Κεφάλαιο 5

Διανυσματικοί χώροι με εσωτερικό γινόμενο

5.1 Διανυσματικοί χώροι με νόρμα

Για το $\mathbb{R}^n$, έχουμε δει το εσωτερικό γινόμενο δύο διανυσμάτων x και y,

$$x \cdot y = x^T y = x_1 y_1 + \cdots + x_n y_n$$

και τη νόρμα

$$||x|| = \sqrt{\langle x,\, y\rangle}\,.$$

Στο $\mathbb{C}$, θα θέλαμε η νόρμα $||z||$ να συμπίπτει με το μέτρο $|z|$ και, εάν $z = x + \mathrm{i}y$ με τη νόρμα του $(x,\, y)$:

$$||z|| = |z| = \sqrt{z\bar{z}} = \sqrt{x^2 + y^2} = ||(x,\, y)||\,.$$

Ανάλογα, στο $\mathbb{C}^n$, εάν θέλουμε $||(z_1, \ldots, z_n)||$ να συμπίπτει με τη νόρμα του διανύσματος $(x_1,\, y_1,\, x_2,\, y_2, \ldots, x_n,\, y_n)$ στο $\mathbb{R}^{2n}$, όπου $z_j = x_j + \mathrm{i}y_j$, πρέπει να ορίσουμε τη νόρμα

$$||z|| = \sqrt{z_1\bar{z}_1 + \cdots + z_n\bar{z}_n}$$

και το εσωτερικό γινόμενο

$$\langle z,\, w\rangle = z_1\bar{w}_1 + \cdots + z_n\bar{w}_n\,.$$

Μετά από αυτές τις παρατηρήσεις, δίδουμε τον ακόλουθο ορισμό.

Ορισμός 5.1. V διανυσματικός χώρος πάνω από το σώμα $\mathbb{K}$ $(= \mathbb{R}$ ή $\mathbb{C})$. Μια απεικόνιση $V \longrightarrow \mathbb{R} : v \mapsto ||v||$ ονομάζεται **νόρμα** (ή **στάθμη**) εάν

N 1. $||v|| = 0$ εάν και μόνον εάν $v = 0$

N 2. Για κάθε $v \in V$ και $a \in \mathbb{K}$, $||av|| = |a|\, ||v||$

N 3. Για κάθε $v,\, w \in V$, $||v + w|| \leq ||v|| + ||w||$ (τριγωνική ανισότητα)

Λήμμα 5.1 *Σε ένα διανυσματικό χώρο V με νόρμα,*

1. *Για κάθε $v \in V$, $||v|| \geq 0$.*
2. *Για κάθε $v, w \in V$, $||v-w|| \geq |\ ||v|| - ||w||\ |$.*

Απόδειξη.

1. Για κάθε $v \in V$,

$$\begin{aligned} ||v|| &= \frac{1}{2}(||v|| + ||v||) = \frac{1}{2}(||v|| + ||-v||) \\ &\geq \frac{1}{2}||v + (-v)|| = \frac{1}{2}||0|| = 0\,. \end{aligned}$$

2. Για κάθε $v, w \in V$,

$$||v|| = ||(v-w)+w|| \leq ||v-w|| + ||w||\,,$$

και συνεπώς

$$||v-w|| \geq ||v|| - ||w||\,.$$

Ανάλογα

$$||v-w|| = ||w-v|| \geq ||w|| - ||v||\,.$$

□

Τα αξιώματα Ν 1 και Ν3 μας επιτρέπουν να θεωρούμε τη νόρμα ως γενίκευση της έννοιας του μήκους ενός γεωμετρικού διανύσματος. Κάθε διάνυσμα έχει μήκος μεγαλύτερο ή ίσο με το μηδέν, και μόνο το μηδενικό διάνυσμα έχει μήκος μηδέν. Μπορούμε να θεωρήσουμε το $||v-w||$ ως στην απόσταση από το διάνυσμα v στο διάνυσμα w, που ικανοποιεί τη βασική ιδιότητα της απόστασης, την τριγωνική ανισότητα $||v-w|| \leq ||v-u|| + ||u-w||$.

