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

N1. $||v|| = 0$ εάν και μόνον εάν $v = 0$

N2. Για κάθε $v \in V$ και $a \in \mathbb{K}$, $||av|| = |a|\,||v||$

N3. 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

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

□

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

$$||x|| = \sqrt{x_1^2 + \cdots + x_n^2}\,, \quad x \in \mathbb{R}^n$$

και

$$||z|| = \sqrt{z_1\bar{z}_1 + \cdots + z_n\bar{z}_n}\,, \quad z \in \mathbb{C}^n\,.$$

Τα αξιώματα Ν1 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 $||v - w|| \leq ||v - u|| + ||u - w||$.

Ορισμός 5.2. Θεωρούμε διανυσματικό χώρο V με νόρμα $\|\cdot\|$. Η **απόσταση μεταξύ των σημείων** v, $w \in V$ **ως προς τη νόρμα** $\|\cdot\|$ είναι $\|v - w\|$.
Η **σφαίρα ακτίνας** r **στο** V **ως προς τη νόρμα** $\|\cdot\|$ είναι το σύνολο $S_{\|\cdot\|}(r) = \{v \in V : \|v\| = r\}$. Η σφαίρα ακτίνας 1 ονομάζεται και μοναδιαία σφαίρα του χώρου V με νόρμα$\|\cdot\|$.

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

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

Ελέγχουμε τα αξιώματα:

N1.

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

N2.

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

N3.

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

Δραστηριότητα 5.1 Βρείτε την απόσταση μεταξύ των σημείων (1, 2) και (5, 3) $\in \mathbb{R}^2$ με την ℓ^1-νόρμα.
Σχεδιάστε τη σφαίρα ακτίνας 1 στο $\mathbb{R}^2$ με την ℓ^1-νόρμα.

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

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

Δραστηριότητα 5.2 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

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

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

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

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

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

Δραστηριότητα 5.3 Δείξτε οτι $||f||_\infty$ ικανοποιεί τα αξιώματα της νόρμας.

Παράδειγμα 5.6 Ο περιορισμός της L^2-νόρμας 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 Ν 2 και Ν 3. Όμως μία συνάρτηση με μηδενικό ολοκλήρωμα δεν είναι υποχρεωτικά ίση με 0. Για παράδειγμα, η συνάρτηση $h : [0,\,1] \longrightarrow \mathbb{R}$ με

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

[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

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.

5.2 Ισοδύναμες Νόρμες

Ορισμός 5.3. Έστω V ένας διανυσματικός χώρος. Δύο νόρμες $\|\cdot\|, \|\cdot\|'$ στον V λέγονται **ισοδύναμες** αν υπάρχουν θετικές σταθερές m και M τέτοιες ώστε, για κάθε $x \in V$,

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

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

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

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

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

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

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

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

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

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 συγκεκριμμένο τρόπο.

Ορισμός 5.4. Μια απεικόνιση $\mathcal{M}(n, \mathbb{R}) \longrightarrow \mathbb{R} : A \mapsto \|A\|$ λέγεται **νόρμα πινάκων**(ή **στάθμη πινάκων**) εαν

ΝΠ1. $\|A\| = 0$ εάν και μόνον εάν $A = 0$

ΝΠ2. Για κάθε $A \in \mathcal{M}(n, \mathbb{R})$ και $a \in \mathbb{R}$, $\|av\| = |a|\,\|v\|$

ΝΠ3. Για κάθε $A, B \in \mathcal{M}(n, \mathbb{R})$, $\|A + B\| \leq \|A\| + \|B\|$ (τριγωνική ανισότητα)

ΝΠ4. 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 :

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

Όταν ο πίνακας A έχει πραγματικές συνιστώσες, $\|A\|_F == \sqrt{\operatorname{tr}(A^T A)}$. Όταν ο πίνακας A έχει μιγαδικές συνιστώσες, $\|A\|_F == \sqrt{\operatorname{tr}(A^T \overline{A})}$.

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

Ορισμός 5.5. 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 οτι η 5.1 ικανοποιεί όλες τις ιδιότητες του ορισμού 5.4. Για κάθε $A, B \in \mathcal{M}(n, \mathbb{R})$ και $\lambda \in \mathbb{R}$,

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

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

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

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

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

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

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

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

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 5.6. 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 $\mu\epsilon\gamma\acute{\alpha}\lambda\eta$ 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 συστημάτος 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).

5.5 Εσωτερικό γινόμενο

Στη συνέχεια θα χρησιμοποιούμε το συμβολισμό του συζυγούς, $\bar{a}$, κατανοώντας οτι εάν το σώμα $\mathbb{K}$ είναι οι πραγματικοί αριθμοί, τότε $\bar{a} = a$.

Ορισμός 5.7. Θεωρούμε διανυσματικό χώρο V πάνω από το σώμα $\mathbb{K}$ ($= \mathbb{R}$ ή $\mathbb{C}$). Μια απεικόνιση

$$V \times V \to \mathbb{K} : (v, w) \mapsto \langle v, w \rangle$$

ονομάζεται **εσωτερικό γινόμενο** εάν

ΕΓ 1. Είναι γραμμική στην πρώτη μεταβλητή, δηλαδή εάν για κάθε $u, v, w \in V$ και $a \in \mathbb{K}$,

$$\langle u + v, w \rangle = \langle u, w \rangle + \langle v, w \rangle$$

και

$$\langle av, w \rangle = a\langle v, w \rangle .$$

ΕΓ 2. Για κάθε $v, w \in V$,

$$\langle v, w \rangle = \overline{\langle w, v \rangle}$$

ΕΓ 3. Για κάθε $v \in V$, εάν $v \neq 0$, τότε $\langle v, v \rangle > 0$.

Εάν $\langle v, w \rangle = 0$, λέμε οτι τα διανύσματα v και w είναι **ορθογώνια**.

Παρατηρούμε οτι εάν το σώμα $\mathbb{K} = \mathbb{R}$, τότε η ιδιότητα ΕΓ 2 σημαίνει οτι το εσωτερικό γινόμενο είναι συμμετρικό, και μαζί με την ΕΓ 1, οτι είναι γραμμικό και στη δεύτερη μεταβλητή.

Αντιθέτως, εάν $\mathbb{K} = \mathbb{C}$, για τη δεύτερη μεταβλητή έχουμε

$$\begin{aligned} \langle v, aw \rangle &= \overline{\langle aw, v \rangle} \\ &= \overline{a\langle w, v \rangle} \\ &= \bar{a}\langle v, w \rangle . \end{aligned}$$

Παράδειγμα 5.11 Στο $\mathbb{R}^n$ ορίζεται το ευκλείδειο εσωτερικό γινόμενο: εάν $x = (x_1, \dots, x_n)$, $y = (y_1, \dots, y_n)$,

$$\langle x, y \rangle = \sum_{i=1}^{n} x_i y_i .$$

Τα διανύσματα (x_1, x_2) και $(-x_2, x_1)$ είναι ορθογώνια διανύσματα στο $\mathbb{R}^2$ με το ευκλείδειο εσωτερικό γινόμενο.

