \quad \|\lambda A\| = |\lambda| \cdot \|A\|$$

$$(N_3) \quad \forall A, B \in \mathbb{R}^{n,n}, \quad \|A+B\| \leq \|A\| + \|B\|$$

$$(N_4) \quad \forall A, B \in \mathbb{R}^{n,n}, \quad \|AB\| \leq \|A\| \cdot \|B\|$$

- Θα ασχοληθούμε μόνο με τις λεγόμενες φυσικές νόρμες πινάκων.

• Ορισμός: Φυσική νόρμα πίνακα:

Έστω $\|\cdot\|$ μια νόρμα στον \mathbb{R}^n .

Η απεικόνιση,

$$\|\cdot\|: \mathbb{R}^{n,n} \rightarrow \mathbb{R}, \quad \|A\| := \max_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|}{\|x\|}$$

(που είναι καλά ορισμένη) λέγεται

φυσική νόρμα πινάκων ή νόρμα παραγόμενη (ή επαγόμενη) από τη νόρμα $\|\cdot\|$ του \mathbb{R}^n .

• Κάθε φυσική νόρμα πινάκων είναι νόρμα πινάκων.

Βασική ιδιότητα:

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

- Για $x=0$ η $(*)$ είναι προφανής.
- Για $x \neq 0$ έχουμε

$$\|Ax\| \leq \|A\| \cdot \|x\| \Leftrightarrow \frac{\|Ax\|}{\|x\|} \leq \|A\|$$

και η τελευταία εκτίμηση ισχύει λόγω του ορισμού της $\|A\|$.

Πρόβλημα: Πώς υπολογίζεται μια φυσική νόρμα πίνακα;

- Αυτό γενικά είναι ένα δύσκολο πρόβλημα.

① • Έστω $c_1 \in \mathbb{R}$ τ.ω.

$$\forall x \in \mathbb{R}^n \quad \|Ax\| \leq c_1 \|x\|$$

Τότε έχουμε $\|A\| \leq c_1$

② • Έστω $y \in \mathbb{R}^n$, $y \neq 0$, τ.ω.

$$\|Ay\| \geq c_2 \|y\|$$

με $c_2 > 0$, τότε $\|A\| \geq c_2$

- Αποδεικνύουμε τις ① και ② με το μικρότερο c_1 , και το μεγαλύτερο c_2 που μπορούμε και έχουμε,

$$c_2 \leq \|A\| \leq c_1$$

Για $c_1 = c_2$, παίρνουμε $\|A\| = c_1$

Παράδειγμα :

1. $(\mathbb{R}^n, \|\cdot\|_1)$

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \quad (\text{νόρμα του αθροίσματος στήλων})$$

2. $(\mathbb{R}^n, \|\cdot\|_\infty)$

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \quad (\text{νόρμα του αθροίσματος γραμμών})$$

(Φυσικές) νόρμες πινάκων

Έστω $\|\cdot\|$ μια νόρμα στον \mathbb{R}^n .

Ορίζουμε,

$$\|\cdot\| : \mathbb{R}^{n,n} \rightarrow \mathbb{R},$$

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

αυτή είναι μια νόρμα πινάκων

• Βασική ιδιότητα:

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

$$\|I_n\| = 1$$

• $\|\cdot\|_1$ στον \mathbb{R}^n

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}| \quad (\text{νόρμα του αθροίσματος στηλών})$$

• $\|\cdot\|_\infty$ στον \mathbb{R}^n

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}| \quad (\text{νόρμα του αθροίσματος γραμμών})$$

• $\|\cdot\|_2$ στον \mathbb{R}^n

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

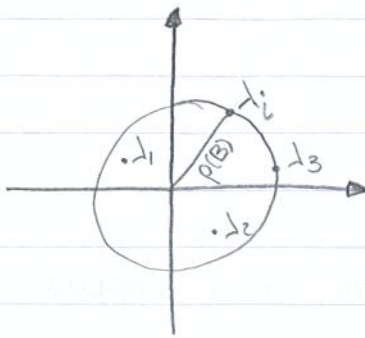
• Φασματική οκτίνα πίνακα

$$B \in \mathbb{R}^{n,n}$$

Έστω $\lambda_1, \dots, \lambda_n \in \mathbb{C}$ οι ιδιοτιμές του

φασματική ακτίνα $\rho(B)$:

$$\rho(B) = \max_{1 \leq i \leq n} |\lambda_i|$$



Ο $A^T A$ είναι συμμετρικός,

$(A^T A)^T = A^T \cdot (A^T)^T = A^T \cdot A$ οπότε οι ιδιοτιμές του είναι πραγματικές.

