

Γραμμικά συστήματα

Τμήμα Ναυπηγών Μηχανικών

Δημήτρης Μητσούδης

Σύντομη επανάληψη στους πίνακες

Πίνακας $m \times n$ (m γραμμές, n στήλες)

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & a_{m3} & \dots & a_{mn} \end{pmatrix}$$

a_{ij} : το στοιχείο που βρίσκεται στην i -γραμμή, και στην j -στήλη.

Διανύσματα

Οι πίνακες που έχουν μόνο μία γραμμή ή μόνο μία στήλη λέγονται **διανύσματα** (vectors). Π.χ.,

$$a = \begin{pmatrix} 1 & 2 & -3 & 0 & \sqrt{2} \end{pmatrix}, \quad b = \begin{pmatrix} -1 \\ 2 \\ 0 \\ -3 \end{pmatrix}.$$

Θα συμβολίζουμε τα διανύσματα με πεζά γράμματα, π.χ. $a, b, \dots$

Τα στοιχεία ενός διανύσματος λέγονται **συνιστώσες** του διανύσματος και αρκεί ένας μόνο δείκτης για να καθορίσει τη θέση του στοιχείου στο διάνυσμα.

Π.χ. στο παραπάνω διάνυσμα γραμμή a , το τρίτο στοιχείο του είναι το $a_3 = -3$, ενώ το πρώτο στοιχείο του διανύσματος στήλη b είναι το $b_1 = -1$.

Ειδικές μορφές πινάκων

- Ένας αριθμός αποτελεί έναν πίνακα 1×1 που λέγεται πίνακας στοιχείο.
- Ένας πίνακας $m \times n$ λέγεται **μηδενικός** αν όλα του τα στοιχεία είναι μηδέν. Συμβολ. $\mathbf{0}_{m \times n}$ (ή, απλώς, $\mathbf{0}$)

$$\mathbf{0}_{2 \times 3} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

- Ένας πίνακας $m \times n$ λέγεται **τετραγωνικός** αν $m = n$.

$$A = \begin{pmatrix} -1 & 0 & 2 \\ 2 & -1 & 3 \\ 0 & 5 & -3 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & 2 \\ 5 & -1 \end{pmatrix}.$$

- Ένας τετραγωνικός $n \times n$ πίνακας λέγεται **διαγώνιος**, αν μόνο η διαγωνιάς του (δηλ. τα στοιχεία της μορφής a_{ii} , για $1 \leq i \leq n$) περιέχει μη μηδενικά στοιχεία. Δηλαδή

$$a_{ij} = 0 \text{ για } i \neq j.$$

Π.χ.,

$$A = \begin{pmatrix} -1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 4 \end{pmatrix}.$$

- Ο διαγώνιος $n \times n$ πίνακας που όλα τα στοιχεία της διαγωνιάς του είναι μονάδες λέγεται **μοναδιαίος** και συμβολίζεται με I_n . Π.χ.,

$$I_3 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \quad I_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}.$$

Σχέσεις και πράξεις πινάκων

- **Ισότητα πινάκων.** Έστω $A = (a_{ij})$, $B = (b_{ij})$, δύο $m \times n$ πίνακες. Λέμε ότι οι πίνακες είναι ίσοι, και συμβολίζουμε $A = B$, αν

$$a_{ij} = b_{ij}, \quad 1 \leq i \leq m, 1 \leq j \leq n.$$

- **Πρόσθεση πινάκων.** Έστω $A = (a_{ij})$, $B = (b_{ij})$, δύο $m \times n$ πίνακες. Ορίζουμε ως άθροισμα των A και B , και συμβολίζουμε $A + B$, τον $m \times n$ πίνακα $C = (c_{ij})$ τα στοιχεία του οποίου δίνονται από τη σχέση

$$c_{ij} = a_{ij} + b_{ij}, \quad 1 \leq i \leq m, 1 \leq j \leq n.$$

Π.χ.

$$\begin{pmatrix} -1 & 2 & 1 \\ 5 & -1 & 2 \end{pmatrix} + \begin{pmatrix} -3 & 0 & 2 \\ -2 & 1 & 3 \end{pmatrix} = \begin{pmatrix} -4 & 2 & 3 \\ 3 & 0 & 5 \end{pmatrix}.$$

- **Γινόμενο αριθμού επί πίνακα (βαθμωτό γινόμενο).** Έστω $A = (a_{ij})$ ένας $m \times n$ πίνακας και λ ένας πραγματικός αριθμός. Το (βαθμωτό) τους γινόμενο συμβολίζεται με λA και είναι ένας $m \times n$ πίνακας $C = (c_{ij})$ με στοιχεία

$$c_{ij} = \lambda a_{ij}, \quad 1 \leq i \leq m, 1 \leq j \leq n.$$

Π.χ.

$$3 \cdot \begin{pmatrix} -1 & 2 \\ 5 & -4 \end{pmatrix} = \begin{pmatrix} -3 & 6 \\ 15 & -12 \end{pmatrix}.$$

- **Γινόμενο πινάκων.** Έστω $A = (a_{ij})$ ένας $m \times n$ πίνακας και $B = (b_{ij})$ ένας $n \times r$ πίνακας. Τότε το γινόμενό τους συμβολίζεται με AB και ορίζεται να είναι ένας $m \times r$ πίνακας C με στοιχεία

$$c_{ij} = \sum_{k=1}^n a_{ik} b_{kj} = a_{i1} b_{1j} + a_{i2} b_{2j} + \cdots + a_{in} b_{nj}, \quad 1 \leq i \leq m, 1 \leq j \leq r.$$

Προσοχή: Για να πολλαπλασιάσουμε δύο πίνακες A και B πρέπει ο αριθμός στηλών του A να είναι ίσος με τον αριθμό γραμμών του B .

Παρατηρήστε επίσης ότι ο AB έχει αριθμό γραμμών ίσο με τον αριθμό γραμμών του πρώτου και αριθμό στηλών ίσο με τον αριθμό στηλών του δεύτερου.

Παραδείγματα

Έστω ότι θέλουμε να πολλαπλασιάσουμε έναν 4×3 πίνακα A με έναν 3×2 πίνακα B . Το αποτέλεσμα θα είναι ένας 4×2 πίνακας C .

$$\underbrace{\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \\ a_{41} & a_{42} & a_{43} \end{pmatrix}}_{4 \times 3} \underbrace{\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}}_{3 \times 2} = \underbrace{\begin{pmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \\ c_{31} & c_{32} \\ c_{41} & c_{42} \end{pmatrix}}_{4 \times 2}$$

Π.χ., το στοιχείο του πίνακα C που βρίσκεται στην 3η γραμμή και 1η στήλη, δηλαδή το c_{31} , υπολογίζεται ως εξής:

Πολλαπλασιάζουμε κάθε στοιχείο της 3ης γραμμής του πίνακα A με το αντίστοιχο στοιχείο της 1ης στήλης του πίνακα B και στη συνέχεια αθροίζουμε τα τρία αυτά γινόμενα. Δηλαδή

$$c_{31} = a_{31}b_{11} + a_{32}b_{21} + a_{33}b_{31}.$$

Παραδείγματα

1

$$\begin{pmatrix} -1 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix} \begin{pmatrix} -1 & 0 & 0 \\ 1 & 2 & 0 \\ -2 & 7 & 5 \end{pmatrix} = \begin{pmatrix} 1 & 11 & 5 \\ -5 & 12 & 10 \end{pmatrix}.$$

2

$$\begin{pmatrix} 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} = 1 \cdot 2 + 0 \cdot 3 + (-1) \cdot 1 = 1.$$

3

$$\begin{pmatrix} 2 \\ 3 \\ 1 \end{pmatrix} \begin{pmatrix} 1 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 2 \cdot 1 & 2 \cdot 0 & 2 \cdot (-1) \\ 3 \cdot 1 & 3 \cdot 0 & 3 \cdot (-1) \\ 1 \cdot 1 & 1 \cdot 0 & 1 \cdot (-1) \end{pmatrix} = \begin{pmatrix} 2 & 0 & -2 \\ 3 & 0 & -3 \\ 1 & 0 & -1 \end{pmatrix}.$$

Αν A είναι ένας τετραγωνικός πίνακας και k είναι ένας φυσικός αριθμός, ορίζουμε

$$A^k = \underbrace{AA \dots A}_{k \text{ φορές}}.$$

Ιδιότητες των πράξεων πινάκων

- ❶ Έστω A, B δύο $m \times n$ πίνακες. Τότε $A + B = B + A$.
- ❷ Έστω A ένας $m \times n$ πίνακας. Τότε $A + \mathbf{0} = \mathbf{0} + A = A$.
- ❸ Έστω A, B, C τρεις $m \times n$ πίνακες. Τότε $(A + B) + C = A + (B + C) = A + B + C$.
- ❹ Έστω A ένας $k \times \ell$ πίνακας, B ένας $\ell \times m$ πίνακας, και C ένας $m \times n$ πίνακας. Τότε $(AB)C = A(BC) = ABC$.
- ❺ Έστω A, B δύο $k \times m$ πίνακες και C ένας $m \times n$ πίνακας. Τότε $(A + B)C = AC + BC$.
- ❻ Αν D είναι ένας $k \times m$ πίνακας και E, F δύο $m \times n$ πίνακες, τότε $D(E + F) = DE + DF$.
- ❼ Αν A είναι τετραγωνικός $n \times n$ πίνακας, τότε $AI_n = I_n A = A$.