Παράδειγμα 5.1 Στο $\mathbb{R}^n$ και στο $\mathbb{C}^n$ η ευκλείδεια νόρμα (ή ℓ^2-νόρμα) είναι η συνήθης νόρμα

$$||x|| = \sqrt{x_1^2 + \cdots + x_n^2}$$

και

$$||z|| = \sqrt{z_1\bar{z}_1 + \cdots + z_n\bar{z}_n}\,.$$

Παράδειγμα 5.2 Η ℓ^1-νόρμα στο $\mathbb{R}^n$ ορίζεται ως

$$||x||_1 = |x_1| + \cdots + |x_n|\,.$$

Ελέγχουμε τα αξιώματα:

N 1.

$$\begin{aligned} ||x||_1 = 0 &\Leftrightarrow |x_1| = \cdots = |x_n| = 0 \\ &\Leftrightarrow x = 0\,. \end{aligned}$$

N 2.

$$||ax||_1 = \sum_{i=1}^{n} |ax_i| = |a| \sum_{i=1}^{n} |x_i| = |a|\,||x||_1\,.$$

N 3.

$$||x+y||_1 = \sum_{i=1}^{n} |x_i + y_i| \le \sum_{i=1}^{n} |x_i| + |y_i| = ||x||_1 + ||y||_1\,.$$

Παράδειγμα 5.3 Η ℓ^∞-νόρμα στο $\mathbb{R}^n$ ορίζεται ως

$$||x||_\infty = \max\{|x_i| \ : \ i = 1, \ldots, n\}\,.$$

Άσκηση 5.1 Δείξτε οτι $||x||_\infty$ ικανοποιεί τα αξιώματα της νόρμας.

Παράδειγμα 5.4 Στο χώρο των πολυωνύμων $\mathbb{K}[x]$, με $\mathbb{K} = \mathbb{R}$ ή $\mathbb{C}$, ορίζουμε τη νόρμα

$$||p(x)|| = \left(\int_0^1 |p(t)|^2 dt\right)^{1/2}\,.$$

Ο έλεγχος των αξιωμάτων N 1 και N 2 είναι εύκολος. Για το N 3 παρατηρούμε οτι

$$\begin{aligned} ||p(x) + q(x)||^2 &= \int_0^1 |p(t) + q(t)|^2 dt \\ &= \int_0^1 |p(t)|^2 dt + \int_0^1 |q(t)|^2 dt + 2\mathrm{Re} \int_0^1 p(t)\overline{q(t)} dt \end{aligned}$$

ενώ

$$(||p(x)|| + ||q(x)||)^2 = ||p(x)||^2 + ||q(x)||^2 + 2||p(x)||\,||q(x)||\,.$$

Άρα για να ισχύει η τριγωνική ανισότητα, αρκεί να ισχύει η ανισότητα

$$\mathrm{Re} \int_0^1 p(t)\overline{q(t)} dt \le ||p(x)||\,||q(x)||\,,$$

την οποία θα αποδείξουμε στην Πρόταση 5.5.

Παράδειγμα 5.5 Στο χώρο $C[a, b]$ των συνεχών συναρτήσεων στο διάστημα $[a, b]$ (με πραγματικές ή μιγαδικές τιμές) ορίζουμε την L^2-νόρμα

$$||f|| = \left(\int_a^b |f(t)|^2 dt\right)^{1/2}$$

και την L^∞-νόρμα

$$||f||_\infty = \max\{|f(x)| \ : \ x \in [a,\, b]\}\,.$$

Άσκηση 5.2 Δείξτε οτι $||f||_\infty$ ικανοποιεί τα αξιώματα της νόρμας.

Παράδειγμα 5.6 Ο περιορισμός στο προηγούμενο παράδειγμα σε συνεχείς συναρτήσεις δεν είναι φυσιολογικός, αφού για τον ορισμό της νόρμας χρησιμοποιούμε μόνο το ολοκλήρωμα. Στην Ανάλυση συχνά εξετάζουμε όρια συναρτήσεων, και το όριο μία ακολουθίας συνεχών συναρτήσεων δεν είναι απαραίτητα συνεχής συνάρτηση. Ας προσπαθήσουμε να ορίσουμε μία νόρμα σε ένα μεγαλύτερο σύνολο συναρτήσεων. Ορίζουμε $\mathcal{L}^2(I)$ ως το σύνολο των συναρτήσεων στο διάστημα $I = [0,\, 1]$ (με πραγματικές ή μιγαδικές τιμές), για τις οποίες το ολοκλήρωμα $\int_0^1 |f(t)|^2 dt < \infty$. Εύκολα δείχνουμε οτι αυτό το σύνολο είναι κλειστό ως προς πολλαπλασιασμό με αριθμό. Μπορούμε επίσης να δείξουμε οτι