Λόγω των τύπων αυτών οι νόρμες $\|\cdot\|_1$, $\|\cdot\|_\infty$ και $\|\cdot\|_2$ είναι ιδιαίτερα εύχρηστες.

Αν $A \in \mathbb{C}^{n,n}$ τότε οι $\|A\|_1$ και $\|A\|_\infty$ δίνονται από τους ίδιους τύπους όπως προηγουμένως.

Επίσης, $\|A\|_2 = (\rho(A^* A))^{1/2}$ με $(A^*)_{ij} = \bar{a}_{ij}$

• Δείκτης κατάστασης:

• $Ax=b$, $A \in \mathbb{R}^{n,n}$ αντιστρέψιμος $b \in \mathbb{R}^n$, $\|\cdot\|$ νόρμα στον \mathbb{R}^n .

Ερώτημα: Τι μπορούμε να πούμε για την κατάσταση του $Ax=b$;

- Υπόθεση $b \neq 0$.

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

Έχουμε,

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

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

Άρα,

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

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

Επομένως,

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

Τώρα,

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

$$\Rightarrow \boxed{\|x\| \geq \frac{\|b\|}{\|A\|}} \textcircled{2}$$

Οι $\textcircled{1}$ και $\textcircled{2}$ δίνουν,

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

$$\Rightarrow \frac{\|\Delta x\|}{\|x\|} \leq \underbrace{\|A\| \cdot \|A^{-1}\|}_{\kappa(A) \text{ cond}(A)} \frac{\|\Delta b\|}{\|b\|} \textcircled{*}$$

Η $\textcircled{*}$ ισχύει ως ισότητα για κατάλληλα $b, \Delta b \neq 0$.

Ο $\kappa(A) = \|A\| \cdot \|A^{-1}\|$ λέγεται δείκτης κατάστασης του A και είναι συντελεστής ευαισθησίας του $Ax=b$ σε μεταβολές στη νόρμα $\|\cdot\|$ (Turkington)

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

• $1 = \|\mathbb{I}_n\| = \|A \cdot A^{-1}\| \leq \|A\| \cdot \|A^{-1}\| = \kappa(A)$

- Όσο μεγαλύτερος ο $\kappa(A)$ τόσο χειρότερη η κατάσταση του συστήματος (του πίνακα) όσο μικρότερος (κοντά στο 1) τόσο καλύτερη η κατάσταση.

• Θεώρημα : Εκτίμηση της σχετικής μεταβολής λύσεων γραμμικών συστημάτων.

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

Έστω $A \in \mathbb{R}^{n,n}$ -αντιστρέψιμος, $\Delta A \in \mathbb{R}^{n,n}$ και $b, \Delta b \in \mathbb{R}^n$ με $b \neq 0$. Τότε, αν $\kappa(A) = \|A\| \cdot \|A^{-1}\|$ έχουμε:

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

b) Αν $\|A^{-1}\| \cdot \|\Delta A\| < 1$, τότε ο $A+\Delta A$ είναι αντιστρέψιμος και

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

γ) Αν $\|A^{-1}\| \cdot \|\Delta A\| < 1$, τότε ο $A + \Delta A$ είναι αντιστρέψιμος και

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

- Οι (α), (β) είναι ειδικές περιπτώσεις της (γ) και δεν βελτιώνονται.

$$\cdot \left(\frac{\|\Delta b\|}{\|b\|} + \frac{\|\Delta A\|}{\|A\|} \right)$$

Απόδειξη

α) ✓

β) Θα αποδείξουμε ότι ο $A + \Delta A$ είναι αντιστρέψιμος και

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

Από Άσκηση 3.36 :

$$B \in \mathbb{R}^{n,n} \quad \|B\| < 1$$

$\Rightarrow I_n - B$ αντιστρέψιμος και

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

Έχουμε,

$$A + \Delta A = A \left(I_n + \underbrace{A^{-1} \Delta A}_B \right)$$

Επίσης,

$$\|A^{-1} \Delta A\| \leq \|A^{-1}\| \cdot \|\Delta A\| < 1,$$

οπότε σύμφωνα με την Άσκηση 3.36 ο πίνακας $I_n + A^{-1} \Delta A$ είναι αντιστρέψιμος και