Προσοχή!!! – 1

Γενικά **δεν** ισχύει ότι $AB = BA$, ακόμη και αν οι πίνακες A και B είναι τετραγωνικοί. Π.χ., ελέγξτε ότι αν

$$A = \begin{pmatrix} -1 & 2 \\ 1 & 0 \end{pmatrix} \quad \text{και} \quad B = \begin{pmatrix} 2 & 1 \\ 0 & 1 \end{pmatrix},$$

τότε

$$AB = \begin{pmatrix} -2 & 1 \\ 2 & 1 \end{pmatrix} \quad \text{αλλά} \quad BA = \begin{pmatrix} -1 & 4 \\ 1 & 0 \end{pmatrix}.$$

Αν για δύο πίνακες A και B ισχύει ότι $AB = BA$, τότε λέμε ότι οι A και B **μετατίθενται**.

Προσοχή!!! – 2

Επιπλέον, αν π.χ.

$$A = \begin{pmatrix} 1 & 1 \\ 2 & 2 \end{pmatrix} \quad \text{και} \quad B = \begin{pmatrix} -1 & 1 \\ 1 & -1 \end{pmatrix},$$

τότε

$$AB = \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \quad \text{και} \quad BA = \begin{pmatrix} 1 & 1 \\ -1 & -1 \end{pmatrix}.$$

Επομένως, ακόμη και στην περίπτωση τετραγωνικών πινάκων, αν $AB = \mathbf{0}$, **δεν** ισχύει απαραίτητα ότι $BA = \mathbf{0}$, ούτε ότι $A = \mathbf{0}$ ή $B = \mathbf{0}$.

Ανάστροφος ενός πίνακα

Ορισμός

Έστω $A = (a_{ij})$ ένας $m \times n$ πίνακας. Ο **ανάστροφος** του A συμβολίζεται με A^T και είναι ένας $n \times m$ πίνακας (b_{ij}) που έχει ως γραμμές τις στήλες του A (και ως στήλες τις γραμμές του A). Δηλαδή θα έχουμε $b_{ij} = a_{ji}$.

Παράδειγμα.

$$\text{Αν } A = \begin{pmatrix} -1 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix}, \quad \text{τότε} \quad A^T = \begin{pmatrix} -1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix}.$$

Ανάστροφος ενός πίνακα

Ορισμός

Έστω $A = (a_{ij})$ ένας $m \times n$ πίνακας. Ο **ανάστροφος** του A συμβολίζεται με A^T και είναι ένας $n \times m$ πίνακας (b_{ij}) που έχει ως γραμμές τις στήλες του A (και ως στήλες τις γραμμές του A). Δηλαδή θα έχουμε $b_{ij} = a_{ji}$.

Παράδειγμα.

$$\text{Αν } A = \begin{pmatrix} -1 & 2 & 1 \\ 0 & -1 & 2 \end{pmatrix}, \quad \text{τότε } A^T = \begin{pmatrix} -1 & 0 \\ 2 & -1 \\ 1 & 2 \end{pmatrix}.$$

Ορισμός

Ένας πίνακας λέγεται **συμμετρικός** αν $A^T = A$.

Ένας συμμετρικός πίνακας θα είναι κατ' ανάγκην τετραγωνικός.

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

Αντίστροφος ενός πίνακα

Ορισμός

Έστω A ένας μη μηδενικός, τετραγωνικός $n \times n$ πίνακας. Ο A θα λέγεται **αντιστρέψιμος** αν υπάρχει ένας πίνακας B τέτοιος ώστε

$$AB = BA = I_n.$$

Αν υπάρχει ένας τέτοιος πίνακας B , τότε εύκολα μπορούμε να δούμε ότι είναι μοναδικός, θα τον συμβολίζουμε A^{-1} και θα τον αποκαλούμε **αντίστροφος** του A .

Παράδειγμα. Ο πίνακας

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

είναι αντιστρέψιμος, και ο αντίστροφός του είναι

$$A^{-1} = \begin{pmatrix} 1 & -1 \\ -1 & 2 \end{pmatrix}.$$

(Επαληθεύστε ότι $AA^{-1} = A^{-1}A = I_2$.)

Ιδιοτιμές και ιδιοδιανύσματα ενός πίνακα

Έστω A ένας μη μηδενικός, τετραγωνικός $n \times n$ πραγματικός πίνακας. Αν $x \in \mathbb{R}^n$ είναι ένα διάνυσμα, τότε το Ax είναι και αυτό ένα διάνυσμα του $\mathbb{R}^n$. Άρα μπορώ να θεωρήσω έναν πίνακα ως μια απεικόνιση από το $\mathbb{R}^n$ στον $\mathbb{R}^n$.

Στην περίπτωση που υπάρχει κάποιο μη μηδενικό διάνυσμα x που είναι παράλληλο στο Ax , τότε αυτό σημαίνει ότι θα υπάρχει μια μη μηδενική σταθερά λ τέτοια ώστε

$$Ax = \lambda x \Leftrightarrow (A - \lambda I)x = 0.$$

Ορισμός

Έστω A ένας μη μηδενικός, τετραγωνικός $n \times n$ πραγματικός πίνακας. Τότε ορίζουμε ως **χαρακτηριστικό πολυώνυμο** του A το πολυώνυμο

$$p(\lambda) = \det(A - \lambda I).$$

Είναι σχετικά εύκολο να δείξουμε ότι το p είναι ένα πολυώνυμο βαθμού n , συνεπώς, από ένα θεμελιώδες θεώρημα της Άλγεβρας γνωρίζουμε ότι έχει το πολύ n διακριτές ρίζες, που κάποιες από αυτές θα είναι μιγαδικές.

... Ιδιοτιμές και ιδιοδιανύσματα ενός πίνακα

Τώρα, αν λ είναι μια ρίζα του p , δηλ. $p(\lambda) = 0 \Leftrightarrow \det(A - \lambda I) = 0$, που σημαίνει ότι το ομογενές σύστημα $(A - \lambda I)x = 0$ θα έχει και μη μηδενικές λύσεις $x \neq 0$.

Ορισμός

- Αν p είναι το χαρακτηριστικό πολυώνυμο του πίνακα A , τότε οι ρίζες του p λέγονται **ιδιοτιμές** του πίνακα A .
- Αν λ είναι μια ιδιοτιμή του A και $x \neq 0$ ένα διάνυσμα που ικανοποιεί

$$(A - \lambda I)x = 0,$$

τότε λέμε ότι το x είναι ένα **ιδιοδιάνυσμα** που αντιστοιχεί στην ιδιοτιμή λ .

Για να υπολογίσουμε τις ιδιοτιμές ενός πίνακα χρησιμοποιούμε το γεγονός ότι

- λ είναι ιδιοτιμή του $A \Leftrightarrow \det(A - \lambda I) = 0$.

Στη συνέχεια, για να υπολογίσουμε το αντίστοιχο ιδιοδιάνυσμα $x \neq 0$

- λύνουμε το ομογενές σύστημα $(A - \lambda I)x = 0$.

Παράδειγμα 1

Βρείτε τις ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα $A = \begin{pmatrix} 1 & -1 \\ 2 & 4 \end{pmatrix}$.

Χαρακτηριστικό πολυώνυμο:

$$p(\lambda) = \det \begin{pmatrix} 1-\lambda & -1 \\ 2 & 4-\lambda \end{pmatrix} = (1-\lambda)(4-\lambda) + 2 = \lambda^2 - 5\lambda + 6.$$

$$\therefore p(\lambda) = 0 \Leftrightarrow \lambda^2 - 5\lambda + 6 = 0 \Leftrightarrow \lambda = 3 \text{ ή } \lambda = 2.$$

Άρα οι ιδιοτιμές είναι $\lambda = 3$ και $\lambda = 2$.

Για να βρω ιδιοδιάνυσμα που αντιστοιχεί στην ιδιοτιμή 3 λύνω το ομογενές σύστημα:

$$\begin{aligned} (A - 3I)x = 0 &\Leftrightarrow \begin{pmatrix} -2 & -1 \\ 2 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ &\Leftrightarrow 2x_1 + x_2 = 0 \Leftrightarrow x_2 = -2x_1. \end{aligned}$$

Άρα τα ιδιοδιανύσματα που αντιστοιχούν στην ιδιοτιμή 3 είναι της μορφής $\begin{pmatrix} x_1 \\ -2x_1 \end{pmatrix}$.

Παίρνοντας, π.χ., $x_1 = 1$ βρίσκουμε το ιδιοδιάνυσμα $\begin{pmatrix} 1 \\ -2 \end{pmatrix}$.

Στη συνέχεια, για να βρω ιδιοδιάνυσμα που αντιστοιχεί στην ιδιοτιμή 2 λύνω το ομογενές σύστημα:

$$\begin{aligned} (A - 2I)x = 0 &\Leftrightarrow \begin{pmatrix} -1 & -1 \\ 2 & 2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ &\Leftrightarrow x_1 + x_2 = 0 \Leftrightarrow x_2 = -x_1. \end{aligned}$$

Άρα τα ιδιοδιανύσματα που αντιστοιχούν στην ιδιοτιμή 2 είναι της μορφής $\begin{pmatrix} x_1 \\ -x_1 \end{pmatrix}$.

Παίρνοντας, π.χ., $x_1 = 1$ βρίσκουμε το ιδιοδιάνυσμα $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

(Φυσικά και κάθε πολλαπλάσιό του είναι και αυτό ιδιοδιάνυσμα.)

Παράδειγμα 2

Ελέγξτε αν ο πίνακας $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ έχει πραγματικές ιδιοτιμές και ιδιοδιανύσματα.

Χαρακτηριστικό πολυώνυμο:

$$p(\lambda) = \det \begin{pmatrix} -\lambda & 1 \\ -1 & -\lambda \end{pmatrix} = \lambda^2 + 1.$$

$$\therefore p(\lambda) = 0 \Leftrightarrow \lambda^2 + 1 = 0 \Leftrightarrow \lambda = i \text{ ή } \lambda = -i.$$

Άρα οι ιδιοτιμές είναι $\lambda = i$ και $\lambda = -i \notin \mathbb{R} !!$

Για να βρω ιδιοδιάνυσμα που αντιστοιχεί στην ιδιοτιμή i λύνω το ομογενές σύστημα:

$$(A - iI)x = 0 \Leftrightarrow \begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \\ \Leftrightarrow -ix_1 + x_2 = 0 \text{ και } -x_1 - ix_2 = 0 \Leftrightarrow x_1 = -ix_2 \quad (*).$$

Προφανώς, αν $x_1, x_2 \in \mathbb{R}$ η (*) είναι αδύνατη!