Παράδειγμα 5.12 Στο $\mathbb{C}^n$ ορίζεται εσωτερικό γινόμενο για $z = (z_1, \dots, z_n)$, $w = (w_1, \dots, w_n)$,

$$\langle z, w \rangle = \sum_{i=1}^{n} z_i \bar{w}_i .$$

Δραστηριότητα 5.4 Τα διανύσματα $(z_1,\, z_2)$ και $(-z_2,\, z_1)$ *δεν* είναι ορθογώνια στο $\mathbb{C}^2$ με αυτό το εσωτερικό γινόμενο. Βρείτε ένα μη μηδενικό διάνυσμα ορθογώνιο στο $(z_1,\, z_2)$

Παράδειγμα 5.13 Θεωρούμε πίνακα $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, και ορίζουμε την απεικόνιση $f : K^2 \times K^2 \longrightarrow \mathbb{K}$,

$$\begin{aligned} f(v,\, w) &= v^T A \bar{w} && (5.14) \\ &= [v_1,\, v_2] \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} \bar{w}_1 \\ \bar{w}_2 \end{bmatrix} && (5.15) \\ &= a v_1 \bar{w}_1 + b v_1 \bar{w}_2 + c v_2 \bar{w}_1 + d v_2 \bar{w}_2\,. && (5.16) \end{aligned}$$

Ελέγξτε οτι ικανοποιείται η ιδιότητα ΕΓ1. Για την ΕΓ2, $a w_1 \bar{v}_1 + b w_1 \bar{v}_2 + c w_2 \bar{v}_1 + d w_2 \bar{v}_2 = \bar{a} \bar{v}_1 w_1 + \bar{b} \bar{v}_1 w_2 + \bar{c} \bar{v}_2 w_1 + \bar{d} \bar{v}_2 w_2$πρέπει να απαιτήσουμε $a = \bar{a}$, $d = \bar{d}$ και $c = \bar{b}$. Τέλος, για να ισχύει η ΕΓ3, πρέπει να ισχύει $a v_1 \bar{v}_1 + b v_1 \bar{v}_2 + c v_2 \bar{v}_1 + d v_2 \bar{v}_2 > 0$ για κάθε $v \in K^2$, $v \neq 0$. Θεωρούμε $v = (1,\, 0)$ ή $v = (0,\, 1)$, και έχουμε $a > 0$, $d > 0$. Μπορούμε να δείξουμε οτι f ικανοποιεί την ιδιότητα ΕΓ3 εάν και μόνον εάν $ad - b\bar{b} > 0$. (Δες την Άσκηση 5.10 για την πραγματική περίπτωση.)

Παράδειγμα 5.14 Στο χώρο των πολυωνύμων $\mathbb{K}[x]$ ορίζουμε το εσωτερικό γινόμενο

$$\langle p(x),\, q(x) \rangle = \int_0^1 p(t)\, \overline{q(t)}\, dt \tag{5.17}$$

Δραστηριότητα 5.5 Ελέγξτε οτι η 5.17 πράγματι ορίζει ένα εσωτερικό γινόμενο.

Τα πολυώνυμα $p(x) = (x - \frac{1}{2})^2$ και $q(x) = (x - \frac{1}{2})^3$ είναι ορθογώνια:

$$\begin{aligned} \langle p(x),\, q(x) \rangle &= \int_0^1 (t - \tfrac{1}{2})^2 (t - \tfrac{1}{2})^3 dt \\ &= \int_0^1 (t - \tfrac{1}{2})^5 dt \\ &= 0\,. \end{aligned}$$

Παράδειγμα 5.15 Στο χώρο των συνεχών συναρτήσεων στο διάστημα $[a,\, b]$, με πραγματικές τιμές, $C[a,\, b]$, ή με μιγαδικές τιμές , $C_{\mathbb{C}}[a,\, b]$, ορίζουμε το εσωτερικό γινόμενο

$$\langle f,\, g \rangle = \int_a^b f(s)\, \overline{g(s)}\, ds \tag{5.18}$$

Δραστηριότητα 5.6 Ελέγξτε οτι η 5.18 πράγματι ορίζει ένα εσωτερικό γινόμενο.

Οι συναρτήσεις sin και cos είναι ορθογώνιες στο $C[0, \pi]$:

$$\int_0^{\pi} \sin t \cos t dt = \frac{1}{2}\int_0^{\pi} \sin 2t dt = 0 \,.$$

Θεώρημα 5.4 (Ανισότητα Cauchy-Schwarz) *Σε ένα χώρο με εσωτερικό γινόμενο ισχύει, για κάθε v, w:*

$$|\langle v,\, w\rangle| \le \sqrt{\langle v,\, v\rangle}\sqrt{\langle w,\, w\rangle}\,.$$

Απόδειξη. Εάν $w = 0$, τότε και οι δύο πλευρές μηδενίζονται και η σχέση επαληθεύεται.

Υποθέτουμε οτι $w \neq 0$ και θεωρούμε, για $a \in \mathbb{K}$, το διάνυσμα $v - aw$:

$$\begin{aligned} 0 &\le \langle v - aw,\, v - aw\rangle \\ &= \langle v,\, v\rangle - \langle v,\, aw\rangle - \langle aw,\, v\rangle + \langle aw,\, aw\rangle \\ &= \langle v,\, v\rangle - \bar{a}\langle v,\, w\rangle - a\overline{\langle v,\, w\rangle} + a\bar{a}\langle w,\, w\rangle\,. \end{aligned}$$

Ειδικότερα, για $a = \frac{\langle v, w\rangle}{\langle w, w\rangle}$ έχουμε

$$\langle v,\, v\rangle \ge \frac{\langle v,\, w\rangle\overline{\langle v,\, w\rangle}}{\langle w,\, w\rangle}$$

ή

$$|\langle v,\, w\rangle|^2 \le \langle v,\, v\rangle\langle w,\, w\rangle\,,$$

και αφού οι πραγματικοί αριθμοί $\langle v,\, v\rangle$, $\langle w,\, w\rangle$ και $|\langle v,\, w\rangle|$ είναι θετικοί ή μηδέν, έχουμε

$$|\langle v,\, w\rangle| \le \sqrt{\langle v,\, v\rangle}\sqrt{\langle w,\, w\rangle}\,.$$

□

Ειδικές περιπτώσεις της ανισότητας Cauchy-Schwarz είναι οι ακόλουθες ανισότητες.

Στο χώρο $\mathbb{R}^n$ ή $\mathbb{C}^n$, με το συνήθες εσωτερικό γινόμενο $\langle x,\, y\rangle = \sum_{i=1}^n x_i\bar{y}_i$,

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

Στο χώρο $C[a,\, b]$, με εσωτερικό γινόμενο $\langle f,\, g\rangle = \int_a^b f(s)g(s)ds$,

$$\left|\int_a^b f(s)g(s)ds\right| \le \left(\int_a^b (f(s))^2\, ds\right)^{1/2} \left(\int_a^b (g(s))^2\, ds\right)^{1/2}.$$

Πρόταση 5.5 *Εάν V είναι χώρος με εσωτερικό γινόμενο, τότε ορίζεται μία νόρμα στο V:*

$$||v|| = \sqrt{\langle v,\, v\rangle}\,.$$

Απόδειξη. Η απόδειξη των Ν 1 και Ν 2 είναι απλή. Για να αποδείξουμε την τριγωνική ανισότητα Ν 3, παρατηρούμε οτι

$$\begin{aligned} ||v+w||^2 &= \langle v+w,\, v+w\rangle \\ &= ||v||^2 + ||w||^2 + 2\mathrm{Re}\,\langle v,\, w\rangle\,, \end{aligned}$$