$$\int_0^1 |f(t)+g(t)|^2 dt \leq \left(\left(\int_0^1 |f(t)|^2 dt \right)^{1/2} + \left(\int_0^1 |g(t)|^2 dt \right)^{1/2} \right)^2 ,$$

και συνεπώς το σύνολο είναι κλειστό ως προς την πρόσθεση συναρτήσεων. Συμπεραίνουμε οτι $\mathcal{L}^2(I)$ είναι διανυσματικός χώρος[1]. Στο $\mathcal{L}^2(I)$ ορίζουμε τη συνάρτηση

$$||f|| = \left(\int_0^1 |f(t)|^2 dt \right)^{1/2} ,$$

η οποία ικανοποιεί τα αξιώματα N 2 και N 3. Όμως μία συνάρτηση με μηδενικό ολοκλήρωμα δεν είναι υποχρεωτικά ίση με 0. Για παράδειγμα, η συνάρτηση $h : [0,\, 1] \longrightarrow \mathbb{R}$ με

$$h(x) = \begin{cases} 1 & x = \frac{1}{2} \\ 0 & x \neq \frac{1}{2} \end{cases} ,$$

είναι ολοκληρώσιμη, με $\int_0^1 |h(t)|^2 dt = 0$ αλλά δεν είναι η μηδενική συνάρτηση.

Για να μπορέσουμε να ορίσουμε μία νόρμα σε αυτή την περίπτωση, χρειάζεται να εξετάσουμε έναν διαφορετικό διανυσματικό χώρο. Το σύνολο όλων των συναρτήσεων $f \in \mathcal{L}^2(I)$ με $||f|| = 0$ αποτελεί διανυσματικό υπόχωρο L_0 του $\mathcal{L}^2(I)$. Θεωρούμε το χώρο πηλίκο του $\mathcal{L}^2(I)$ με το διανυσματικό υπόχωρο L_0. Τότε το σύνολο L_0 αντιστοιχεί στο μηδενικό διάνυσμα του $\mathcal{L}^2(I)/\,L_0$. Εάν $h \in L_0$, τότε για κάθε συνάρτηση $f \in \mathcal{L}^2(I)$, $||f+h|| = ||f||$. Συμπεραίνουμε οτι η συνάρτηση $(f + L_0) \mapsto ||f||$ είναι καλά ορισμένη στο χώρο πηλίκο $\mathcal{L}^2(I)/\,L_0$, και οτι ικανοποιεί τα αξιώματα μίας νόρμας. Ο διανυσματικός χώρος $\mathcal{L}^2(I)/\,L_0$ με αυτή τη νόρμα συμβολίζεται $L^2(I)$, και η νόρμα ονομάζεται L^2-νόρμα.

[1] Στον ορισμό του χώρου $\mathcal{L}^2(I)$ χρησιμοποιούμε έναν άλλο ορισμό του ολοκληρώματος, το ολοκλήρωμα Lebesgue, το οποίο όμως έχει ιδιότητες παρόμοιες με αυτές του ολοκληρώματος Riemann που έχετε γνωρίσει στον Απειροστικό Λογισμό Ι

5.2 Ισοδύναμες Νόρμες

Ορισμός 5.2. Έστω V ένας διανυσματικός χώρος. Δύο νόρμες $\|\cdot\|, \|\cdot\|'$ στον V λέγονται **ισοδύναμες** αν υπάρχουν θετικές σταθερές m και M τέτοιες ώστε, για κάθε $x \in V$,

$$m\|x\| \le \|x\|' \le M\|x\|\,.$$

Από τον παραπάνω ορισμό προκύπτει άμεσα οτι για κάθε $x \in V$,

$$\frac{1}{M}\|x\|' \le \|x\| \le \frac{1}{m}\|x\|'\,.$$

Παράδειγμα 5.7 Οι νόρμες $\|\cdot\|_1$ και $\|\cdot\|_\infty$ του $\mathbb{R}^n$ είναι ισοδύναμες μεταξύ τους.