Αν όμως, π.χ., πάρουμε $x_2 = i$, τότε $x_1 = 1$ και το $\begin{pmatrix} 1 \\ i \end{pmatrix}$ είναι ένα (μιγαδικό) ιδιοδιάνυσμα που αντιστοιχεί στην ιδιοτιμή i . (Ελέγξτε ότι $\begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ i \end{pmatrix} = i \begin{pmatrix} 1 \\ i \end{pmatrix}$.)

Συστήματα γραμμικών εξισώσεων

Θέλουμε να λύσουμε συστήματα n γραμμικών εξισώσεων με n αγνώστους:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n = b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n = b_2 \\ \vdots \\ a_{n1}x_1 + a_{n2}x_2 + \cdots + a_{nn}x_n = b_n \end{cases}$$

όπου τα a_{ij} , και b_i , $1 \leq i \leq n$, $1 \leq j \leq n$, είναι δεδομένοι αριθμοί, και x_j είναι οι άγνωστοι που πρέπει να προσδιορίσουμε ώστε να ικανοποιούνται όλες οι εξισώσεις. Συνήθως τα a_{ij} λέγονται **συντελεστές του συστήματος**, ενώ τα b_i **σταθεροί όροι**.

Συστήματα γραμμικών εξισώσεων

Ισοδύναμα, μπορούμε να γράψουμε το σύστημα στη μορφή:

$$Ax = b,$$

όπου

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix},$$

είναι ο πίνακας των συντελεστών του συστήματος, και

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix},$$

είναι τα διανύσματα (στήλη) που περιέχουν τους άγνωστους και τους σταθερούς όρους, αντίστοιχα.

Ξέρουμε από τη Γραμμική Άλγεβρα, ότι αν ο πίνακας A είναι αντιστρέψιμος, δηλ. αν υπάρχει ένας πίνακας A^{-1} , τέτοιος ώστε

$$AA^{-1} = A^{-1}A = I_n,$$

τότε το σύστημα $Ax = b$ έχει **μοναδική λύση** $x = A^{-1}b$.

Όμως, τις περισσότερες φορές, δεν είναι απαραίτητο να υπολογίζουμε τον αντίστροφο του A για να λύσουμε το σύστημα.

Παράδειγμα: Έστω η εξίσωση

$$7x = 14.$$

Ο καλύτερος τρόπος να λύσω το σύστημα είναι, προφανώς, να κάνω μια διαίρεση

$$x = \frac{14}{7} = 2.$$

Αν λύσουμε την εξίσωση αυτή με τον “αντίστροφο πίνακα” θα έχουμε

$$x = \left(\frac{1}{7}\right) 14 = (0.142857) 14 = 1.999998.$$

(Απαιτούνται περισσότερες πράξεις και παράγεται μια λιγότερο ακριβής λύση.)

Ξέρουμε από τη Γραμμική Άλγεβρα ότι τα ακόλουθα είναι ισοδύναμα:

- Το σύστημα $Ax = b$ έχει μοναδική λύση.
- Ο πίνακας A είναι αντιστρέψιμος.
- $\det(A) \neq 0$.
- Το 0 δεν είναι ιδιοτιμή του πίνακα A .
- Το αντίστοιχο ομογενές σύστημα $Ax = 0$, έχει μόνο τη μηδενική λύση ($x = 0$).

Άμεσες Μέθοδοι – Απαλοιφή Gauss

Θέλουμε να λύσουμε το γραμμικό σύστημα $Ax = b$, όπου ο $n \times n$ πίνακας A είναι αντιστρέψιμος. Θα χρησιμοποιήσουμε **απαλοιφή Gauss**.

Παράδειγμα: Θεωρήστε το σύστημα

$$\begin{cases} 10x_1 - 7x_2 = 7 \\ -3x_1 + 2x_2 + 6x_3 = 4 \\ 5x_1 - x_2 + 5x_3 = 6 \end{cases}.$$

Ισοδύναμα,

$$\begin{pmatrix} 10 & -7 & 0 \\ -3 & 2 & 6 \\ 5 & -1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \\ 6 \end{pmatrix}$$

Γενικά, η απαλοιφή Gauss έχει δύο φάσεις:

- Τη φάση της τριγωνοποίησης, όπου μετατρέπουμε μέσω συγκεκριμένων πράξεων τον πίνακα A σε **άνω τριγωνικό**.
- Τη φάση της “οπισθοδρόμησης” (προς τα πίσω αντικατάστασης).

1η φάση: τριγωνποίηση

$$\underbrace{\begin{pmatrix} 10 & -7 & 0 \\ -3 & 2 & 6 \\ 5 & -1 & 5 \end{pmatrix}}_{A^{(1)}} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 7 \\ 4 \\ 6 \end{pmatrix}}_{b^{(1)}}$$

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

1ο βήμα της απαλοιφής: Θέλω να μηδενίσω τα στοιχεία του πίνακα που βρίσκονται κάτω από το στοιχείο (1, 1), δηλ. κάτω από το 10.

- Πολλαπλασιάζω την πρώτη εξίσωση επί $m_{21} = \frac{-3}{10} = -0.3$ και την αφαιρώ από τη δεύτερη.
- Πολλαπλασιάζω την πρώτη εξίσωση επί $m_{31} = \frac{5}{10} = 0.5$ και την αφαιρώ από την τρίτη.

$$\underbrace{\begin{pmatrix} 10 & -7 & 0 \\ 0 & -0.1 & 6 \\ 0 & 2.5 & 5 \end{pmatrix}}_{A^{(2)}} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 7 \\ 6.1 \\ 2.5 \end{pmatrix}}_{b^{(2)}}$$

$$A^{(2)}x = b^{(2)}$$

$$\begin{pmatrix} 10 & -7 & 0 \\ 0 & -0.1 & 6 \\ 0 & 2.5 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 6.1 \\ 2.5 \end{pmatrix}$$

2ο βήμα της απαλοιφής: Θέλω να μηδενίσω τα στοιχεία του πίνακα που βρίσκονται κάτω από το στοιχείο (2, 2), δηλ. κάτω από το -0.1 .

- Πολλαπλασιάζω τη δεύτερη εξίσωση επί $m_{32} = \frac{2.5}{-0.1} = -25$ και την αφαιρώ από την τρίτη.

$$\underbrace{\begin{pmatrix} 10 & -7 & 0 \\ 0 & -0.1 & 6 \\ 0 & 0 & 155 \end{pmatrix}}_{A^{(3)}} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \underbrace{\begin{pmatrix} 7 \\ 6.1 \\ 155 \end{pmatrix}}_{b^{(3)}}$$

$$A^{(1)}x = b^{(1)} \Leftrightarrow A^{(2)}x = b^{(2)} \Leftrightarrow A^{(3)}x = b^{(3)}$$

Παρατηρήστε ότι ο πίνακας $A^{(3)}$ είναι άνω τριγωνικός. Ολοκληρώθηκε η φάση της τριγωνποίησης.

2η φάση: οπισθοδρόμηση

$$\begin{pmatrix} 10 & -7 & 0 \\ 0 & -0.1 & 6 \\ 0 & 0 & 155 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 6.1 \\ 155 \end{pmatrix}$$

- Λύνουμε την τελευταία εξίσωση ως προς x_3 :

$$155 x_3 = 155 \Rightarrow x_3 = 1.$$

- Αντικαθιστούμε στη δεύτερη, και λύνουμε ως προς x_2 :

$$-0.1 x_2 + 6 \cdot 1 = 6.1 \Rightarrow x_2 = -1.$$

- Αντικαθιστούμε στην πρώτη, και λύνουμε ως προς x_1 :

$$10 x_1 - 7(-1) + 0 \cdot 1 = 7 \Rightarrow 10 x_1 = 0 \Rightarrow x_1 = 0.$$

Ο αλγόριθμος σε γλώσσα πινάκων

- Θέτουμε $M_1 = \begin{pmatrix} 1 & 0 & 0 \\ 0.3 & 1 & 0 \\ -0.5 & 0 & 1 \end{pmatrix}$. Τότε

$$M_1 A = \begin{pmatrix} 10 & -7 & 0 \\ 0 & -0.1 & 6 \\ 0 & 2.5 & 5 \end{pmatrix} \quad \text{και} \quad M_1 b = \begin{pmatrix} 7 \\ 6.1 \\ 2.5 \end{pmatrix}$$

- Θέτουμε $M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 25 & 1 \end{pmatrix}$. Τότε

$$M_2 M_1 A = \begin{pmatrix} 10 & -7 & 0 \\ 0 & -0.1 & 6 \\ 0 & 0 & 155 \end{pmatrix} \quad \text{και} \quad M_2 M_1 b = \begin{pmatrix} 7 \\ 6.1 \\ 155 \end{pmatrix}$$

Αν θέσουμε $M_2 M_1 = M$, και τον τελικό άνω τριγωνικό πίνακα $A^{(3)} = U$, τότε

$$MA = U \Leftrightarrow A = \underbrace{M^{-1}}_L U \Leftrightarrow A = LU, \quad (\text{Ανάλυση LU})$$

όπου ο πίνακας L είναι κάτω τριγωνικός και ο πίνακας U άνω τριγωνικός.