και οτι

$$(||v|| + ||w||)^2 = ||v||^2 + ||w||^2 + 2||v||\,||w||\,.$$

Αλλά από την ανισότητα Cauchy-Schwarz, $\langle v,\, w\rangle \leq ||v||\,||w||$ και συνεπώς

$$||v+w||^2 \leq (||v|| + ||w||)^2\;.$$

Αφού οι πραγματικοί αριθμοί $||v+w||$, $||v||$ και $||w||$ είναι θετικοί ή μηδέν, έχουμε

$$||v+w|| \leq ||v|| + ||w||\,.$$

□

Με τον συμβολισμό της νόρμας, η ανισότητα Cauchy-Schwarz γράφεται στη μορφή

$$|\langle v,\, w\rangle| \leq ||v||\,||w||\,.$$

Θεωρώντας τη νόρμα ως το μήκος του διανύσματος, μπορούμε να ορίσουμε τη **γωνία μεταξύ δύο διανυσμάτων** σε οποιοδήποτε πραγματικό διανυσματικό χώρο με εσωτερικό γινόμενο. Από την ανισότητα Cauchy-Schwarz έχουμε

$$-1 \leq \frac{\langle v,\, w\rangle}{||v||\,||w||} \leq 1\,,$$

και συνεπώς υπάρχει $\vartheta \in [0,\, \pi]$ τέτοιο ώστε

$$\cos\vartheta = \frac{\langle v,\, w\rangle}{||v||\,||w||}\,. \tag{5.19}$$

Αυτός ο ορισμός γωνίας ταιριάζει με την έννοια της ορθογωνιότητας, που έχουμε ορίσει σε οποιοδήποτε χώρο με εσωτερικό γινόμενο, πραγματικό ή μιγαδικό. Όταν $\langle v,\, w\rangle = 0$, $\vartheta = \frac{\pi}{2}$ και τα διανύσματα v και w σχηματίζουν ορθή γωνία.

Στην ευκλείδεια γεωμετρία, ο 'νόμος του παραλληλογράμμου' λέει οτι σε ένα παραλληλόγραμμο, το άθροισμα των τετραγώνων των τεσσάρων πλευρών ισούται με το άθροισμα των τετραγώνων των διαγωνίων. Όταν το παραλληλόγραμμο είναι ορθογώνιο, αυτό είναι ισοδύναμο με το Πυθαγόρειο Θεώρημα. Θα δούμε οτι ένα ανάλογο αποτέλεσμα ισχύει σε κάθε διανυσματικό χώρο με εσωτερικό γινόμενο.

Πρόταση 5.6 *Έστω V ένας χώρος με εσωτερικό γινόμενο.*

1. ***Νόμος του Παραλληλογράμμου.*** *Για κάθε $v,\, w \in V$ ισχύει η ισότητα*

$$||v + w||^2 + ||v - w||^2 = 2||v||^2 + 2||w||^2\,.$$

2. ***Πυθαγόρειο Θεώρημα.*** *Εάν $v_1,\, \ldots,\, v_k \in V$ και $\langle v_i,\, v_j\rangle = 0$ όταν $i \neq j$, τότε*

$$||v_1 + \cdots + v_k||^2 = ||v_1||^2 + \cdots + ||v_k||^2\,.$$

Απόδειξη. Σε ένα χώρο με εσωτερικό γινόμενο έχουμε $||v||^2 = \langle v,\, v\rangle$. Συνεπώς

$$||v + w||^2 = \langle v + w,\, v + w\rangle = \langle v,\, v\rangle + \langle v,\, w\rangle + \langle w,\, v\rangle + \langle w,\, w\rangle\,, \tag{5.20}$$

$$||v - w||^2 = \langle v - w,\, v - w\rangle = \langle v,\, v\rangle - \langle v,\, w\rangle - \langle w,\, v\rangle + \langle w,\, w\rangle\,. \tag{5.21}$$

Προσθέτοντας τις 5.20 και 5.21 έχουμε τον 'νόμο του παραλληλογράμμου'.

$$||v + w||^2 + ||v - w||^2 = 2\langle v,\, v\rangle + 2\langle w,\, w\rangle\,.$$

Για $k = 2$ το 'Πυθαγόρειο Θεώρημα' προκύπτει από την 5.20, αφού $\langle v_1,\, v_2\rangle = 0$. Η γενική περίπτωση αποδεικνύεται με επαγωγή στο k. Αφού $\langle v_1 + \cdots + v_{k-1},\, v_k\rangle = 0$,

$$\begin{aligned} ||v_1 + \cdots + v_k||^2 &= ||v_1 + \cdots + v_{k-1}||^2 + ||v_k||^2 \\ &= ||v_1||^2 + \cdots + ||v_{k-1}||^2 + ||v_k||^2\,. \end{aligned}$$

□

Παράδειγμα 5.16 Θα δείξουμε οτι η ℓ^1 νόρμα στο $\mathbb{R}^2$ δεν ικανοποιεί το νόμο του παραλληλογράμμου. Θεωρούμε τα διανύσματα $v = (1,\, 0)$ και $w = (0,\, 1)$. Τότε $||v||_1 = |1| + |0| = 1$, $||w||_1 = 1$, και $||v + w||_1 = 2$, $||v - w||_1 = 2$. Άρα $2||v||^2 + 2||w||^2 = 4$ ενώ $||v + w||^2 + ||v - w||^2 = 8$.

Αφού κάθε νόρμα που προέρχεται από εσωτερικό γινόμενο ικανοποιεί το Νόμο του Παραλληλογράμμου, συμπεραίνουμε οτι η νόρμα ℓ^1 δεν προκύπτει από εσωτερικό γινόμενο.

5.6 Ορθοκανονικά σύνολα διανυσμάτων

Ένα σύνολο διανυσμάτων $S = \{v_1, \ldots, v_n\}$ ονομάζεται **ορθογώνιο** εάν τα στοιχεία του είναι ορθογώνια ανά δύο, δηλαδή εάν για κάθε $i,\, j = 1, \ldots, n, i \neq j$,

$$\langle v_i,\, v_j\rangle = 0\,.$$

Εάν επί πλέον, για κάθε $i = 1, \ldots, n$, $||v_i|| = 1$, το σύνολο ονομάζεται **ορθοκανονικό**. Χρησιμοποιώντας το συμβολισμό δ του Kronecker,

$$\delta_{ij} = \left[\begin{array}{ll} 0 & i \neq j \\ 1 & i = j \end{array}\right.$$

βλέπουμε οτι το σύνολο S είναι ορθοκανονικό εάν και μόνον εάν, για κάθε $i,\, j = 1, \ldots, n$,

$$\langle v_i,\, v_j \rangle = \delta_{ij}\,.$$

Λήμμα 5.7 *Ένα ορθοκανονικό σύνολο είναι γραμμικά ανεξάρτητο.*

Απόδειξη. Υποθέτουμε οτι $\{v_1, \ldots, v_n\}$ είναι ένα ορθοκανονικό σύνολο, και οι αριθμοί $a_1, \ldots, a_n$ είναι τέτοιοι ώστε

$$a_1 v_1 + \cdots + a_n v_n = 0\,.$$

Για κάθε $j = 1, \ldots, n$, έχουμε

$$\begin{aligned} 0 &= \langle 0,\, v_j \rangle \\ &= \left\langle \sum_{i=1}^{n} a_i v_i,\, v_j \right\rangle \\ &= \sum_{i=1}^{n} a_i \langle v_i,\, v_j \rangle \\ &= \sum_{i=1}^{n} a_i \delta_{ij} \\ &= a_j\,. \end{aligned}$$

Συνεπώς $a_j = 0$ για κάθε $j = 1, \ldots, n$, και το σύνολο $\{v_1, \ldots, v_n\}$ είναι γραμμικά ανεξάρτητο.