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

Alja,

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

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

$$\stackrel{\oplus}{\leq} \frac{1}{1 - \|A^{-1} \Delta A\|} \|A^{-1}\|$$

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

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

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

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

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

$$\stackrel{\oplus\oplus}{\leq} \frac{\|A^{-1}\|}{1 - \|A^{-1}\| \cdot \|\Delta A\|} \cdot \|\Delta A\|$$

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

Απόδειξη του γ) :

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

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

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

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

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

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

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

• Επαναληπτικές μέθοδοι

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

Μέθοδοι

- Με αρχική προσέγγιση $x^{(0)} \in \mathbb{R}^n$ αυτές οι μέθοδοι δίνουν αναδρομικά μια ακολουθία προσεγγίσεων $(x^{(m)})_{m \in \mathbb{N}}$ της x .

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

Μέθοδος του Jacobi

Μέθοδος των Gauss-Seidel

$$(Ax)_i = b_i \Rightarrow \sum_{j=1}^n a_{ij} x_j = b_i \Rightarrow$$

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

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

• Μέθοδος Jacobi:

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

• Μέθοδος Gauss-Seidel:

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

• Επαναληπτικές μέθοδοι:

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

• Υπόθεση: $a_{ii} \neq 0, i = 1, \dots, n$

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

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

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

• Η μέθοδος του Jacobi μπορεί να υλοποιηθεί παράλληλα ενώ η μέθοδος των Gauss-Seidel δεν μπορεί.

• Αν συγκρίνουν οι μέθοδοι είναι ιδιαίτερα κατάλληλες αν ο A είναι σπαιός.

• Για τη θεωρητική μελέτη έχει πλεονεκτήματα να γράφουμε τις μεθόδους λίγο διαφορετικά. Με

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

$$U = \begin{pmatrix} 0 & a_{12} & a_{13} & \dots & a_{1n} \\ & 0 & a_{23} & \dots & a_{2n} \\ & & 0 & & \\ & & & \ddots & \\ & & & & a_{n-1,n} \\ 0 & \dots & & & 0 \end{pmatrix}$$

και έχουμε $A = L + D + U$
(Αυτοί οι πίνακες L και U δεν έχουν καμία σχέση με την ανάλυση LU).

Jacobi: $Dx^{(m+1)} = b - Lx^{(m)} - Ux^{(m)}$

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

Gauss-Seidel:

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

• Γενική επαναληπτική μέθοδος

$A = M - N$

$M, N \in \mathbb{R}^{n,n}$, M αντιστρέψιμος

$Ax = b$ (=) $(M - N)x = b$ (=) $Mx = Nx + b$

Μέθοδος:

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

↑
Λύνουμε ένα σύστημα με πίνακα M .

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

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

• Μελέτη της γενικής επαναληπτικής μεθόδου.

$$x^{(0)} \in \mathbb{R}^n \quad Mx^{(m+1)} = Nx^{(m)} + b, \quad m \in \mathbb{N}_0$$

→ Πότε ισχύει $x^{(m)} \rightarrow x, m \rightarrow \infty$ (για κάθε $x^{(0)} \in \mathbb{R}^n$);

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

Ισχυρισμός: $y = x$

$$\left. \begin{array}{l} Mx^{(m+1)} = Nx^{(m)} + b \\ Mx = Nx + b \end{array} \right\} \Rightarrow$$

$$\begin{aligned} M(x^{(m+1)} - x) &= N(x^{(m)} - x) \xrightarrow{m \rightarrow \infty} \\ M(y - x) &= N(y - x) (=) \end{aligned}$$

$$(M - N)(y - x) = 0 (=) \quad A(y - x) = 0 (=) \boxed{y = x}$$

→ Πότε συγκλίνει μια επαναληπτική μέθοδος;

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

$$x^{(m+1)} - x = \underbrace{M^{-1}N}_{=G} (x^{(m)} - x)$$

- Ο G λέγεται πίνακας επανάληψης της μεθόδου.