Τα στοιχεία του πίνακα M που βρίσκονται κάτω από τη διαγώνιο είναι οι πολλαπλασιαστές $-m_{ij}$, ενώ τα διαγώνια στοιχεία του πίνακα U λέγονται **οδηγοί**.

Στο παράδειγμά μας οι οδηγοί είναι 10, -0.1 και 155.

Χρησιμότητα ανάλυσης LU

Έστω A τετραγωνικός ($n \times n$) αντιστρέψιμος πίνακας. Τότε υπάρχουν πίνακες L, U με L κάτω τριγωνικό (με μονάδες στη διαγώνιο) και U άνω τριγωνικό:

$$A = LU.$$

Στόχος: Θέλω να λύσω το γραμμ. συστ. $Ax = b$

$Ax = b \Leftrightarrow LUx = b$

① Θέτω $Ux = y$: $L(Ux) = b \Rightarrow L y = b$ ↓ κάτω τριγων.

→ Βρίσκω το y .
(λύνοντας από πάνω προς τα κάτω)

$\begin{pmatrix} \text{κάτω τριγων.} & 0 \end{pmatrix} \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} b_1 \\ \vdots \\ b_n \end{pmatrix}$
 $L \cdot y = b$

② $Ux = y$ ↑ άνω τριγωνικός

$\begin{pmatrix} 0 & \text{άνω τριγων.} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ → Βρίσκω το x .
(λύνοντας από κάτω προς τα πάνω)

$U \cdot x = y$

Απαλοιφή Gauss με μερική οδήγηση

Παρατηρήστε ότι οι οδηγοί εμφανίζονται ως παρονομαστές τόσο στους πολλαπλασιαστές όσο και κατά την οπισθοδρομηση στον υπολογισμό των x_i . Είναι λογικό να αναμένουμε προβλήματα (λόγω της πεπερασμένης ακρίβειας στον Η/Υ) όταν κάποιος οδηγός είναι σχεδόν μηδέν, δηλαδή πολύ μικρός κατ' απόλυτη τιμή.

Παράδειγμα: Θεωρήστε το σύστημα

$$\begin{cases} 10^{-4}x_1 + x_2 = 1 \\ x_1 + x_2 = 2 \end{cases} \quad (\text{Μοναδική λύση: } x_1 = 1.0001 \dots, x_2 = 0.9998 \dots)$$

Λύνουμε αυτό το σύστημα με απαλοιφή Gauss σε έναν Η/Υ με $\beta = 10$, $t = 3$, $U = -L = 20$ και στρογγύλευση.

$$\begin{aligned} m_{21} &= 1/10^{-4} = 10^4, \\ a_{22}^{(2)} &= \text{fl}(1 - m_{21} \cdot 1) = \text{fl}(1 - 10^4) = \text{fl}(-0.9999 \cdot 10^4) = -10^4, \\ b_2^{(2)} &= \text{fl}(2 - m_{21} \cdot 1) = \text{fl}(2 - 10^4) = \text{fl}(-0.9998 \cdot 10^4) = -10^4, \\ \begin{cases} 10^{-4}x_1 + x_2 = 1 \\ -10^4x_2 = -10^4 \end{cases} & \quad (\text{Μοναδική λύση: } x_1 = 0, x_2 = 1 !!!) \end{aligned}$$

Εναλλακτικά, προσπαθούμε να κάνουμε τους υπολογισμούς με μικρό πολλαπλασιαστή m_{21} , δηλαδή με όσο το δυνατόν μεγαλύτερο οδηγό.

Κάνουμε εναλλαγή γραμμών, και θεωρούμε το σύστημα:

$$\begin{cases} x_1 + x_2 = 2 \\ 10^{-4}x_1 + x_2 = 1 \end{cases}.$$

Κάνοντας απαλοιφή Gauss στον ίδιο Η/Υ (όπως πριν) παίρνουμε

$$\begin{cases} x_1 + x_2 = 2 \\ 1 \cdot x_2 = 1 \end{cases} \Rightarrow x_1 = 1, x_2 = 1$$

Μερική οδήγηση ή οδήγηση κατά γραμμές: Στο r στάδιο της τριγωνοποίησης εξετάζουμε τα στοιχεία $a_{kr}^{(r)}$, $k = r, r+1, \dots, n$, δηλαδή αυτά που βρίσκονται στην r στήλη, από το διαγώνιο στοιχείο $a_{rr}^{(r)}$ και κάτω, βρίσκουμε ένα από αυτά που έχει τη μεγαλύτερη απόλυτη τιμή και με εναλλαγή γραμμών το φέρνουμε στη θέση του οδηγού.

Παράδειγμα απαλοιφής Gauss με μερική οδήγηση

Θεωρήστε και πάλι το σύστημα

$$\begin{pmatrix} 10 & -7 & 0 \\ -3 & 2 & 6 \\ 5 & -1 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 4 \\ 6 \end{pmatrix}$$

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

Ελέγχουμε την πρώτη στήλη. Παρατηρούμε ότι το μεγαλύτερο κατ' απόλυτη τιμή στοιχείο είναι το 10 και βρίσκεται στη θέση (1, 1), άρα προχωρούμε χωρίς εναλλαγές γραμμών, και καταλήγουμε (όπως πριν) στο σύστημα:

$$\begin{pmatrix} 10 & -7 & 0 \\ 0 & -0.1 & 6 \\ 0 & 2.5 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 6.1 \\ 2.5 \end{pmatrix}$$

$A^{(2)}$

$b^{(2)}$

$$A^{(2)}x = b^{(2)}$$

$$\begin{pmatrix} 10 & -7 & 0 \\ 0 & -0.1 & 6 \\ 0 & 2.5 & 5 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 6.1 \\ 2.5 \end{pmatrix}$$

Ελέγχουμε τα στοιχεία $a_{22}^{(2)} = -0.1$ και $a_{23}^{(2)} = 2.5$. Το μεγαλύτερο κατ' απόλυτη τιμή είναι το 2.5, οπότε εναλλάσσουμε τις γραμμές 2 και 3.

$$\begin{pmatrix} 10 & -7 & 0 \\ 0 & 2.5 & 5 \\ 0 & -0.1 & 6 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 2.5 \\ 6.1 \end{pmatrix}$$

Τώρα θέλω να μηδενίσω τα στοιχεία του πίνακα που βρίσκονται κάτω από το στοιχείο (2, 2), δηλ. κάτω από το 2.5.

- Πολλαπλασιάζω τη δεύτερη εξίσωση επί $m_{32} = \frac{-0.1}{2.5} = 0.04$ και την αφαιρώ από την τρίτη.

$$\begin{pmatrix} 10 & -7 & 0 \\ 0 & 2.5 & 5 \\ 0 & 0 & 6.2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 2.5 \\ 6.2 \end{pmatrix}$$

$A^{(3)}$

$b^{(3)}$

$$A^{(3)}x = b^{(3)}$$

2η φάση: οπισθοδρόμηση

$$\begin{pmatrix} 10 & -7 & 0 \\ 0 & 2.5 & 5 \\ 0 & 0 & 6.2 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \end{pmatrix} = \begin{pmatrix} 7 \\ 2.5 \\ 6.2 \end{pmatrix}$$

- Λύνουμε την τελευταία εξίσωση ως προς x_3 :

$$6.2 x_3 = 6.2 \Rightarrow x_3 = 1.$$

- Αντικαθιστούμε στη δεύτερη, και λύνουμε ως προς x_2 :

$$2.5 x_2 + 5 \cdot 1 = 2.5 \Rightarrow x_2 = -1.$$

- Αντικαθιστούμε στην πρώτη, και λύνουμε ως προς x_1 :

$$10 x_1 - 7(-1) + 0 \cdot 1 = 7 \Rightarrow 10 x_1 = 0 \Rightarrow x_1 = 0.$$

Απαλοιφή Gauss (“στο χαρτί”) με χρήση του επαυξημένου πίνακα

Παράδειγμα. Να λυθεί το σύστημα

$$\begin{cases} x_1 + 2x_2 - x_3 = -2 \\ x_1 + x_3 = 4 \\ 2x_1 + x_2 + 2x_3 = -1 \end{cases}.$$

Ο επαυξημένος πίνακας του συστήματος είναι

$$(A|b) = \left(\begin{array}{ccc|c} 1 & 2 & -1 & -2 \\ 1 & 0 & 1 & 4 \\ 2 & 1 & 2 & -1 \end{array} \right).$$

Με απαλοιφή Gauss με μερική οδήγηση θα φέρουμε τον επαυξημένο πίνακα στην κατάλληλη μορφή μετατρέποντας τον A σε άνω τριγωνικό.

$$\begin{aligned}
(A|b) &= \left(\begin{array}{ccc|c} 1 & 2 & -1 & -2 \\ 1 & 0 & 1 & 4 \\ 2 & 1 & 2 & -1 \end{array} \right) \\
&\xrightarrow[r_1 \leftrightarrow r_3]{} \left(\begin{array}{ccc|c} 2 & 1 & 2 & -1 \\ 1 & 0 & 1 & 4 \\ 1 & 2 & -1 & -2 \end{array} \right) \quad \left[\begin{array}{l} m_{21} = 1/2 \\ m_{31} = 1/2 \end{array} \right] \\
&\xrightarrow[r_2 \rightarrow r_2 - m_{21}r_1]{r_3 \rightarrow r_3 - m_{31}r_1} \left(\begin{array}{ccc|c} 2 & 1 & 2 & -1 \\ 0 & -1/2 & 0 & 9/2 \\ 0 & 3/2 & -2 & -3/2 \end{array} \right) \\
&\xrightarrow[r_2 \leftrightarrow r_3]{} \left(\begin{array}{ccc|c} 2 & 1 & 2 & -1 \\ 0 & 3/2 & -2 & -3/2 \\ 0 & -1/2 & 0 & 9/2 \end{array} \right) \quad \left[m_{32} = \frac{-\frac{1}{2}}{\frac{3}{2}} = -\frac{1}{3} \right] \\
&\xrightarrow[r_3 \rightarrow r_3 - m_{32}r_2]{} \left(\begin{array}{ccc|c} 2 & 1 & 2 & -1 \\ 0 & 3/2 & -2 & -3/2 \\ 0 & 0 & -2/3 & 4 \end{array} \right) = (A'|b').
\end{aligned}$$

Άρα καταλήξαμε στο ισοδύναμο σύστημα

$$\begin{cases} 2x_1 + x_2 + 2x_3 &= -1 \\ \frac{3}{2}x_2 - 2x_3 &= -\frac{3}{2} \\ -\frac{2}{3}x_3 &= 4 \end{cases}.$$

- Λύνουμε την τελευταία ως προς x_3 : $x_3 = -6$.
- Αντικαθιστούμε στη δεύτερη, και λύνουμε ως προς x_2 : $x_2 = -9$.
- Αντικαθιστούμε στην πρώτη, και λύνουμε ως προς x_1 : $x_1 = 10$.