□

Ένα ορθοκανονικό σύνολο αποτελεί μία ιδιαίτερα χρήσιμη βάση για το χώρο τον οποίο παράγει. Οι συντεταγμένες ενός διανύσματος ως προς μία ορθοκανονική βάση δίδονται απλώς από τα εσωτερικά γινόμενα του διανύσματος με τα διανύσματα της βάσης. Πράγματι, εάν $\{v_1, \ldots, v_n\}$ είναι ορθοκανονική βάση, και

$$v = a_1 v_1 + \cdots + a_n v_n$$

τότε

$$\begin{aligned} \langle v,\, v_j \rangle &= \left\langle \sum_{i=1}^{n} a_i v_i\,,\, v_j \right\rangle \\ &= \sum_{i=1}^{n} a_i \langle v_i,\, v_j \rangle \\ &= \sum_{i=1}^{n} a_i \delta_{ij} \\ &= a_j\,. \end{aligned}$$

Σε ένα χώρο πεπερασμένης διάστασης με εσωτερικό γινόμενο, μπορούμε πάντα να κατασκευάσουμε μια ορθοκανονική βάση, εφαρμόζοντας τη διαδικασία **ορθοκανονικοποίησης Gram-Schmidt**.

Θεώρημα 5.8 (Ορθοκανονικοποίηση Gram-Schmidt) *Θεωρούμε χώρο V με εσωτερικό γινόμενο, και ένα γραμμικά ανεξάρτητο σύνολο $\{v_1, \ldots, v_n\}$. Τότε υπάρχει ορθοκανονικό σύνολο $\{e_1, \ldots, e_n\}$ τέτοιο ώστε για κάθε $i = 1, \ldots, n$*

$$e_i \in \langle v_1, \ldots, v_i \rangle$$

και

$$\langle e_1, \ldots, e_n \rangle = \langle v_1, \ldots, v_n \rangle \, .$$

Στην απόδειξη του Θεωρήματος ο συμβολισμός $\langle \ldots \rangle$ χρησιμοποιείται για να δηλώσει τόσο το εσωτερικό γινόμενο όσο και τον παραγόμενο υπόχωρο, αλλά η διάκριση είναι συνήθως εύκολη από τα συμφραζόμενα.

Απόδειξη. Για κάθε j ορίζουμε πρώτα το διάνυσμα e'_j ορθογώνιο προς τα e_i για $i = 1, \ldots, j-1$, και μετά διαιρούμε με τη νόρμα του e'_j για να πάρουμε το μοναδιαίο διάνυσμα e_j. ένα σύνολο μη μηδενικών ορθογωνίων διανυσμάτων $e'_1, \ldots, e'_n$, και στη συνέχεια θα ορίσουμε τα μοναδιαία διανύσματα

$$e_i = \frac{1}{||e'_i||} \, e'_i \, .$$

Αφού τα $\{v_1, \ldots, v_n\}$ είναι γραμμικά ανεξάρτητα, $v_1 \neq 0$, και ορίζουμε

$$e'_1 = v_1 \, , \qquad e_1 = \frac{1}{||e'_1||} \, e'_1 \, .$$

Το e'_2 προκύπτει από το v_2 αφαιρώντας κατάλληλο πολλαπλάσιο του e_1 ώστε e'_2 να είναι ορθογώνιο προς το e_1.

$$e'_2 = v_2 - \langle v_2, e_1 \rangle e_1 \, .$$

Πράγματι

$$\begin{aligned}
\langle e'_2, e'_1 \rangle &= \langle v_2 - \langle v_2, e_1 \rangle \, e_1, e_1 \rangle \\
&= \langle v_2, e_1 \rangle - \langle v_2, e_1 \rangle \, \langle e_1, e_1 \rangle \\
&= 0
\end{aligned}$$

Παρατηρούμε οτι αφού τα e_1, v_2 είναι γραμμικά ανεξάρτητα, $e'_2 \neq 0$ και μπορούμε να διαιρέσουμε το e'_2 με τη νόρμα του για να πάρουμε το μοναδιαίο

$$e_2 = \frac{1}{||e'_2||} \, e'_2 \, .$$

Παρατηρούμε οτι τα e_1, e_2 παράγουν τον ίδιο υπόχωρο που παράγουν τα v_1, v_2.

Στη συνέχεια, για $j = 3, \ldots, n$, ορίζουμε αναδρομικά τα μη μηδενικά διανύσματα

$$e'_j = v_j - \sum_{i=1}^{j-1} \langle v_j, e_i \rangle \, e_i \, , \tag{5.22}$$

τα οποία ικανοποιούν $\langle e_i, e'_j \rangle = 0$ για $i = 1, \ldots, j-1$. Το e'_j δεν είναι μηδενικό, άρα μπορούμε να το διαιρέσουμε με τη νόρμα του για να πάρουμε το διάνυσμα

$$e_j = \frac{1}{||e'_j||} e'_j \,.$$

Από την 5.22 είναι φανερό οτι

$$e_j \in \langle v_1, \ldots, v_j \rangle$$

και οτι

$$v_j \in \langle e_1, \ldots, e_j \rangle \,.$$

Συνεπώς, για κάθε $j = 1, \ldots, n$

$$\langle e_1, \ldots, e_j \rangle = \langle v_1, \ldots, v_j \rangle \,.$$

□

Παρατήρηση Όταν κάνουμε υπολογισμούς με το χέρι, είναι συχνά προτιμότερο να υπολογίσουμε όλα τα e'_j, ακολουθώντας τη διαδικασία Gram – Schmidt όπως την περιγράψαμε στο Παράδειγμα 1.9, και στο τέλος να διαρέσουμε κάθε διάνυσμα e'_i με τη νόρμα του. Με αυτόν τον τρόπο αποφεύγουμε να μεταφέρουμε σε όλη τη διαδικασία τις τετραγωνικές ρίζες που προκύπτουν από τη νόρμα.

Πολυώνυμα Legendre

Η συνήθης βάση του χώρου των πολυωνύμων, $\{1, x, x^2, \ldots\}$, δεν είναι ορθογώνια για οποιοδήποτε εσωτερικό γινόμενο της μορφής $\langle p, q \rangle = \int_a^b p(t)\bar{q}(t)dt$.

Για να ορίσουμε μία ορθοκανονική οικογένεια πολυωνύμων, που θα μας επιτρέπει να προσεγγίζουμε συναρτήσεις από τις προβολές τους στα στοιχεία αυτής της οικογένειας, διευκολύνει να χρησιμοποιήσουμε ένα εσωτερικό γινόμενο που ορίζεται στο διάστημα $[-1, 1]$. Με αυτή την επιλογή, κάθε άρτιο πολυώνυμο, για το οποίο $p(x) = p(-x)$, είναι εξ αρχής ορθογώνιο σε κάθε περιττό πολυώνυμο, για το οποίο $p(x) = -p(-x)$.

Θεωρούμε λοιπόν το εσωτερικό γινόμενο στο χώρο $\mathbb{R}[x]$,

$$\langle p, q \rangle = \int_{-1}^{1} p(t)q(t)dt \,,$$

και τα πολυώνυμα $p_0(x) = 1$, $p_1(x) = x$, $p_2(x) = x^2$, $\ldots$, $p_n(x) = x^n$, $\ldots$.