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

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

↑
επαγωγικά

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

$$\Rightarrow \|x^{(m)} - x\| \leq \|G^m\| \cdot \|x^{(0)} - x\|$$

↑
Για δεδομένο m
υπάρχει $x^{(0)} \in \mathbb{R}^n$
τ.ω. αυτή η
ανισότητα να ισχύει
ως ισότητα

θέλω να ισχύει $\forall x$ άρα δεν
το μεταβάλλω

→ Η μέθοδος συγκλίνει αν και μόνο αν $G^m \rightarrow 0, m \rightarrow \infty$

• Λήμμα (Σχέση φυσικής νόρμας και φασματικής ακτίνας πίνακα)

- Έστω $\|\cdot\|$ οποιαδήποτε νόρμα στον \mathbb{C}^n . Τότε για κάθε $P \in \mathbb{C}^{n,n}$ για την αντίστοιχη φυσική νόρμα έχουμε,

$$(*) \quad \rho(P) \leq \|P\|$$

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

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

Απόδειξη (του πρώτου μέρους):

Έστω λ ιδιοτιμή του P , $\delta n \lambda$.

$$Pz = \lambda z, \text{ με } z \neq 0$$

Τότε,

$$\|\lambda z\| = \|Pz\| \leq \|P\| \cdot \|z\|$$

$$\| \lambda \cdot \|z\|$$

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

Από εδώ επεται αμέσως η $(*)$

• Θεώρημα (Ίκανές και αναγκαίες συνθήκες σύγκλισης επαναληπτικών μεθόδων).

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

α) Η επαναληπτική μέθοδος $Mx^{(m+1)} = Nx^{(m)} + b$ συγκλίνει, δηλαδή για κάθε $x^{(0)} \in \mathbb{C}^n$ έχουμε $x^{(m)} \rightarrow x$ $m \rightarrow \infty$.

β) $\rho(G) < 1$, με $G = M^{-1} \cdot N$ τον πίνακα επανάληψης της μεθόδου.

γ) Υπάρχει φυσική νόρμα πινάκων τ.ω. $\|G\| < 1$.

δ) $\lim_{n \rightarrow \infty} G^n = 0$.

Απόδειξη:

$$a) \Rightarrow b) \Rightarrow \gamma) \Rightarrow \delta) \Rightarrow a)$$

a) \Rightarrow β): Έστω $x^{(m)} \rightarrow x$, για $m \rightarrow \infty$.

Τότε σύμφωνα με την,

$$x^{(m)} - x = G^m (x^{(0)} - x), \text{ συμπεραίνουμε ότι,}$$

$$G^m (x^{(0)} - x) \rightarrow 0, m \rightarrow \infty \text{ για οποιοδήποτε } x^{(0)} \in \mathbb{C}^n.$$

Για $y \in \mathbb{C}^n$, με $x^{(0)} = x + y$, συμπεραίνουμε ότι $\bigoplus G^m y \rightarrow 0, m \rightarrow \infty$

Έστω λ ιδιοτιμή του G και $z \in \mathbb{C}^n$ αντίστοιχο ιδιοδιάνυσμα.

Τότε, $G^m \cdot z = \lambda^m \cdot z$ \rightarrow $Gz = \lambda z$

$$\begin{aligned} G \cdot Gz &= \lambda \underbrace{Gz} = \lambda^2 z \\ G^2 z &= \lambda^2 z \\ G^m z &= \lambda^m z \end{aligned}$$

άρα λ^m ιδιοτιμή του πίνακα G^m .

οπότε η \oplus δίνει,

$$\lambda^m z \rightarrow 0, m \rightarrow \infty$$

Αλλά,

$$\|\lambda^m z\| = |\lambda^m| \cdot \|z\| = |\lambda|^m \underbrace{\|z\|}_{\neq 0}$$

οπότε η,

$$|\lambda|^m \cdot \|z\| \rightarrow 0, m \rightarrow \infty$$

συνεπάγεται ότι $|\lambda| < 1$. Επομένως, $\rho(G) < 1$.