Πράγματι για $x \in \mathbb{R}^n$ έχουμε

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

Έστω k ο δείκτης για τον οποίο $|x_k| = \|x\|_\infty$, τότε έχουμε

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

Συνδυάζοντας τις παραπάνω σχέσεις, έχουμε σύμφωνα με τον ορισμό για $m = 1,\ M = n$,

$$\|x\|_\infty \le \|x\|_1 \le n\|x\|_\infty$$

Λήμμα 5.2 *Κάθε νόρμα $\|\cdot\|$ στον $\mathbb{R}^n$ είναι ισοδύναμη με τη νόρμα μεγίστου $\|\cdot\|_\infty$ του $\mathbb{R}^n$.*

□

Χρησιμοποιώντας το παραπάνω Λήμμα μπορούμε να δείξουμε το ακόλουθο πιο γενικό αποτέλεσμα.

Λήμμα 5.3 *Όλες οι νόρμες στον $\mathbb{R}^n$ είναι ισοδύναμες μεταξύ τους.*

Γενικότερα μπορεί να δείξει κανείς οτι σε χώρους πεπερασμένης διάστασης όλες οι νόρμες πάνω στο χώρο είναι ισοδύναμες μεταξύ τους. Σε αυτή την περίπτωση οι αντίστοιχες σταθερές m, M του ορισμού, μπορεί να εξαρτώνται από τη διάσταση του χώρου.

5.3 Νόρμες Πινάκων

Εισάγουμε τώρα τις νόρμες πινάκων. Στον γραμμικό χώρο $\mathcal{M}(n, \mathbb{R})$ των πινάκων μπορούμε να ορίσουμε διάφορες νόρμες. Όμως, θεωρούμε τους πίνακες ως *απεικονίσεις* από τον $\mathbb{R}^n$ στον $\mathbb{R}^n$ και ορίζουμε τις νόρμες πινάκων με συγκεκριμένο τρόπο.

Ορισμός 5.3. Μια απεικόνιση $\mathcal{M}(n, \mathbb{R}) \longrightarrow \mathbb{R} : A \mapsto \|A\|$ λέγεται **νόρμα πινάκων**(ή **στάθμη πινάκων**) εαν

ΝΠ 1. $\|A\| = 0$ εάν και μόνον εάν $A = 0$

ΝΠ 2. Για κάθε $A \in \mathcal{M}(n, \mathbb{R})$ και $a \in \mathbb{R}$, $\|av\| = |a| \, \|v\|$

ΝΠ 3. Για κάθε $A, B \in \mathcal{M}(n, \mathbb{R})$, $\|A + B\| \leq \|A\| + \|B\|$ (τριγωνική ανισότητα)

ΝΠ 4. Για κάθε $A, B \in \mathcal{M}(n, \mathbb{R})$, $\|A\,B\| \leq \|A\| \, \|B\|$

Παρατηρούμε οτι οι τρεις πρώτες ιδιότητες της νόρμας πίνακα ταυτίζονται με τις αντίστοιχες της νόρμας διανύσματος.

Παράδειγμα 5.8 Ένα παράδειγμα νόρμας πίνακα είναι η νόρμα Frobenius :

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

Στη συνέχεια θα επικεντρωθούμε σε μια συγκεκριμμένη κατηγορία νορμών πινάκων, τις λεγόμενες **φυσικές νόρμες** ή **νόρμες τελεστών**.

Ορισμός 5.4. Έστω $\mathbb{R}^n \longrightarrow \mathbb{R} : u \mapsto \|u\|_v$ μια νόρμα διανυσμάτων. Η απεικόνιση $\mathcal{M}(n, \mathbb{R}) \longrightarrow \mathbb{R} : A \mapsto \|A\|_m$

$$\|A\|_m = \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \frac{\|Ax\|_v}{\|x\|_v} \tag{5.1}$$

λέγεται **φυσική νόρμα πίνακα** παραγόμενη από την νόρμα $\| \cdot \|_v$ του $\mathbb{R}^n$.