Ορίζουμε

$$w_0(x) = p_0(x) = 1 \,.$$

Έχουμε $\langle p_1, w_0 \rangle = \int_{-1}^{1} t\, dt = 0$. Άρα

$$w_1(x) = p_1(x) - \frac{\langle p_1, w_0 \rangle}{\langle w_0, w_0 \rangle} w_0(x) = x \,.$$

Για να βρούμε το w_2, υπολογίζουμε τα εσωτερικά γινόμενα

$$\begin{aligned}
\langle p_2,\, w_1\rangle &= \int_{-1}^{1} t^3\, dt = 0\,,\\
\langle p_2,\, w_0\rangle &= \int_{-1}^{1} t^2\, dt = \frac{1}{3}t^3\Big|_{-1}^{1} = \frac{2}{3}\,,\\
\langle w_0,\, w_0\rangle &= \int_{-1}^{1} 1\, dt = 2\,,\\
\langle w_1,\, w_1\rangle &= \int_{-1}^{1} t^2\, dt = \frac{2}{3}\,.
\end{aligned}$$

Έτσι έχουμε

$$\begin{aligned}
w_2(x) &= p_2(x) - \frac{\langle p_2,\, w_1\rangle}{\langle w_1,\, w_1\rangle} w_1(x) - \frac{\langle p_2,\, w_0\rangle}{\langle w_0,\, w_0\rangle} w_0(x)\\
&= x^2 - \frac{2/3}{2} = x^2 - \frac{1}{3}\,.
\end{aligned}$$

και

$$\langle w_2,\, w_2\rangle = \int_{-1}^{1}\left(t^2 - \frac{1}{3}\right)^2 dt = \frac{8}{45}\,.$$

Όταν διαιρέσουμε τα w_0, w_1, w_2 με τη νόρμα τους, έχουμε

$$\begin{aligned}
q_0(x) &= \frac{1}{\sqrt{2}}\,,\\
q_1(x) &= \frac{\sqrt{3}}{\sqrt{2}}x\,,\\
q_2(x) &= \frac{3\sqrt{5}}{\sqrt{8}}\left(x^2 - \frac{1}{3}\right) = \frac{\sqrt{5}}{\sqrt{8}}(3x^2 - 1)\,.
\end{aligned}$$

Αυτά συνδέονται με τα πολυώνυμα Legendre,

$$P_m(x) = \frac{1}{2^m\, m!}\frac{d^m}{dx^m}(x^2 - 1)^m\,.$$

Για $m = 0,\, \ldots,\, n-1$, τα πολυώνυμα

$$q_m(x) = \sqrt{\frac{2m+1}{2}}P_m(x)$$

είναι η ορθοκανονική βάση του $\mathbb{R}_{n-1}[x]$ που προκύπτει με τη διαδικασία Gram – Schmidt, για το εσωτερικό γινόμενο στο διάστημα $[-1,\, 1]$, από τη συνήθη βάση $\{1,\, x,\, x^2,\, \ldots,\, x^{n-1}\}$.

Άσκηση 5.1 Θεωρούμε διανύσματα $v = (v_1, v_2)$ και $w = (w_1, w_2)$ στο $\mathbb{R}^2$.

1. Δείξτε οτι η συνάρτηση
$$(v, w) = 4v_1w_1 + 9v_2w_2$$
ορίζει εσωτερικό γινόμενο στο $\mathbb{R}^2$.

2. Δείξτε οτι η συνάρτηση
$$(v, w) = 2v_1w_1 - v_2w_2$$
δεν ορίζει εσωτερικό γινόμενο.

Άσκηση 5.2 Θεωρούμε το χώρο $C[0, 1]$ των συνεχών συναρτήσεων στο διάστημα $[0, 1]$, με εσωτερικό γινόμενο
$$(f, g) = \int_0^1 f(t)\, g(t)\, dt$$

1. Βρείτε το εσωτερικό γινόμενο των $f(x) = 2x + 1$, $g(x) = 3x - 2$.

2. Δείξτε οτι οι συναρτήσεις $f(x) = x^2$ και $g(x) = 4x - 3$ είναι ορθογώνιες.

3. Βρείτε μία συνάρτηση ορθογώνια προς την $f(x) = 6x + 12$

Άσκηση 5.3 Θεωρούμε το μιγαδικό διανυσματικό χώρο $\mathbb{C}^2$, με το σύνηθες εσωτερικό γινόμενο. Βρείτε τα (u, v), $||u||$, $||v||$ και την απόσταση $d(u, v) = ||u - v||$ για τα διανύσματα:

1. $u = (2 - \mathrm{i}, 3 + 2\mathrm{i})$, $\quad v = (3 - 2\mathrm{i}, 2 + \mathrm{i})$.

2. $u = (2 - 3\mathrm{i}, -2 + 3\mathrm{i})$, $\quad v = (1, 1)$.

Άσκηση 5.4 Στο χώρο $\mathbb{R}^3$, με το ευκλείδειο εσωτερικό γινόμενο, βρείτε την ορθοκανονική βάση που προκύπτει από την εφαρμογή της διαδικασίας Gram-Schmidt στη βάση
$$\{(1, 1, 0), (2, 1, 0), (0, 1, 2)\}\,.$$

Άσκηση 5.5 Θεωρήστε θετική συνεχή συνάρτηση $f : [0, 1] \to \mathbb{R}$. Δείξτε οτι εάν $p(x)$, $q(x)$ είναι πολυώνυμα, τότε

$$\langle p, q \rangle_f = \int_0^1 f(t)p(t)q(t)dt\,,$$

ορίζει εσωτερικό γινόμενο στο χώρο των πολυωνύμων με πραγματικούς συντελεστές. Εάν $f(x) = x + 1$, βρείτε ορθοκανονική βάση για το χώρο των πολυωνύμων βαθμού το πολύ 1, με το εσωτερικό γινόμενο $\langle\ ,\ \rangle_f$.

Άσκηση 5.6 Βρείτε όλα τα διαφορετικά εσωτερικά γινόμενα που ορίζονται σε ένα διανυσματικό χώρο διάστασης 2 πάνω από το $\mathbb{R}$.

Άσκηση 5.7 Θεωρούμε διανυσματικό χώρο V πάνω από το $\mathbb{R}$, με εσωτερικό γινόμενο, και δύο διαφορετικά διανύσματα $a, b \in V$. Αποδείξτε οτι εάν $x \in V$ και $||x - a|| + ||x - b|| = ||a - b||$, τότε $x = \lambda a + \mu b$, με $\lambda, \mu \in \mathbb{R}$ και $\lambda + \mu = 1$.

Άσκηση 5.8 Θεωρούμε τον διανυσματικό χώρο $\mathbb{R}^3$, με το κανονικό εσωτερικό γινόμενο.

1. Επαληθεύστε οτι τα διανύσματα

   $$v_1 = (1, 1, 1),\ v_2 = (1, 2, -3),\ v_3 = (5, -4, -1)$$

   είναι ανά δύο ορθογώνια, και βρείτε μία ορθοκανονική βάση του V, διαφορετική από την κανονική.

2. Βρείτε τα μοναδιαία διανύσματα τα οποία είναι ταυτόχρονα ορθογώνια με τα $v_1 - v_2$ και $v_1 + v_3$.

3. Βρείτε τα διανύσματα τα οποία είναι ορθογώνια στο $2v_2 + v_3$ και ανήκουν στον γραμμικό υπόχωρο που παράγεται από τα $v_1 - v_2$, $v_1 + v_3$.

Άσκηση 5.9 Θεωρούμε τον διανυσματικό χώρο $\mathbb{R}^4$ με το κανονικό εσωτερικό γινόμενο, και τον υπόχωρο X που παράγεται από τα διανύσματα $u_1 = (1, 1, 0, 0)$ και $u_2 = (0, 1, -1, 1)$.
Βρείτε μία ορθοκανονική βάση του ορθογωνίου συμπληρώματος $X^{\perp}$, και συμπληρώστε την σε μια ορθοκανονική βάση του $\mathbb{R}^4$.

Άσκηση 5.10 Έστω 2×2 πραγματικός πίνακας A. Δείξτε οτι η απεικόνιση

$$\langle x,\, y \rangle = x^T A y = [x_1,\, x_2] \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}$$

ορίζει ένα εσωτερικό γινόμενο στο $\mathbb{R}^2$ εάν και μόνον εάν $A^T = A$, $\det A > 0$ και $\operatorname{tr} A > 0$, (δηλαδή εάν $b = c$, $ad - b^2 > 0$ και $a > 0$).