Κατάσταση γραμμικών συστημάτων

Κατά την επίλυση ενός γραμμικού συστήματος της μορφής $Ax = b$ (με A αντιστρέψιμο πίνακα) στον Η/Υ είναι συχνά αναπόφευκτο, λόγω των σφαλμάτων στρογγύλευσης, να υπολογίζουμε μια λύση $\tilde{x}$, η οποία θα διαφέρει κάπως από την ακριβή λύση $x = A^{-1}b$.

Π.χ., αν οι συνιστώσες της λύσης x δεν είναι αριθμοί μηχανής, τότε το $\tilde{x} \neq x$.

Για να εκτιμήσουμε τη διαφορά αυτή χρησιμοποιούμε δύο ποσότητες:

α) το **σφάλμα**: $e = x - \tilde{x}$, και

β) το **υπόλοιπο**: $r = b - A\tilde{x}$.

Δεν είναι δύσκολο να αποδείξουμε ότι, αφού ο A είναι αντιστρέψιμος, αν το σφάλμα είναι μηδέν, τότε και το υπόλοιπο θα είναι μηδέν, και αντίστροφα. **Από την άλλη μεριά όμως δεν είναι απαραίτητα και τα δύο ταυτοχρόνως «μικρά».**

Παράδειγμα

Θεωρήστε το γραμμικό σύστημα

$$\begin{pmatrix} 0.913 & 0.659 \\ 0.780 & 0.563 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.254 \\ 0.217 \end{pmatrix}.$$

Το σύστημα αυτό έχει μοναδική λύση $x = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$.

Απαλοιφή Gauss με μερική οδήγηση σε έναν Η/Υ με $\beta = 10$, $t = 3$, $U = -L = 20$ και αποκοπή. Έχουμε

$$m_{21} = \frac{0.780}{0.913} = 0.854,$$

$$a_{22}^{(2)} = \text{fl}(0.780 - \text{fl}(m_{21} \cdot 0.913)) = \text{fl}(0.780 - 0.779) = 0.001,$$

$$b_2^{(2)} = \text{fl}(0.217 - \text{fl}(m_{21} \cdot 0.254)) = \text{fl}(0.217 - 0.216) = 0.001.$$

Καταλήξαμε δηλαδή στο σύστημα

$$\begin{pmatrix} 0.913 & 0.659 \\ 0 & 0.001 \end{pmatrix} \begin{pmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{pmatrix} = \begin{pmatrix} 0.254 \\ 0.001 \end{pmatrix}.$$

Λύνουμε με οπισθοδρόμηση και έχουμε

$$\begin{aligned} \tilde{x}_2 &= \frac{0.001}{0.001} = 1 \\ \tilde{x}_1 &= \text{fl} \left(\frac{0.254 - \text{fl}(0.659x_2)}{0.913} \right) = -0.443. \end{aligned}$$

Δηλαδή χρησιμοποιήσαμε έναν ευσταθή αλγόριθμο και καταλήξαμε σε μια εντελώς ανακριβή προσεγγιστική λύση $\tilde{x}$.

Δεν είναι δύσκολο να ελέγξουμε ότι αν δουλεύαμε σε έναν υπολογιστή με μεγαλύτερη ακρίβεια, $t = 6$, θα παίρναμε την ακριβή λύση $x = (1, -1)^T$.

Ακρίβεια λύσης

Υπόλοιπο:

$$\begin{aligned} r = b - A\tilde{x} &= \begin{pmatrix} 0.254 - (0.913)(-0.443) - (0.659)(1.00) \\ 0.217 - (0.780)(-0.443) - (0.563)(1.00) \end{pmatrix} \\ &= \begin{pmatrix} -0.000541 \\ -0.000460 \end{pmatrix}. \end{aligned}$$

(Εκτελέσαμε τις πράξεις ακριβώς.)

Οι συνιστώσες του υπολοίπου είναι μικρές, για την ακρίβεια, μικρότερες από 10^{-3} .

Σφάλμα:

$$e = x - \tilde{x} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} - \begin{pmatrix} -0.443 \\ 1.000 \end{pmatrix} = \begin{pmatrix} 1.443 \\ -2.000 \end{pmatrix},$$

Το σφάλμα είναι μεγαλύτερο από την ίδια τη λύση.

Τι φταίει;

Ας κρατήσουμε σταθερό τον πίνακα του συστήματος και ας μεταβάλουμε λίγο το δεύτερο μέλος. Συγκεκριμένα, θεωρούμε το ελαφρά διαταραγμένο σύστημα

$$\begin{pmatrix} 0.913 & 0.659 \\ 0.780 & 0.563 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 0.253 \\ 0.218 \end{pmatrix}.$$

Η ακριβής λύση αυτού του συστήματος είναι $x = \begin{pmatrix} 1223 \\ -1694 \end{pmatrix}$.

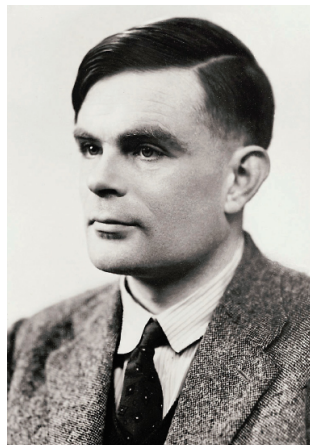
Βλέπουμε δηλαδή ότι μια «μικρή» μεταβολή στα δεδομένα του προβλήματος προκαλεί μια «μεγάλη» μεταβολή στη λύση του προβλήματος.

Άρα το πρόβλημα αυτό είναι **πολύ ευαίσθητο σε μικρές μεταβολές των δεδομένων του**, και για το λόγο αυτό λέμε ότι έχει **κακή κατάσταση**.

Παρατήρηση

Αποδεικνύεται ότι:

Η απαλοιφή Gauss με μερική οδήγηση παράγει πάντα προσεγγιστικές λύσεις με μικρό υπόλοιπο. Όπως όμως είδαμε, **ένα μικρό υπόλοιπο r δεν μας εξασφαλίζει ότι και το σφάλμα e είναι μικρό**. Η σχέση μεταξύ του υπολοίπου και του σφάλματος καθορίζεται από μια ποσότητα που λέγεται **δείκτης κατάστασης** του πίνακα



Alan Turing (1912 – 1954)

Νόρμα διανύσματος

Για να μπορέσουμε να μετρήσουμε το μέγεθος των στοιχείων ενός διανύσματος χρειαζόμαστε την έννοια της **νόρμας**.

Ορισμός

Μια διανυσματική νόρμα είναι μια απεικόνιση από τον $\mathbb{R}^n$ στο $\mathbb{R}$, η οποία ικανοποιεί τις ακόλουθες τρεις ιδιότητες:

- (1) $\|x\| = 0 \Leftrightarrow x = 0$,
- (2) $\|\lambda x\| = |\lambda| \|x\|$, για κάθε αριθμό λ ,
- (3) $\|x + y\| \leq \|x\| + \|y\|$, (τριγωνική ανισότητα),

όπου $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix}$, $y = \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix}$ είναι διανύσματα του $\mathbb{R}^n$ και $\lambda \in \mathbb{R}$.

Παραδείγματα νορμών

Έστω $x = \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \in \mathbb{R}^n$. Τότε οι ακόλουθες είναι νόρμες

$$\|x\|_1 = \sum_{i=1}^n |x_i|, \quad \ell_1\text{-νόρμα},$$

$$\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|, \quad \ell_\infty\text{-νόρμα ή νόρμα μεγίστου}$$

$$\|x\|_2 = \left(\sum_{i=1}^n |x_i|^2 \right)^{1/2} \quad \ell_2\text{-νόρμα ή Ευκλείδεια νόρμα}.$$

Παράδειγμα

Έστω $x = \begin{pmatrix} -3 \\ 1 \\ -2 \end{pmatrix}$. Τότε

$$\|x\|_1 = \sum_{i=1}^3 |x_i| = |-3| + |1| + |-2| = 3 + 1 + 2 = 6.$$

$$\|x\|_\infty = \max_{1 \leq i \leq 3} |x_i| = 3.$$

$$\|x\|_2 = \left(\sum_{i=1}^3 |x_i|^2 \right)^{1/2} = \sqrt{(-3)^2 + 1^2 + (-2)^2} = \sqrt{14} \approx 3.742.$$

Νόρμες πινάκων

Έστω $\mathbb{R}^{n \times n}$ το σύνολο των τετραγωνικών $n \times n$ πινάκων με στοιχεία πραγματικούς αριθμούς.