Δεν είναι δύσκολο να δει κανείς οτι μια *φυσική νόρμα πίνακα* είναι μια νόρμα πινάκων. Πράγματι μπορούμε να επαληθεύσουμε ότι η 5.1 ικανοποιεί όλες τις ιδιότητες του ορισμού 5.3. Για κάθε $A, B \in \mathcal{M}(n, \mathbb{R})$ και $\lambda \in \mathbb{R}$,

$$\|A\| = 0 \iff \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \tfrac{\|Ax\|}{\|x\|} \iff \forall x \in \mathbb{R}^n \; Ax = 0 \iff A = 0$$

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

$$\|A + B\| = \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \tfrac{\|(A+B)x\|}{\|x\|} \leq \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \tfrac{\|Ax\|}{\|x\|} + \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \tfrac{\|Bx\|}{\|x\|} = \|A\| + \|B\|$$

$$\|AB\| = \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \tfrac{\|ABx\|}{\|x\|} \leq \sup_{\substack{x \in \mathbb{R}^n \\ x \neq 0}} \tfrac{\|A\| \|Bx\|}{\|x\|} = \|A\| \; \|B\|$$

Η τελευταία σχέση προκύπτει από την

$$\|Ax\| \leq \|A\| \; \|x\|$$

για κάθε $x \in \mathbb{R}^n$, η οποία είναι άμεση συνέπεια του ορισμού 5.1.

5.3.1 Παραδείγματα φυσικών νορμών πινάκων

1. Στον $\mathbb{R}^n$ θεωρούμε την νόρμα μεγίστου $\|\cdot\|_\infty$. Η αντίστοιχη φυσική νόρμα πίνακα που παράγεται ονομάζεται *νόρμα αθροίσματος γραμμών* και δίνεται από

$$A \in \mathcal{M}(n, \mathbb{R}), \quad \|A\|_\infty = \max_{1\le i\le n} \sum_{j=1}^{n} |a_{ij}| \tag{5.2}$$

2. Στον $\mathbb{R}^n$ θεωρούμε την νόρμα $\|\cdot\|_1$. Η αντίστοιχη φυσική νόρμα πίνακα που παράγεται ονομάζεται *νόρμα αθροίσματος στηλών* και δίνεται από

$$A \in \mathcal{M}(n, \mathbb{R}), \quad \|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{n} |a_{ij}| \tag{5.3}$$

3. Στον $\mathbb{R}^n$ θεωρούμε την Ευκλείδια νόρμα $\|\cdot\|_2$. Η αντίστοιχη φυσική νόρμα πίνακα που παράγεται δίνεται από

$$A \in \mathcal{M}(n, \mathbb{R}), \quad \|A\|_2 = \big(\rho(A^T\, A)\big)^{1/2} \tag{5.4}$$

όπου $\rho(A)$ είναι η *φασματική ακτίνα* του πίνακα A και ορίζεται ως το μέγιστο τών απολύτων τιμών των ιδιοτιμών του πίνακα A.

Παράδειγμα 5.9 Για τον πίνακα

$$A = \begin{bmatrix} 1 & 2 & -2 \\ 2 & 3 & -4 \\ -1 & -2 & 0 \end{bmatrix} \tag{5.5}$$

έχουμε ότι $\|A\|_\infty = 9,\ \|A\|_1 = 7,\ \|A\|_2 = 6,3543,\ \|A\|_F = 6,5574$

Τέλος να αναφέρουμε οτι η νόρμα Frobenius που ορίσαμε παραπάνω δεν είναι φυσική νόρμα.