5.7 Ερμιτιανοί τελεστές

Για έναν τελεστή L σε ένα διανυσματικό χώρο πεπερασμένης διάστασης, έχουμε δει οτι εάν υπάρχει βάση του V η οποία αποτελείται από ιδιοδιανύσματα του L τότε ο πίνακας του L ως προς αυτή τη βάση είναι διαγώνιος. Τώρα θα δούμε οτι σε χώρους με εσωτερικό γινόμενο μπορούμε να δώσουμε συγκεκριμένα κριτήρια για να συμβαίνει αυτό και μάλιστα η βάση να αποτελείται από ορθογώνια διανύσματα.

Σε αυτό το κεφάλαιο όλοι οι διανυσματικοί χώροι είναι πάνω από το σώμα $\mathbb{C}$ ή το σώμα $\mathbb{R}$.

Ορισμός 5.8. Θεωρούμε ένα διανυσματικό χώρο V με εσωτερικό γινόμενο και ένα γραμμικό τελεστή $L : V \to V$. Ο τελεστής L ονομάζεται **ερμιτιανός** εάν για κάθε $u, v \in V$,

$$\langle L(u), v\rangle = \langle u, L(v)\rangle .$$

Ένας ερμιτιανός τελεστής σε ένα πραγματικό διανυσματικό χώρο ονομάζεται **συμμετρικός**.

Παράδειγμα 5.17 Θεωρούμε τον τελεστή $L : \mathbb{R}^2 \to \mathbb{R}^2$, $L(x, y) = (x+2y, 2x)$. Ως προς το ευκλείδειο εσωτερικό γινόμενο έχουμε

$$\langle L(u_1, u_2), (v_1, v_2)\rangle = \langle (u_1 + 2u_2, 2u_1), (v_1, v_2)\rangle = u_1v_1 + 2u_2v_1 + 2u_1v_2$$

και

$$\langle (u_1, u_2), L(v_1, v_2)\rangle = \langle (u_1, u_2), (v_1 + 2v_2, 2v_1)\rangle = u_1v_1 + 2u_1v_2 + 2u_2v_1 .$$

Ο τελεστής L είναι συμμετρικός.

Παράδειγμα 5.18 Θεωρούμε τον τελεστή $M : \mathbb{C}^2 \to \mathbb{C}^2$, $L(z, w) = (z + \mathrm{i}w, -\mathrm{i}z)$. Ως προς το συνηθισμένο εσωτερικό γινόμενο στο $\mathbb{C}^2$ έχουμε

$$\langle L(u_1, u_2), (v_1, v_2)\rangle = \langle (u_1 + \mathrm{i}u_2, -\mathrm{i}u_1), (v_1, v_2)\rangle = u_1\bar{v}_1 + \mathrm{i}u_2\bar{v}_1 - \mathrm{i}u_1\bar{v}_2$$

και

$$\langle (u_1, u_2), L(v_1, v_2)\rangle = \langle (u_1, u_2), (v_1 + \mathrm{i}v_2, -\mathrm{i}v_1)\rangle = u_1\bar{v}_1 + u_1\bar{\mathrm{i}}\bar{v}_2 + u_2(-\bar{\mathrm{i}})\bar{v}_1 .$$

Ο τελεστής L είναι ερμιτιανός.

Ορισμός 5.9. Θεωρούμε έναν τετραγωνικό μιγαδικό πίνακα $A = [a_{ij}]$. Ο **συζυγής** (ή **αναστροφοσυζυγής**) του πίνακα A είναι ο πίνακας $A^* = [b_{ij}]$, όπου

$$b_{ij} = \bar{a}_{ji} .$$

Δηλαδή οι όροι του πίνακα A^* είναι οι μιγαδικοί συζυγείς των όρων του ανάστροφου του A. Εάν ο πίνακας A είναι πραγματικός, τότε $A^* = A^T$.

Παράδειγμα 5.19 Έστω ο πίνακας

$$A = \begin{bmatrix} 1 & \mathrm{i} \\ 2 & 3+\mathrm{i} \end{bmatrix}.$$

Τότε

$$A^T = \begin{bmatrix} 1 & 2 \\ \mathrm{i} & 3+\mathrm{i} \end{bmatrix}, \quad \overline{A} = \begin{bmatrix} 1 & -\mathrm{i} \\ 2 & 3-\mathrm{i} \end{bmatrix}$$

και

$$A^* = (\overline{A})^T = \begin{bmatrix} 1 & 2 \\ -\mathrm{i} & 3-\mathrm{i} \end{bmatrix}.$$

Ορισμός 5.10. Ένας τετραγωνικός μιγαδικός πίνακας A ονομάζεται **ερμιτιανός** εάν είναι ίσος με τον συζυγή του, $A^* = A$.
Ένας ερμιτιανός πίνακας του οποίου όλοι οι όροι είναι πραγματικοί αριθμοί είναι **συμμετρικός**.

Παρατηρούμε οτι τα διαγώνια στοιχεία ενός ερμιτιανού πίνακα είναι πραγματικοί αριθμοί.

Λήμμα 5.9 *Θεωρούμε ερμιτιανό τελεστή $L : V \to V$ σε χώρο πεπερασμένης διάστασης, και ορθοκανονική βάση $\mathcal{B}$ του V. Τότε ο πίνακας $A = [a_{ij}]$ του L ως προς τη βάση $\mathcal{B}$ είναι ερμιτιανός,*

$$a_{ij} = \bar{a}_{ji}\,.$$

Απόδειξη. Θεωρούμε την ορθοκανονική βάση $\mathcal{B} = \{u_1, \ldots, u_n\}$. Εάν $A = [a_{ij}]$ είναι ο πίνακας του L ως προς τη βάση $\mathcal{B}$, τότε για κάθε $j = 1, \ldots, n$,

$$L(u_j) = \sum_{k=1}^{n} a_{kj} u_k\,.$$

Αλλά τότε

$$\begin{aligned} \langle L(u_j), u_i \rangle &= \left\langle \sum_{k=1}^{n} a_{kj} u_k, u_i \right\rangle \\ &= \sum_{k=1}^{n} \langle a_{kj} u_k, u_i \rangle \\ &= \sum_{k=1}^{n} a_{kj} \langle u_k, u_i \rangle \\ &= a_{ij}\,, \end{aligned}$$

αφού η βάση είναι ορθοκανονική. Εξ άλλου,

$$\langle u_j, L(u_i) \rangle = \left\langle u_j, \sum_{k=1}^{n} a_{ki} u_k \right\rangle$$

$$
\begin{aligned}
&= \sum_{k=1}^{n} \langle u_j, \, a_{ki} u_k \rangle \\
&= \sum_{k=1}^{n} \bar{a}_{ki} \langle u_j, \, u_k \rangle \\
&= \bar{a}_{ji} \, .
\end{aligned}
$$

Αφού ο L είναι ερμιτιανός, $a_{ij} = \langle L(u_j), \, u_i \rangle = \langle u_j, \, L(u_i) \rangle = \bar{a}_{ji}$ και ο πίνακας A είναι ίσος με τον συζυγή του.

□

Πρόταση 5.10 *Θεωρούμε μιγαδικό διανυσματικό χώρο V και τελεστή $L : V \to V$. Εάν ο L είναι ερμιτιανός, τότε οι ιδιοτιμές του L είναι πραγματικοί αριθμοί.*

Απόδειξη. Έστω $\lambda \in \mathbb{C}$ μία ιδιοτιμή του L, και $v \in V$ ένα ιδιοδιάνυσμα για την ιδιοτιμή λ, $L(v) = \lambda v$. Τότε

$$\langle L(v), \, v \rangle = \langle \lambda v, \, v \rangle = \lambda \langle v, \, v \rangle \, ,$$

και

$$\langle v, \, L(v) \rangle = \langle v, \, \lambda v \rangle = \bar{\lambda} \langle v, \, v \rangle \, .$$

Αφού $\langle v, \, v \rangle \neq 0$ και ο L είναι ερμιτιανός, $\lambda = \bar{\lambda}$. Άρα η ιδιοτιμή λ είναι πραγματικός αριθμός.