$\beta) \Rightarrow \gamma)$: Έστω $0 < \varepsilon < \underbrace{1 - \rho(G)}_{> 0}$

Σύμφωνα με το προηγούμενο λήμμα υπάρχει φυσική νόρμα $\|\cdot\|$ στο $\mathbb{C}^{n,n}$ τ.ω.,

$$\|G\| \leq \rho(G) + \varepsilon$$

Αλλά,

$$\rho(G) + \varepsilon < \cancel{\rho(G)} + 1 - \cancel{\rho(G)} = 1$$

$$\text{οπότε } \|G\| < 1$$

$\gamma) \Rightarrow \delta)$:

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

$$\|AB\| \leq \|A\| \cdot \|B\|$$

$$\|G^2\| \leq \|G\| \cdot \|G\| = \|G\|^2$$

$$\begin{aligned} \|x^{(m)} - x\| &\leq \|G^m\| \cdot \|x^{(0)} - x\| \\ &\leq \|G\|^m \end{aligned}$$

$$\Rightarrow \|x^{(m)} - x\| \leq \underbrace{\|G\|^m}_{\rightarrow 0 \text{ καθώς } m \rightarrow \infty} \cdot \|x^{(0)} - x\|$$

γιατί $\|G\| < 1$

δ) ⇒ α) :

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

$$\Rightarrow \|x^{(m)} - x\| \leq \underbrace{\|G^m\|}_{\rightarrow 0, m \rightarrow \infty} \cdot \|x^{(0)} - x\|$$

$$\Rightarrow x^{(m)} - x \rightarrow 0, m \rightarrow \infty$$

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

$$G = \begin{pmatrix} 0 & 2 \\ 0 & 0 \end{pmatrix} \quad \Bigg| \quad G^2 = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}$$

$$\|G\|_{\infty} = 2$$

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

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

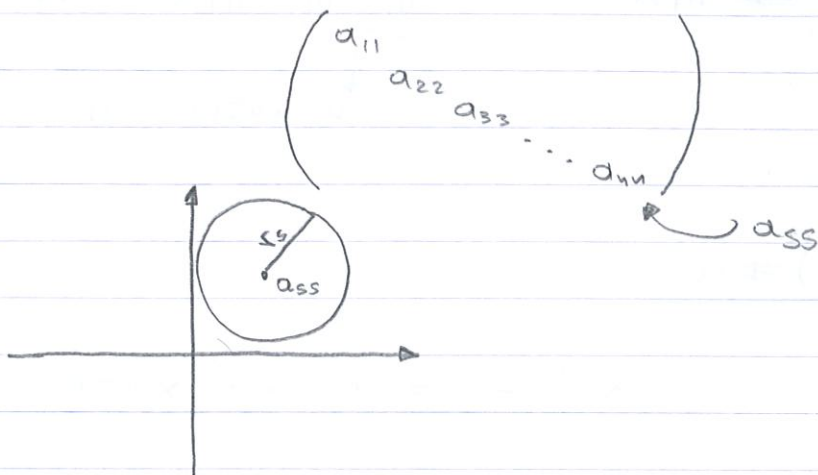
$\rho(G) = 0$ (η μόνη ιδιοτιμή του είναι $\lambda = 0$).

Λήμμα (Ανισότητα του Gerschgorin) :

- Έστω $A = (a_{ij}) \in \mathbb{C}^{n,n}$ ένας πίνακας και λ ιδιοτιμή του.

→ Τότε υπάρχει $s \in \{1, \dots, n\}$ τ.ω.

$$|a_{ss} - \lambda| \leq \underbrace{\sum_{\substack{j=1 \\ j \neq s}}^n |a_{sj}|}_{= r_s}$$



Απόδειξη:

Έστω z αντίστοιχο ιδιοδιάνυσμα,

$$Az = \lambda z$$

Τότε,

$$(Az)_i = \lambda z_i \Rightarrow$$

$$\sum_{j=1}^n a_{ij} z_j = \lambda z_i \Rightarrow$$

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

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

- Έστω $s \in \{1, \dots, n\}$ τ.ω. $\underbrace{|z_s|}_{\neq 0} = \|z\|_\infty = \max_{1 \leq i \leq n} |z_i|$

Για $i=s$, η $*$ δίνει,

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

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

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

• Πινακες με αυστηρά κυριαρχική διαχώνιο

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

Ισχυρισμός: Αν ο $A \in \mathbb{C}^{n,n}$ έχει αυστηρά κυριαρχική διαχώνιο, τότε είναι αντιστρέψιμος.

$$Az = 0z \quad \text{με } z \neq 0$$

• $Az = 0 \Rightarrow z = 0$

Συμπέρασμα: Ένας πίνακας είναι αντιστρέψιμος, αν και μόνο αν το 0 δεν είναι ιδιοτιμή του.