Ορισμός

Μια νόρμα πινάκων είναι μια απεικόνιση από τον $\mathbb{R}^{n \times n}$ στο $\mathbb{R}$, η οποία ικανοποιεί τις ακόλουθες τρεις ιδιότητες:

- (1) $\|A\| = 0 \Leftrightarrow A = 0$,
- (2) $\|\lambda A\| = |\lambda| \|A\|$, για κάθε αριθμό λ ,
- (3) $\|A + B\| \leq \|A\| + \|B\|$, (τριγωνική ανισότητα),

και, επιπλέον, πληροί την ιδιότητα:

$$(4) \quad \|AB\| \leq \|A\| \|B\|,$$

όπου $A, B \in \mathbb{R}^{n \times n}$ και $\lambda \in \mathbb{R}$.

Φυσική νόρμα πίνακα

Ορισμός

Έστω $\|\cdot\|$ μια νόρμα στον $\mathbb{R}^n$. Τότε η αντίστοιχη **φυσική νόρμα πινάκων** ή νόρμα πινάκων που παράγεται από τη νόρμα $\|\cdot\|$ του $\mathbb{R}^n$ ορίζεται ως η απεικόνιση

$$\|\cdot\| : \mathbb{R}^{n \times n} \rightarrow \mathbb{R}, \quad \|A\| = \sup_{0 \neq x \in \mathbb{R}^n} \frac{\|Ax\|}{\|x\|}. \quad (1)$$

Έστω $X \subset \mathbb{R}$. Τότε ως **supremum** του συνόλου X ορίζουμε το **ελάχιστο άνω φράγμα** του X , και το συμβολίζουμε ως $\sup X$.

Συγκεκριμένα, ένας αριθμός $\beta = \sup X$ αν για κάθε $x \in X$ ισχύει $x \leq \beta$ και για κάθε $\beta' < \beta$ υπάρχει ένα $x' \in X$ τέτοιο ώστε $x' > \beta'$.

Π.χ., $\sup[a, b] = b$, $\sup(a, b) = b$, $\sup\{x \in \mathbb{Q} : x^2 < 2\} = \sqrt{2}$.

Φυσική νόρμα πίνακα

Πρόταση

Αν $\|\cdot\|$ είναι μια φυσική νόρμα πίνακα, τότε ισχύουν οι ακόλουθες ιδιότητες:

$$\|I\| = 1, \text{ όπου } I \text{ είναι ο μοναδιαίος } n \times n \text{ πίνακας.} \quad (2)$$

$$\|Ax\| \leq \|A\|\|x\|, \text{ για κάθε } x \in \mathbb{R}^n. \quad (3)$$

Παραδείγματα φυσικών νορμών πινάκων

- ❶ Θεωρούμε στον $\mathbb{R}^n$ τη νόρμα μεγίστου $\|\cdot\|_\infty$. Η φυσική νόρμα στον $\mathbb{R}^{n \times n}$ που παράγεται από την $\|\cdot\|_\infty$ είναι η λεγόμενη «νόρμα του αθροίσματος γραμμών» και για $A \in \mathbb{R}^{n \times n}$ ορίζεται ως

$$\|A\|_\infty = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|. \quad (4)$$

- ❷ Θεωρούμε στον $\mathbb{R}^n$ την ℓ_1 -νόρμα $\|\cdot\|_1$. Η φυσική νόρμα στον $\mathbb{R}^{n \times n}$ που παράγεται από την $\|\cdot\|_1$ είναι η λεγόμενη «νόρμα του αθροίσματος στηλών» και για $A \in \mathbb{R}^{n \times n}$ ορίζεται ως

$$\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|. \quad (5)$$

- ❸ Θεωρούμε στον $\mathbb{R}^n$ την Ευκλείδεια νόρμα $\|\cdot\|_2$. Η φυσική νόρμα στον $\mathbb{R}^{n \times n}$ που παράγεται από την $\|\cdot\|_2$ είναι η λεγόμενη «φασματική νόρμα» και για $A \in \mathbb{R}^{n \times n}$ ορίζεται ως

$$\|A\|_2 = \sqrt{\rho(A^T A)}, \quad (6)$$

όπου για έναν πίνακα $B \in \mathbb{R}^{n \times n}$ συμβολίζουμε με $\rho(B)$ την λεγόμενη *φασματική ακτίνα* του B , που ορίζεται ως το μέγιστο των απολύτων τιμών των ιδιοτιμών του πίνακα B .

Παράδειγμα

Έστω $A = \begin{pmatrix} -3 & 1 & -2 \\ 0 & 4 & 6 \\ -5 & 2 & -3 \end{pmatrix}$. Τότε

$$\begin{aligned} \|A\|_1 &= \max\{|-3| + 0 + |-5|, 1 + 4 + 2, |-2| + 6 + |-3|\} \\ &= \max\{8, 7, 11\} = 11. \end{aligned}$$

$$\begin{aligned} \|A\|_\infty &= \max\{|-3| + 1 + |-2|, 0 + 4 + 6, |-5| + 2 + |-3|\} \\ &= \max\{6, 10, 10\} = 10. \end{aligned}$$

Ορισμός

Ορίζουμε ως **δείκτη κατάστασης του πίνακα** A , και συμβολίζουμε με $\kappa(A)$, την ποσότητα:

$$\kappa(A) = \|A\| \|A^{-1}\|. \quad (7)$$

Από τον ορισμό του δείκτη κατάστασης συμπεραίνουμε εύκολα ότι

$$\kappa(A) \geq 1.$$

Πράγματι, έχουμε: $1 = \|I\| = \|AA^{-1}\| \leq \|A\| \|A^{-1}\| = \kappa(A)$.

Ο δείκτης κατάστασης ενός πίνακα είναι ένα μέτρο του πόσο κοντά βρίσκεται ο πίνακας αυτός στο σύνολο των μη αντιστρέψιμων πινάκων.

Αν ο $\kappa(A)$ είναι μεγάλος, τότε ο A είναι κοντά στο σύνολο των μη αντιστρέψιμων πινάκων, και τότε είναι λογικό να αναμένουμε δυσκολίες κατά την επίλυση του γραμμικού συστήματος $Ax = b$.

Κατάσταση συστημάτων

Θεωρούμε το σύστημα $Ax = b$ (A αντιστρέψιμος), και το σύστημα που προκύπτει αν διαταράξουμε το δεύτερο μέλος:

$$A(x + \delta x) = b + \delta b \Leftrightarrow Ax + A\delta x = b + \delta b.$$

Αφαιρούμε κατά μέλη και καταλήγουμε στο σύστημα: $A\delta x = \delta b \Leftrightarrow \delta x = A^{-1}\delta b$.

Συμβολίζουμε $\|\cdot\|$ μια νόρμα στον $\mathbb{R}^n$, καθώς και την αντίστοιχη φυσική νόρμα πίνακα. Τότε

$$Ax = b \Rightarrow \|b\| = \|Ax\| \stackrel{(3)}{\leq} \|A\| \|x\| \Rightarrow \frac{1}{\|x\|} \leq \|A\| \frac{1}{\|b\|},$$

$$\text{και } \delta x = A^{-1}\delta b \Rightarrow \|\delta x\| = \|A^{-1}\delta b\| \stackrel{(3)}{\leq} \|A^{-1}\| \|\delta b\| \Rightarrow \|\delta x\| \leq \|A^{-1}\| \|\delta b\|.$$

Πολλαπλασιάζουμε κατά μέλη τις δύο αυτές σχέσεις και έχουμε:

$$\frac{\|\delta x\|}{\|x\|} \leq \|A\| \|A^{-1}\| \frac{\|\delta b\|}{\|b\|} \Leftrightarrow \frac{\|\delta x\|}{\|x\|} \leq \kappa(A) \frac{\|\delta b\|}{\|b\|}.$$

Το κλάσμα $\|\delta b\|/\|b\|$ είναι η σχετική διαταραχή που επιφέραμε στο δεύτερο μέλος, και το κλάσμα $\|\delta x\|/\|x\|$ είναι το σχετικό σφάλμα εξαιτίας της διαταραχής αυτής. Ο δείκτης κατάστασης είναι ένας συντελεστής μεγέθυνσης της σχετικής διαταραχής. (Το ίδιο συμβαίνει και αν διαταράξουμε τα στοιχεία του πίνακα.)

Παράδειγμα

Θεωρήστε το γραμμικό σύστημα $Ax = b$, με

$$A = \begin{pmatrix} 4.1 & 2.8 \\ 9.7 & 6.6 \end{pmatrix} \quad \text{και} \quad b = \begin{pmatrix} 4.1 \\ 9.7 \end{pmatrix}.$$

Το σύστημα αυτό έχει μοναδική λύση $x = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$. Θα χρησιμοποιήσουμε τη νόρμα $\|\cdot\|_1$. Έχουμε

$$\|b\|_1 = 13.8 \quad \text{και} \quad \|x\|_1 = 1.$$

Αν μεταβάλλουμε λίγο το δεύτερο μέλος του συστήματος σε $\tilde{b} = \begin{pmatrix} 4.11 \\ 9.70 \end{pmatrix}$, τότε η ακριβής λύση του συστήματος είναι $\tilde{x} = \begin{pmatrix} 0.34 \\ 0.97 \end{pmatrix}$. Θέτουμε

$$\delta b = b - \tilde{b} = \begin{pmatrix} -0.01 \\ 0 \end{pmatrix} \quad \text{και} \quad \delta x = x - \tilde{x} = \begin{pmatrix} 0.66 \\ -0.97 \end{pmatrix}.$$

Τότε

$$\|\delta b\|_1 = 0.01 \quad \text{και} \quad \|\delta x\|_1 = 1.63.$$

Αν και μεταβάλαμε ελάχιστα το b βλέπουμε ότι το x αλλάζει τελείως. Συγκεκριμένα, οι σχετικές μεταβολές είναι

$$\frac{\|\delta b\|_1}{\|b\|_1} = 0.0007246 \quad \text{και} \quad \frac{\|\delta x\|_1}{\|x\|_1} = 1.63.$$

Επειδή το $\kappa_1(A)$ είναι ο μέγιστος συντελεστής μεγέθυνσης έχουμε ότι

$$\kappa_1(A) \geq \frac{1.63}{0.0007246} = 2249.4,$$

όπου υπολογίσαμε το δείκτη κατάστασης ως προς τη νόρμα $\|\cdot\|_1$.