□

Λήμμα 5.11 *Εάν $L : V \to V$ είναι ερμιτιανός τελεστής, τότε τα ιδιοδιανύσματα που αντιστοιχούν σε διαφορετικές ιδιοτιμές είναι ορθογώνια.*

Απόδειξη. Θεωρούμε ιδιοτιμές λ και μ του L και αντίστοιχα ιδιοδιανύσματα u και v. Τότε

$$\lambda \langle u, \, v \rangle = \langle L(u), \, v \rangle = \langle u, \, L(v) \rangle = \bar{\mu} \langle u, \, v \rangle \, .$$

Αφού οι ιδιοτιμές είναι πραγματικές, $\lambda \langle u, \, v \rangle = \mu \langle u, \, v \rangle$, και εάν $\lambda \neq \mu$, $\langle u, \, v \rangle = 0$.

□

5.8 Μοναδιαίοι τελεστές

Ορισμός 5.11. Θεωρούμε ένα διανυσματικό χώρο V με εσωτερικό γινόμενο και ένα γραμμικό τελεστή $L : V \to V$. Ο τελεστής L ονομάζεται **μοναδιαίος** (ή **ορθομοναδιαίος**) εάν διατηρεί το εσωτερικό γινόμενο, δηλαδή εάν για κάθε $u, \, v \in V$,

$$\langle L(u), \, L(v) \rangle = \langle u, \, v \rangle \, .$$

Ένας μοναδιαίος τελεστής σε ένα πραγματικό διανυσματικό χώρο ονομάζεται **ορθογώνιος**.

Παράδειγμα 5.20 Ο τελεστής $L(x, y) = (x\cos\theta - y\sin\theta,\, x\sin\theta + y\cos\theta)$ είναι ορθογώνιος. Ο τελεστής $M(z, w) = (e^{i\theta}z,\, e^{-i\theta}w)$ είναι μοναδιαίος. Ελέγξτε οτι διατηρούν το εσωτερικό γινόμενο στο $\mathbb{R}^2$ και στο $\mathbb{C}^2$ αντίστοιχα.

Ορισμός 5.12. Ένας τετραγωνικός μιγαδικός πίνακας A ονομάζεται **μοναδιαίος** εάν $A^*A = \mathbf{I}_n$.
Ένας μοναδιαίος πίνακας του οποίου όλοι οι όροι είναι πραγματικοί αριθμοί ονομάζεται **ορθογώνιος**.

Παράδειγμα 5.21 Ένας 2×2 πραγματικός πίνακας είναι ορθογώνιος εάν και μόνον εάν είναι της μορφής

$$\begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix} \quad \text{ή} \quad \begin{bmatrix} \cos\theta & \sin\theta \\ \sin\theta & -\cos\theta \end{bmatrix}$$

για κάποιο θ.

Λήμμα 5.12 *Θεωρούμε μοναδιαίο τελεστή $L : V \to V$ σε χώρο πεπερασμένης διάστασης n, και ορθοκανονική βάση $\mathcal{B}$ του V. Τότε ο πίνακας $A = [a_{ij}]$ του L ως προς τη βάση $\mathcal{B}$ είναι μοναδιαίος,*

$$\sum_{k=1}^{n} \bar{a}_{ki} a_{kj} = \delta_{ij}\,.$$

Απόδειξη. Θεωρούμε την ορθοκανονική βάση $\mathcal{B} = \{u_1, \ldots, u_n\}$. Εάν $A = [a_{ij}]$ είναι ο πίνακας του L ως προς τη βάση $\mathcal{B}$, τότε για κάθε $j = 1, \ldots, n$,

$$L(u_j) = \sum_{k=1}^{n} a_{kj} u_k\,.$$

Αφού ο L είναι μοναδιαίος και η βάση $\mathcal{B}$ είναι ορθοκανονική, $\langle L(u_i), L(u_j)\rangle = \langle u_i, u_j\rangle = \delta_{ij}$. Εξ άλλου

$$\begin{aligned}
\langle L(u_i), L(u_j)\rangle &= \left\langle \sum_{k=1}^{n} a_{ki}u_k,\, \sum_{\ell=1}^{n} a_{\ell j}u_\ell \right\rangle \\
&= \sum_{k=1}^{n}\sum_{\ell=1}^{n} a_{ki}\bar{a}_{\ell j}\langle u_k, u_\ell\rangle \\
&= \sum_{k=1}^{n}\sum_{\ell=1}^{n} a_{ki}\bar{a}_{\ell j}\delta_{k\ell} \\
&= \sum_{k=1}^{n} a_{ki}\bar{a}_{kj} \\
&= (A^*A)_{ji}\,.
\end{aligned}$$

Συμπεραίνουμε οτι $A^*A = \mathbf{I}_n$ και ο πίνακας A είναι μοναδιαίος.

□

Πρόταση 5.13 *Θεωρούμε μιγαδικό διανυσματικό χώρο V με εσωτερικό γινόμενο, και τελεστή $L : V \longrightarrow V$. Εάν ο L είναι μοναδιαίος, τότε οι ιδιοτιμές του L είναι μιγαδικοί αριθμοί μέτρου* 1.

Απόδειξη. Έστω $\lambda \in \mathbb{C}$ μία ιδιοτιμή του L, και $v \in V$ ένα ιδιοδιάνυσμα για την ιδιοτιμή λ, $L(v) = \lambda v$. Τότε

$$\langle v, v\rangle = \langle L(v), L(v)\rangle = \langle \lambda v, \lambda v\rangle = \lambda\bar{\lambda}\langle v, v\rangle .$$

Αφού $\langle v, v\rangle \neq 0$, $\lambda\bar{\lambda} = 1$.

□

5.9 Τριγωνοποίηση τελεστών σε χώρο με εσωτερικό γινόμενο.

Στο επόμενο Θεώρημα θα δείξουμε οτι όταν έχουμε χώρο με εσωτερικό γινόμενο, η βάση ως προς την οποία ο πίνακας ενός τελεστή είναι άνω τριγωνικός, μπορεί να επιλεγεί να είναι ορθοκανονική.

Θεώρημα 5.14 (Λήμμα Schur.) *Θεωρούμε διανυσματικό χώρο V πεπερασμένης διάστασης, με εσωτερικό γινόμενο πάνω από το $\mathbb{C}$, και γραμμικό τελεστή $L : V \to V$. Τότε υπάρχει ορθοκανονική βάση του V ως προς την οποία ο L έχει άνω τριγωνικό πίνακα.*

Απόδειξη. Η απόδειξη ακολουθεί τα βήματα του Θεωρήματος Τριγωνοποίησης, Θεώρημα 3.25. Πρέπει να δείξουμε οτι εάν V είναι χώρος με εσωτερικό γινόμενο, μπορούμε να επιλέξουμε τη βάση $\mathcal{B}'$ να είναι ορθοκανονική.