5.4 Δείκτης κατάστασης πίνακα

Έστω $A \in \mathcal{M}(n, \mathbb{R}),\ b \in \mathbb{R}^n$. Υποθέτουμε ότι ο πίνακας A είναι αντιστρέψιμος και θεωρούμε τη λύση του γραμμικού συστήματος

$$Ax = b. \tag{5.6}$$

Έστω οτι στα στοιχεία του b εισάγεται κάποια διαταραχή Δb. Αυτό έχει ως συνέπεια να διαταράξει την λύση του 5.6 σε $x + \Delta x$,

$$A(x + \Delta x) = b + \Delta b. \tag{5.7}$$

Ερώτημα : Μπορούμε να εκτιμήσουμε το μέγεθος της σχετικής μεταβολής $\|\Delta x\|/\|x\|$ της λύσης του συστήματος λόγω της μεταβολής του δεξιού μέλους;

Με την χρήση των νορμών μπορούμε να απαντήσουμε θετικά στο παραπάνω ερώτημα. Ας δούμε πως ! Αφαιρούμε κατά μέλη τις σχέσεις 5.6 και 5.7 παίρνουμε

$$A\ \Delta x = \Delta b \Rightarrow \Delta x = A^{-1}\ \Delta b \Rightarrow \|\Delta x\| \le \|A^{-1}\|\|\Delta b\| \tag{5.8}$$

για κάποια νόρμα $\|\cdot\|$ στον $\mathbb{R}^n$ και την παραγόμενη φυσική νόρμα πίνακα. Υποθέτοντας οτι $b \neq 0$, άρα και $x \neq 0$, από την 5.6 έχουμε ότι

$$\|b\| \le \|A\|\|x\| \tag{5.9}$$

Συνδυάζοντας τις σχέσεις 5.8, 5.9 παίρνουμε την εκτίμηση

$$\frac{\|\Delta x\|}{\|x\|} \le \|A\|\,\|A^{-1}\|\frac{\|\Delta b\|}{\|b\|} \tag{5.10}$$

Δοσμένου οτι για κάθε πίνακα A και νόρμα $\|\cdot\|$ μπορούμε να βρούμε b και Δb έτσι ώστε η 5.10 να ισχύει ως ισότητα καταλήγουμε οτι ο αριθμός

$$\kappa(A) = \|A\|\,\|A^{-1}\| \tag{5.11}$$

είναι ένας συντελεστής ¨ευαισθησίας-μεγένθυσησ' που εκφράζει τη μέγιστη δυνατή μεταβολή της λύσης $\|\Delta x\|/\|x\|$ ως προς την μεταβολή του δεξιού μέλους $\|\Delta b\|/\|b\|$

Ορισμός 5.5. Ο αριθμός $\kappa(A) = \|A\|\ \|A^{-1}\|$ ονομάζεται **δείκτης κατάστασης** του πίνακα A.

Παρατηρούμε ότι ο $\kappa(A)$ είναι θετικός αριθμός και μάλιστα είναι μεγαλύτερος ή ίσος με την μονάδα,

$$1 = \|I\| = \|A\ A^{-1}\| \le \|A\|\ \|A^{-1}\| = \kappa(A).$$

Η τιμή του $\kappa(A)$ εξαρτάται από την νόρμα πίνακα που χρησιμοποιούμε. Για το πίνακα στο παράδειγμα 5.5 έχουμε ότι $\kappa_\infty(A) = 63,\ \kappa_1(A) = 45,5,\ \kappa_2(A) = 32,45,\ \kappa_F(A) = 33,76.$

Ας δούμε τώρα με ένα πάραδειγμα πώς επηρεάζει ο δείκτης κατάστασης ενός πίνακα τη λύση ενός γραμμικού συστήματος.

Παράδειγμα 5.10 Θεωρούμε το παρακάτω γραμμικό σύστημα

$$\begin{bmatrix} 0,913 & 0,659 \\ 0,780 & 0,563 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0,254 \\ 0,217 \end{bmatrix} \tag{5.12}$$

το οποίο έχει μοναδική λύση $x = (1,-1)^T$, αν φυσικά κάνουμε τις πράξεις ακριβώς, δηλαδή χρησιμοποιώντας άπειρα δεκαδικά ψηφία. Όμως αν θελούμε να λύσουμε το συστήμα σε ένα υπολογιστή οι πράξεις θα γίνουνε με αριθμητική πεπερασμένης ακρίβειας. Ο λόγος είναι ότι ακόμα και οι πιο σύγχρονοι υπολογιστές χρησιμοποιούν πεπερασμένο αριθμό δεκαδικών ψηφίων για την αναπαράσταση των αριθμών.