(Στο συγκεκριμένο παράδειγμα επιλέξαμε τα b και δb που δίνουν το μέγιστο, επομένως εδώ έχουμε ότι $\kappa_1(A) = 2249.4$.)

Παράδειγμα στο MATLAB

```
>> H = hilb(12);  
>> x = ones(12,1);  
>> b = H*x;  
>> xapp = H\b  
xapp =  
    1.0000  
    1.0000  
    0.9999  
    1.0013  
    0.9905  
    1.0426  
    0.8790  
    1.2225  
    0.7355  
    1.1962  
    0.9175  
    1.0150
```

```
>> norm(x - xapp)/norm(x)  
ans =  
    0.1230  
>> norm(b - H*xapp)  
ans =  
    5.2074e-16  
>> cond(H)  
ans =  
    1.6776e+16
```

Τριδιαγώνια συστήματα

Σε πολλές εφαρμογές απαιτείται να επιλύσουμε γραμμικά συστήματα όπου ο πίνακας του συστήματος είναι **αραιός**, δηλαδή τα περισσότερα στοιχεία του είναι μηδέν.

Εμείς θα μελετήσουμε **τριδιαγώνια συστήματα**, δηλαδή συστήματα όπου τα μη μηδενικά στοιχεία του πίνακα του συστήματος βρίσκονται στη διαγώνιο και στην υπερ- και υποδιαγώνιό του.

Έστω ότι θέλουμε να λύσουμε το σύστημα

$$Ax = f, \quad f \in \mathbb{R}^n, \quad (8)$$

όπου A είναι ο τριδιαγώνιος $n \times n$ πίνακας

$$A = \begin{pmatrix} a_1 & b_1 & & & 0 \\ c_2 & a_2 & b_2 & & \\ & \ddots & \ddots & \ddots & \\ & & c_{n-1} & a_{n-1} & b_{n-1} \\ 0 & & & c_n & a_n \end{pmatrix}, \quad (9)$$

με $a_i, b_i, c_i \in \mathbb{R}$. Υποθέτουμε, επιπλέον, ότι

$$a_i, b_i, c_i \neq 0, \quad |a_1| > |b_1|, \quad |a_n| > |c_n|, \quad \text{και} \quad |a_i| \geq |b_i| + |c_i|, \quad 2 \leq i \leq n-1. \quad (10)$$

Θ.δ.ο. ο A αντιστρέφεται και μπορεί να αναλυθεί στη μορφή

$$A = LU, \quad (11)$$

όπου L και U είναι οι διδιαγώνιοι πίνακες

$$L = \begin{pmatrix} d_1 & & & & 0 \\ c_2 & d_2 & & & \\ & \ddots & \ddots & \ddots & \\ & & c_{n-1} & d_{n-1} & \\ 0 & & & c_n & d_n \end{pmatrix}, \quad U = \begin{pmatrix} 1 & e_1 & & & 0 \\ & 1 & e_2 & & \\ & & \ddots & \ddots & \\ & & & 1 & e_{n-1} \\ 0 & & & & 1 \end{pmatrix}.$$

Η ανάλυση αυτή είναι μοναδική.

Θέλουμε να υπολογίσουμε τα άγνωστα στοιχεία $d_i, i = 1, \dots, n$ και $e_i, i = 1, \dots, n-1$.

Εξισώνουμε τα αντίστοιχα μη μηδενικά στοιχεία που βρίσκονται στη διαγώνιο και στην υπερδιαγώνιο του αριστερού και δεξιού μέλους της ισότητας (11):

$$\begin{aligned} a_1 = d_1 &\Rightarrow d_1 = a_1 & b_1 = d_1 e_1 &\Rightarrow e_1 = b_1/d_1 \\ a_2 = c_2 e_1 + d_2 &\Rightarrow d_2 = a_2 - c_2 e_1 & b_2 = d_2 e_2 &\Rightarrow e_2 = b_2/d_2 \\ & & \vdots & \end{aligned}$$

Αλγόριθμος της ανάλυσης του A σε LU

```
Είσοδος:  $a_i, i = 1, 2, \dots, n, \quad b_i, i = 1, 2, \dots, n-1, \quad c_i, i = 2, 3, \dots, n$   
 $d_1 \leftarrow a_1$   
 $e_1 \leftarrow b_1/d_1$   
for  $k = 2, \dots, n-1$  do  
     $d_k \leftarrow a_k - c_k \cdot e_{k-1}$   
     $e_k \leftarrow b_k/d_k$   
end for  
 $d_n \leftarrow a_n - c_n \cdot e_{n-1}$ 
```

Πλήθος πράξεων Αλγ. = $(2n-2)$ πολ/σμοί-διαιρέσεις και $(n-1)$ προσθέσεις.

Ο αλγόριθμος θα είναι καλά ορισμένος $\Leftrightarrow d_i \neq 0, \forall i = 1, 2, \dots, n$.

Αν $d_i \neq 0 \Rightarrow L$ θα είναι αντιστρέψιμος $\Rightarrow A$ θα είναι και αυτός αντιστρέψιμος ως γινόμενο των αντιστρέψιμων πινάκων L και U .

Πρόταση

Έστω A, L, U οι προηγούμενοι πίνακες, και τα στοιχεία του τριδιαγώνιου πίνακα A ικανοποιούν τις υποθέσεις (10). Τότε τα $d_i, i = 1, 2, \dots, n$, είναι μη μηδενικά και, επομένως ο προηγούμενος αλγόριθμος είναι καλά ορισμένος. Επιπλέον, $|e_i| < 1, \forall i = 1, \dots, n-1$.

Απόδειξη. Η απόδειξη θα γίνει **επαγωγικά**.

Αρχικά παρατηρούμε ότι: $d_1 = a_1$ και $a_1 \neq 0 \Rightarrow d_1 \neq 0$.

Λόγω της υπόθεσης $|a_1| > |b_1|$ έχουμε ότι $|e_1| = \left| \frac{b_1}{d_1} \right| = \left| \frac{b_1}{a_1} \right| < 1$.

Έστω τώρα ότι $d_{k-1} \neq 0$ και $|e_{k-1}| < 1$ για κάποιο $k, k = 2, \dots, n$.

Θ.δ.ο. $d_k \neq 0$ και $|e_k| < 1$. Όμως

$$|d_k| = |a_k - c_k e_{k-1}| \geq |a_k| - |c_k| |e_{k-1}| > |a_k| - |c_k| \stackrel{(10)}{\geq} |b_k| \stackrel{(10)}{>} 0,$$

Άρα

$$|d_k| > |b_k| \Rightarrow \frac{|b_k|}{|d_k|} < 1 \Rightarrow |e_k| < 1.$$

Τώρα, για να λύσουμε το σύστημα (8) εργαζόμαστε ως εξής:

$$Ax = f \Leftrightarrow LUx = f \Leftrightarrow L(\underbrace{Ux}_{=y}) = f.$$

Λύνουμε αρχικά το σύστημα

$$Ly = f,$$

Είσοδος: $c_i, i = 2, 3, \dots, n, d_i, i = 1, 2, \dots, n, f_i, i = 1, 2, \dots, n$
 $y_1 \leftarrow f_1/d_1$
for $k = 2, \dots, n$ **do**
 $y_k \leftarrow (f_k - c_k \cdot y_{k-1})/d_k$
end for

Πλήθος πράξεων Αλγ. = $(2n - 1)$ πολ/σμοί-διαιρέσεις και $(n - 1)$ προσθέσεις.

Στη συνέχεια, λύνουμε το σύστημα

$$Ux = y$$

Είσοδος: $y_i, i = 1, 2, \dots, n, e_i, i = 1, 2, \dots, n - 1$
 $x_n \leftarrow y_n$
for $k = n - 1, n - 2, \dots, 1$ **do**
 $x_k \leftarrow y_k - e_k \cdot x_{k+1}$
end for

Πλήθος πράξεων Αλγ. = $(n - 1)$ πολ/σμοι και $(n - 1)$ προσθέσεις.

Κόστος της μεθόδου σε πράξεις και θέσεις μνήμης

Συνολικά η μέθοδος απαιτεί $5n - 4$ πολλαπλασιασμούς-διαιρέσεις και $3n - 3$ προσθέσεις, δηλαδή πράξεις $O(n)$.

Θυμηθείτε ότι η απαλοιφή Gauss για έναν πλήρη $n \times n$ πίνακα απαιτεί $O(n^3)$ πράξεις.

Μνήμη: Χρειαζόμαστε περίπου $4n$ θέσεις μνήμης για την αποθήκευση των διανυσμάτων a, b, c, f . Μετά

$$a_i \leftarrow d_i, \quad b_i \leftarrow e_i, \quad f \leftarrow y \leftarrow x$$

Άρα συνολικά οι απαιτούμενες θέσεις μνήμης είναι μόνο $4n$.

Επαναληπτικές μέθοδοι

Θέλουμε και πάλι να λύσουμε το σύστημα:

$$Ax = b,$$

όπου

$$A = \begin{pmatrix} a_{11} & a_{12} & a_{13} & \dots & a_{1n} \\ a_{21} & a_{22} & a_{23} & \dots & a_{2n} \\ \vdots & \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & a_{n3} & \dots & a_{nn} \end{pmatrix},$$

και

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}, \quad b = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{pmatrix}.$$

Στις επαναληπτικές μεθόδους θέλουμε να κατασκευάσουμε μια ακολουθία διανυσμάτων $x^{(m)}$, η οποία, υπό ορισμένες προϋποθέσεις, θα συγκλίνει στη λύση x , καθώς $m \rightarrow \infty$.