• Έστω ότι το $\lambda = 0$ είναι ιδιοτιμή του A . Τότε, σύμφωνα με την ανισότητα του Gerschgorin.

$$|a_{ss} - 0| \leq \sum_{\substack{j=1 \\ j \neq s}}^n |a_{sj}| \quad \text{για κάποιο } s \in \{1, \dots, n\}$$

$A \in \mathbb{C}^{n,n}$ έχει αυστηρά κυρίαρχικό διαγώνιο:

$$(*) \quad |a_{ii}| > \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|, \quad i=1, 2, \dots, n$$

• Πρόταση (Σύγκλιση των μεθόδων του Jacobi και των Gauss-Seidel στην περίπτωση πίνακα με αυστηρά κυρίαρχικό διαγώνιο) (ικανή συνθήκη)

Έστω $A \in \mathbb{C}^{n,n}$ ένας πίνακας με αυστηρά κυρίαρχικό διαγώνιο.

Τότε:

α) οι πίνακες επανάληψης,

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

των μεθόδων του Jacobi και των Gauss-Seidel, αντίστοιχα ικανοποιούν τις ανισότητες,

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

β) οι μέθοδοι του Jacobi και Gauss-Seidel συγκλίνουν για γραμμικά συστήματα $Ax=b$ με τέτοιους πίνακες.

• Απόδειξη (του α)

Όπως ξέρουμε, ο A είναι αντιστρέψιμος. Επίσης, σύμφωνα με την $(*)$ έχουμε $a_{ii} \neq 0, i=1, \dots, n$, οπότε οι μέθοδοι αυτές μπορούν να εφαρμοστούν.

Θέτουμε,
$$G := \max_{1 \leq i \leq n} \frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}|$$

Σύμφωνα με την \ast , έχουμε $C < 1$

Jacobi :

$$D^{-1} = \begin{pmatrix} \frac{1}{a_{11}} & 0 & \dots & 0 \\ 0 & \frac{1}{a_{22}} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & \dots & \dots & \frac{1}{a_{nn}} \end{pmatrix}$$

Άρα,

$$-D^{-1}(L+U) = - \begin{pmatrix} 0 & \frac{a_{12}}{a_{11}} & \dots & \frac{a_{1n}}{a_{11}} \\ \vdots & 0 & \dots & \vdots \\ \vdots & \vdots & \ddots & \vdots \\ \frac{a_{n1}}{a_{nn}} & \frac{a_{n2}}{a_{nn}} & \dots & \frac{a_{n-1,n}}{a_{nn}} \end{pmatrix}$$

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

Gauss-Seidel : Για να παρακάμψουμε το πρόβλημα του υπολογισμού του $(L+D)^{-1}$ θέτουμε,

$$u = G_{GS} \cdot y, \text{ με } y \in \mathbb{C}^n$$

Τότε,

$$u = -(L+D)^{-1} U y (=)$$

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

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

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

Ισχυρισμός:

$$(**) \quad |u_i| \leq C \|y\|_\infty, \quad i=1, \dots, n$$

Απόδειξη της (**) επαγωγικά:

$$\underline{i=1} : \text{ Έχουμε, } u_1 = \frac{1}{a_{11}} \left(- \sum_{j=2}^n a_{1j} y_j \right)$$

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

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

$$\Rightarrow |u_1| \leq C \|y\|_\infty$$

Επαγωγικό βήμα: Έστω ότι η **(**)** ισχύει για $1, \dots, i-1$
 $\leq C \|y\|_\infty$ (υπόθεση επαγωγής)

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

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

$$\Rightarrow \underbrace{\left(\frac{1}{|a_{ii}|} \sum_{\substack{j=1 \\ j \neq i}}^n |a_{ij}| \right)}_{\leq C} \cdot \|y\|_\infty$$

$$\Rightarrow |u_i| \leq C \cdot \|y\|_\infty$$

Γράφουμε την ****** στη μορφή $\underbrace{\max_{1 \leq i \leq n} |u_i|}_{= \|u\|_\infty} \leq C \|y\|_\infty$

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

$$\Rightarrow \boxed{\|G_{GS}\|_\infty \leq C}$$

Το (b) έπεται άμεσα από το (a).