Έστω λοιπόν ότι θέλουμε να λύσουμε το παραπάνω σύστημα σε ένα υπολογιστή που χρησιμοποιεί 3 δεκαδικά ψηφία για αναπαράσταση των αριθμών. Τότε η απαλοιφή Gauss σε αυτόν τον υπολογιστή θα δώσει

$$\begin{bmatrix} 0,913 & 0,659 \\ 0 & 0,001 \end{bmatrix} \begin{bmatrix} \tilde{x}_1 \\ \tilde{x}_2 \end{bmatrix} = \begin{bmatrix} 0,254 \\ 0,001 \end{bmatrix}$$

και η ανάδρομη αντικατάσταση θα δώσει την λύση $\tilde{x} = (1, -0,443)$!!! Ο λόγος αποτυχίας θα πεί κανείς είναι η αριθμητική πεπερασμένης ακρίβειας των τριών δεκαδικών ψηφίων του συγκεκριμένου υπολογιστή. Αυτό είναι εν μέρει σωστό και αν χρησιμοποιήσουμε έναν υπολογιστή με ακρίβεια 6 δεκαδικών ψηφίων θα πάρουμε την ακριβή λύση του συστήματος $x = (1, -1)^T$.

Ας εξετάσουμε όμως λίγο πιο προσεκτικά το σύστημα αυτό. Ας κρατήσουμε τον πίνακα του συστήματος σταθερό και ας μεταβάλλουμε το δεξί μέλος ελάχιστα, και ας θεωρήσουμε το εξής διαταραγμένο σύστημα

$$\begin{bmatrix} 0,913 & 0,659 \\ 0,780 & 0,563 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 0,253 \\ 0,218 \end{bmatrix}. \tag{5.13}$$

Παρατηρήστε οτι η διαταραχή του δεξιού μέλους του 5.13 σε σχέση με το αντίστοιχο του 5.12 είναι $\Delta b = (-0,001, 0,001)$. Η ακριβής λύση του 5.13 είναι $y = (1223, -1694)^T$!!! Μια μικρή μεταβολή του δεξιού μέλους προκάλεσε μια *μεγάλη* μεταβολή στη λύση x. Η εκτίμηση 5.10 και μερικοί απλοί υπολογισμοί θα μας βοηθήσουν να καταλάβουμε γιατί συνέβη αυτό.

Θεωρούμε την νόρμα $\|\cdot\|_1$ στον $\mathbb{R}^n$ και την παραγόμενη νόρμα πίνακα 5.3. Τότε έχουμε

$$\|\Delta x\|_1 = \|y - x\|_1 = |1223 - 1| + |-1694 + 1| = 2915, \ \|x\|_1 = 2 \Rightarrow \tfrac{\|\Delta x\|_1}{\|x\|_1} \sim 1,5 \times 10^3$$
$$\|\Delta b\|_1 = 2 \times 10^{-3}, \ \|b\|_1 = 0,471 \Rightarrow \tfrac{\|\Delta b\|_1}{\|b\|_1} \sim 4 \times 10^{-3}$$

Επίσης

$$\|A\|_1 = \max(1,693,\ 1,222) = 1,693, \qquad \|A^{-1}\|_1 = 10^6 \max(1,572,\ 1,343) = 1,572 \times 10^6$$

Άρα ο δείκτης κατάστασης του πίνακα είναι $\kappa_1(A) = 2,66 \times 10^6$!!! Από την εκτίμηση 5.10 βλέπουμε αμέσως πως μια μικρή μεταβολή της τάξεως 10^{-3} στο δεξί μέλος του συστήματος πολλαπλασιάστηκε από το δείκτη κατάστασης του πίνακα, τάξεως 10^6, για να δώσει μια μεγένθυνση τάξεως 10^3 στη λύση.

Από τον ορισμό του δείκτη κατάστασης $\kappa(A)$ βλέπουμε οτι αυτός ορίζεται για αντιστρέψιμους πίνακες. Μπορεί να δειχθεί οτι ο λόγος $1/\kappa(A)$ αποτελεί ένα μέτρο για την απόσταση του A από το σύνολο των μη-αντιστρέψιμων πινάκων. Ειδικότερα μπορούμε να δείξουμε την παρακάτω εκτίμηση

$$\frac{1}{\kappa(A)} \leq \inf\left\{\frac{\|A - B\|}{\|A\|}, \quad B \in \mathcal{M}(n, \mathbb{R}) \text{ μη-αντιστρέψιμος}\right\}$$

Άρα όταν $\kappa(A) \to \infty$ ο πίνακας A τείνει να γίνει μη-αντιστρέψιμος. Για συγκεκριμμένες νόρμες πινάκων η προηγούμενη ανισότητα ισχύει και ως ισότητα.