«Σπάμε» τον πίνακα A ως εξής:

$$A = L + D + U,$$

όπου

$$L = \begin{pmatrix} 0 & & & & \\ a_{21} & 0 & & & \\ a_{31} & a_{32} & 0 & & \\ \vdots & \vdots & & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{n,n-1} & 0 \end{pmatrix}, \quad D = \begin{pmatrix} a_{11} & & & & \\ & a_{22} & & & \\ & & a_{33} & & \\ & & & \ddots & \\ & 0 & & & a_{nn} \end{pmatrix}$$

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

Μέθοδος Jacobi

$$\begin{aligned} Ax = b &\Rightarrow (L + D + U)x = b \Rightarrow Dx = -(L + U)x + b \\ &\Rightarrow x = -D^{-1}(L + U)x + D^{-1}b. \end{aligned}$$

Τότε για κάθε $m \in \mathbb{N}$ ορίζουμε

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

όπου $x^{(0)}$ είναι ένα δεδομένο αρχικό διάνυσμα.

Αν γράψουμε την ίδια σχέση κατά συνιστώσα θα έχουμε:

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

Θα λέμε ότι ο πίνακας A έχει **αυστηρά κυριαρχική διαγώνιο**, αν

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

Αποδεικνύεται ότι αν ένας πίνακας A έχει αυστηρά κυριαρχική διαγώνιο, τότε η μέθοδος Jacobi συγκλίνει.

Παράδειγμα

Να λυθεί με τη μέθοδο Jacobi το σύστημα

$$\begin{cases} 3x_1 + 6x_2 + 2x_3 = 0 \\ 3x_1 - x_2 + x_3 = 1 \\ 3x_1 + 3x_2 + 7x_3 = 4 \end{cases}.$$

Παρατηρήστε ότι ο πίνακας του συστήματος είναι:

$$\begin{pmatrix} 3 & 6 & 2 \\ 3 & -1 & 1 \\ 3 & 3 & 7 \end{pmatrix}$$

και δεν έχει αυστηρά κυριαρχική διαγώνιο.

Αν όμως εναλλάξουμε τις δύο πρώτες εξισώσεις έχουμε:

$$\begin{cases} 3x_1 - x_2 + x_3 = 1 \\ 3x_1 + 6x_2 + 2x_3 = 0 \\ 3x_1 + 3x_2 + 7x_3 = 4 \end{cases},$$

αλλά τώρα ο πίνακας του συστήματος είναι:

$$\begin{pmatrix} 3 & -1 & 1 \\ 3 & 6 & 2 \\ 3 & 3 & 7 \end{pmatrix}$$

και έχει αυστηρά κυριαρχική διαγώνιο, αφού

$$|3| > |-1| + |1|$$

$$|6| > |3| + |2|$$

$$|7| > |3| + |3|$$

Θα εφαρμόσουμε τη μέθοδο Jacobi στο σύστημα

$$\begin{cases} 3x_1 - x_2 + x_3 = 1 \\ 3x_1 + 6x_2 + 2x_3 = 0 \\ 3x_1 + 3x_2 + 7x_3 = 4 \end{cases} \Leftrightarrow \begin{cases} x_1 = \frac{1}{3}(1 + x_2 - x_3) \\ x_2 = \frac{1}{6}(-3x_1 - 2x_3) \\ x_3 = \frac{1}{7}(4 - 3x_1 - 3x_2) \end{cases}$$

Επομένως

$$\begin{aligned} x_1^{(m+1)} &= \frac{1}{3} \left(1 + x_2^{(m)} - x_3^{(m)} \right) \\ x_2^{(m+1)} &= \frac{1}{6} \left(-3x_1^{(m)} - 2x_3^{(m)} \right) \\ x_3^{(m+1)} &= \frac{1}{7} \left(4 - 3x_1^{(m)} - 3x_2^{(m)} \right) \end{aligned}$$

Θεωρούμε ως αρχική προσέγγιση το διάνυσμα $x^{(0)} = (0, 0, 0)^T$, και θα υπολογίσουμε τις επόμενες δύο προσεγγίσεις.

- 1η προσέγγιση:

$$\begin{aligned}x_1^{(1)} &= \frac{1}{3} (1 + 0 - 0) = \frac{1}{3} \\x_2^{(1)} &= \frac{1}{6} (-3 \cdot 0 - 2 \cdot 0) = 0 \\x_3^{(1)} &= \frac{1}{7} (4 - 3 \cdot 0 - 3 \cdot 0) = \frac{4}{7}\end{aligned}$$

- 2η προσέγγιση:

$$\begin{aligned}x_1^{(2)} &= \frac{1}{3} \left(1 + 0 - \frac{4}{7} \right) = \frac{1}{7} \approx 0.1429 \\x_2^{(2)} &= \frac{1}{6} \left(-3 \cdot \frac{1}{3} - 2 \cdot \frac{4}{7} \right) = -\frac{15}{42} \approx -0.3571 \\x_3^{(2)} &= \frac{1}{7} \left(4 - 3 \cdot \frac{1}{3} - 3 \cdot 0 \right) = \frac{3}{7} \approx 0.4286\end{aligned}$$

Άρα η δεύτερη προσέγγιση της λύσης με τη μέθοδο Jacobi είναι το διάνυσμα:
 $\left(\frac{1}{7}, -\frac{15}{42}, \frac{3}{7} \right)^T \approx (0.1429, -0.3571, 0.4286)^T$.

Μέθοδος Gauss–Seidel

Ιδέα: Για τον υπολογισμό του $x_i^{(m+1)}$ ($i > 1$) αντί για τα $x_1^{(m)}, \dots, x_{i-1}^{(m)}$ θα χρησιμοποιήσουμε τα $x_1^{(m+1)}, \dots, x_{i-1}^{(m+1)}$, που έχουν ήδη υπολογισθεί.

Για κάθε $m \in \mathbb{N}$ θα υπολογίζουμε προσεγγίσεις με τη μέθοδο **Gauss–Seidel** μέσω του τύπου:

$$x_i^{(m+1)} = \frac{1}{a_{ii}} \left(b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(m+1)} - \sum_{j=i+1}^n a_{ij} x_j^{(m)} \right), \quad i = 1, \dots, n. \quad (12)$$

όπου $x^{(0)}$ είναι ένα δεδομένο αρχικό διάνυσμα.

Παράδειγμα επίλυσης με Gauss–Seidel

$$\begin{cases} 3x_1 - x_2 + x_3 = 1 \\ 3x_1 + 6x_2 + 2x_3 = 0 \\ 3x_1 + 3x_2 + 7x_3 = 4 \end{cases} \Leftrightarrow \begin{cases} x_1 = \frac{1}{3}(1 + x_2 - x_3) \\ x_2 = \frac{1}{6}(-3x_1 - 2x_3) \\ x_3 = \frac{1}{7}(4 - 3x_1 - 3x_2) \end{cases}$$

Επομένως

$$\begin{aligned} x_1^{(m+1)} &= \frac{1}{3} \left(1 + x_2^{(m)} - x_3^{(m)} \right) \\ x_2^{(m+1)} &= \frac{1}{6} \left(-3x_1^{(m+1)} - 2x_3^{(m)} \right) \\ x_3^{(m+1)} &= \frac{1}{7} \left(4 - 3x_1^{(m+1)} - 3x_2^{(m+1)} \right) \end{aligned}$$

Αρχική προσέγγιση $x^{(0)} = (0, 0, 0)^T$.

• 1η προσέγγιση:

$$\begin{aligned} x_1^{(1)} &= \frac{1}{3} \left(1 + x_2^{(0)} - x_3^{(0)} \right) = \frac{1}{3} (1 + 0 - 0) = \frac{1}{3} \\ x_2^{(1)} &= \frac{1}{6} \left(-3x_1^{(1)} - 2x_3^{(0)} \right) = \frac{1}{6} \left(-3 \cdot \frac{1}{3} - 2 \cdot 0 \right) = -\frac{1}{6} \\ x_3^{(1)} &= \frac{1}{7} \left(4 - 3x_1^{(1)} - 3x_2^{(1)} \right) = \frac{1}{7} \left(4 - 3 \cdot \frac{1}{3} - 3 \cdot \left(-\frac{1}{6}\right) \right) = \frac{1}{2} \end{aligned}$$

• 2η προσέγγιση:

$$\begin{aligned} x_1^{(2)} &= \frac{1}{3} \left(1 + x_2^{(1)} - x_3^{(1)} \right) = \frac{1}{3} \left(1 + \left(-\frac{1}{6}\right) - \frac{1}{2} \right) = \frac{1}{9} \approx 0.1111 \\ x_2^{(2)} &= \frac{1}{6} \left(-3x_1^{(2)} - 2x_3^{(1)} \right) = \frac{1}{6} \left(-3 \cdot \frac{1}{9} - 2 \cdot \frac{1}{2} \right) = -\frac{2}{9} \approx -0.2222 \\ x_3^{(2)} &= \frac{1}{7} \left(4 - 3x_1^{(2)} - 3x_2^{(2)} \right) = \frac{1}{7} \left(4 - 3 \cdot \frac{1}{9} - 3 \cdot \left(-\frac{2}{9}\right) \right) = \frac{13}{21} \approx 0.6190 \end{aligned}$$

Άρα η δεύτερη προσέγγιση της λύσης με τη μέθοδο Gauss–Seidel είναι το διάνυσμα:

$$\left(\frac{1}{9}, -\frac{2}{9}, \frac{13}{21} \right)^T \approx (0.1111, -0.2222, 0.6190)^T.$$

Σημειώστε ότι η ακριβής λύση αυτού του συστήματος είναι: $(0.0351, -0.2368, 0.6579)^T$.