Παρατηρήσεις

1. Για τον υπολογισμό του $\kappa(A)$ απαιτείται η γνώση του αντίστροφου A^{-1} του οποίου το κόστος υπολογισμού είναι πολύ υψηλό για μεγαλές τιμές του n. Στη πράξη χρησιμοποιούνται ειδικοί αλγόριθμοι οι οποίοι προσεγγίζουν την $\|A^{-1}\|$. Συναρτήσεις υπολογισμού του $\kappa(A)$ υπάρχουν σε όλα τα σύγχρονα πακέτα-βιβλιοθήκες επίλυσης γραμμικών συστημάτων. Στη Python η συνάρτηση $numpy.linalg.cond(A, p)$ προσεγγίζει τον $\kappa(A)$ με μεγάλη ακρίβεια στη νόρμα πίνακα που προσδιορίζεται απο την παράμετρο p.

2. Όπως είδαμε, τόσο η ορίζουσα $det(A)$ όσο και ο λόγος $1/\kappa(A)$ αποτελούν μέτρα για το εαν ένας πίνακας είναι αντιστρέψιμος ή οχι. Όμως για την επίλυση ενός γραμμικού συστήματος και κατά πόσο αυτό επηρεάζεται από μικρές διαταραχές στα δεδομένα, μεγαλή σημασία έχει ο δείκτης κατάστασης του πίνακα και όχι η ορίζουσα του. Αναφέρουμε το εξής χαρακτηριστικό παράδειγμα : Θεωρήστε την λύση ενός γραμμικού συστήματος ο πίνακας του οποίου είναι διαγώνιος και όλα τα στοιχεία του ειναι ίσα με $1/2$. Τότε $det(A) = 1/2^n \to 0, n \to \infty$ άρα για μεγαλές τιμές του n ο πίνακας τείνει να γίνει μη-αντιστρέψιμος. Από την άλλη μεριά όμως $\kappa(A) = 1,\ \forall n$ και σε οποιάδηποτε νόρμα. Άρα ο πίνακας είναι ιδανικός για να λύνουμε γραμμικά συστήματα με την έννοια ότι τυχόν μικρές διαταραχές στα δεδομένα ή στις πράξεις δεν θα επιρεάσουν την λύση.

3. Σε πολλές εφαρμογές καταλήγουμε να λύσουμε ένα πολύ μεγάλο γραμμικό σύστημα $Ax = b$, και η χρήση της απαλοιφής Gauss είναι συνήθως ο αλγόριθμος που χρησιμοποιείται. Λόγω της πεπερασμένης ακρίβειας του υπολογιστή εισάγονται σφάλματα-διαταραχές τόσο στη αναπαράσταση των στοιχείων του συστήματος όσο και κατά την εκτέλεση των πράξεων. Η τελική λύση $\tilde{x}$ που δίνει ο υπολογιστής είναι μια προσέγγιση της ακριβούς λύσης x. Το σχετικό σφάλμα της προσεγγίσης αυτής μπορεί να εκτιμηθεί και δίνεται από
$$\frac{\|x - \tilde{x}\|_\infty}{\|x\|_\infty} \leq \frac{2\ u\ \kappa_\infty(A)}{1 - u\ \kappa_\infty(A)}$$
όπου u είναι το *μοναδιαίο σφάλμα στρογγύλευσης* με $u \sim 10^{-7}$ για απλή ακρίβεια και $u \sim 10^{-15}$ για διπλή ακρίβεια. Από την παραπάνω εκτίμηση περιμένουμε ότι αν $u\ \kappa_\infty(A) \sim 1$, δηλαδή αν $\kappa_\infty(A) >>> 1$, τότε το σχετικό σφάλμα στη λύση θα είναι πολυ μεγαλό. (Δείτε το προηγούμενο παράδειγμα με τον 2×2 πίνακα).