Υποθέτουμε οτι $\dim V = n \geq 2$. Ο τελεστής L έχει μία ιδιοτιμή λ_1. Έστω u_1 ένα ιδιοδιάνυσμα με $||u_1|| = 1$. Συμπληρώνουμε σε ορθοκανονική βάση του V, $\mathcal{B} = \{u_1, u_2, \ldots, u_n\}$ και θεωρούμε τον πίνακα $A = {}_{\mathcal{B}}L_{\mathcal{B}}$. Η πρώτη στήλη του A είναι το διάνυσμα συντεταγμένων του $L(u_1) = \lambda_1 u_1$, και συνεπώς ο A έχει τη μορφή

$$A = \begin{bmatrix} \lambda_1 & a_{12} & \ldots & a_{1n} \\ 0 & & & \\ \vdots & & D & \\ 0 & & & \end{bmatrix} .$$

Θεωρούμε το V ως ευθύ άθροισμα των $V_1 = \langle u_1\rangle$ και $U = \langle u_2, \ldots, u_n\rangle$ και τις κανονικές απεικονίσεις $j_2 : U \to V_1 \oplus U$ και $p_2 : V_1 \oplus U \to U$, σελ. 128. Ορίζουμε την απεικόνιση

$$M = p_2 \circ L \circ j_2 : U \to U$$

και παρατηρούμε οτι ο πίνακας που παριστάνει την απεικόνιση M ως προς τη βάση $\{u_2, \ldots, u_n\}$ είναι ο D.

Από την επαγωγική υπόθεση, υπάρχει βάση $\mathcal{W} = \{w_2, \ldots, w_n\}$, ως προς την οποία ο πίνακας T της απεικόνισης M είναι άνω τριγωνικός. Παρατηρούμε οτι ο πίνακας του L ως προς τη βάση $\mathcal{B}' = \{u_1, w_2, \ldots, w_n\}$ έχει τη μορφή

$$B = \begin{bmatrix} \lambda_1 & b_{12} & \ldots & b_{1n} \\ 0 & & & \\ \vdots & & T & \\ 0 & & & \end{bmatrix},$$

δηλαδή είναι άνω τριγωνικός.

□

Πόρισμα 5.15 *Εάν V είναι διανυσματικός χώρος πάνω από το $\mathbb{R}$ και το χαρακτηριστικό πολυώνυμο χ_L του τελεστή $L : V \longrightarrow V$ αναλύεται σε παράγοντες πρώτου βαθμού πάνω από το $\mathbb{R}$, τότε υπάρχει ορθοκανονική βάση ως προς την οποία ο L έχει άνω τριγωνικό πίνακα.*

Απόδειξη. Αφού $\chi_L(x) = (-1)^n(x-\lambda_1)\cdots(x-\lambda_n)$, οι ιδιοτιμές του L είναι οι πραγματικοί αριθμοί $\lambda_1, \ldots, \lambda_n$, και υπάρχει τουλάχιστον ένα ιδιοδιάνυσμα $u_1 \in V$, με $||u_1|| = 1$, έστω για την ιδιοτιμή λ_1.

Για να εφαρμόσουμε την επαγωγή όπως στο Θεώρημα 5.14, αρκεί να παρατηρήσουμε οτι $\chi_L(x) = -(x-\lambda_1)\chi_M(x)$, και συνεπώς χ_M επίσης αναλύεται σε παραγόντες πρώτου βαθμού.

□

Παράδειγμα 5.22 θεωρούμε τον πίνακα

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

Το χαρακτηριστικό πολυώνυμο του A είναι $\chi_A(x) = -(x-1)^2(x-2)$, το οποίο είναι γινόμενο παραγόντων πρώτου βαθμού. Θα βρούμε ορθογώνιο πίνακα ο οποίος τριγωνοποιεί τον A. Οι ιδιοτιμές του A είναι $\lambda_1 = 1$, με αλγεβρική πολλαπλότητα 2 και ιδιοδιάνυσμα $v_1 = (1, -1, 1)$, και $\lambda_2 = 2$ με αλγεβρική πολλαπλότητα 1 και ιδιοδιάνυσμα $v_2 = (2, -1, 3)$. Σύμφωνα με την απόδειξη του Λήμματος Schur επιλέγουμε ένα ιδιοδιάνυσμα και βρίσκουμε το ορθογώνιο συμπλήρωμα του χώρου που παράγεται από αυτό. Εάν επιλέξουμε το v_1, το ορθογώνιο συμπλήρωμα είναι ο χώρος $W = \left\langle \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix} \right\rangle^{\perp} = \langle w_1, w_2 \rangle$, με

$$w_1 = \begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}, \quad w_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}.$$

Ορίζουμε τον τελεστή $M : W \longrightarrow W$, $M(w) = P \circ L(w)$, όπου P είναι η ορθογώνια προβολή του $\mathbb{R}^3$ στο W. Για να υπολογίσουμε τα $M(w_1)$, $M(w_2)$, παρατηρούμε οτι $L(w_1) = (0, -2, 1) = v_1 - w_2$ και $L(w_2) = (2, 2, 3) = v_1 + 2w_1 + 3w_2$, άρα $M(w_1) = -w_2$ και $M(w_2) = 2w_1 + 3w_2$. Δηλαδή ο πίνακας του M ως προς τη βάση $\{w_1, w_2\}$ είναι ο $C = \begin{bmatrix} 0 & 2 \\ -1 & 3 \end{bmatrix}$. Ο C έχει ιδιοτιμές $\mu_1 = 2$ με ιδιοδιάνυσμα $(1, 1)$, και $\mu_2 = 1$ με ιδιοδιάνυσμα $(2, 1)$. Δηλαδή τα ιδιοδιανύσματα του τελεστή M έχουν διάνυσμα συντεταγμένων ως προς τη βάση $\{w_1, w_2\}$ τα $(1, 1)$ και $(2, 1)$. Χρησιμοποιώντας το διάνυσμα συντεταγμένων $(1, 1)$ βρίσκουμε οτι ένα ιδιοδιάνυσμα του τελεστή M είναι το $w' = w_1 + w_2 = (0, 1, 1)$. Αυτό επιλέγουμε ως δεύτερο διάνυσμα της ορθογώνιας βάσης. Ως τρίτο διάνυσμα μπορούμε να επιλέξουμε οποιοδήποτε διάνυσμα ορθογώνιο στα v_1 και w'. Άρα έχουμε τη βάση

$$\mathcal{B} = \left\{ \begin{bmatrix} 1 \\ -1 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} -2 \\ -1 \\ 1 \end{bmatrix} \right\}$$

και τον ορθογώνιο πίνακα

$$Q = \begin{bmatrix} \frac{1}{\sqrt{3}} & 0 & \frac{-2}{\sqrt{6}} \\ \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{-1}{\sqrt{6}} \\ \frac{1}{\sqrt{3}} & \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

για τον οποίο $Q^T A Q$ είναι άνω τριγωνικός.

5.10 Διαγωνιοποίηση ερμιτιανών τελεστών.

Θεώρημα 5.16 (Φασματικό Θεώρημα) *Κάθε ερμιτιανός τελεστής σε ένα διανυσματικό χώρο πεπερασμένης διάστασης με εσωτερικό γινόμενο έχει μία βάση από ορθογώνια ιδιοδιανύσματα. Ο πίνακας του τελεστή ως προς αυτή τη βάση είναι διαγώνιος, με τις (πραγματικές) ιδιοτιμές στη διαγώνιο.*

Απόδειξη. Αφού ο τελεστής L είναι ερμιτιανός, οι ιδιοτιμές του είναι πραγματικοί αριθμοί, και συνεπώς το χαρακτηριστικό πολυώνυμο είναι γινόμενο παραγόντων πρώτου βαθμού και στην περίπτωση που το σώμα είναι οι πραγματικοί αριθμοί.

Από το Λήμμα του Schur, υπάρχει ορθοκανονική βάση ως προς την οποία ο πίνακας A του L είναι άνω τριγωνικός. Τότε ο πίνακας A^* είναι κάτω τριγωνικός. Αλλά αφού ο L είναι ερμιτιανός, $A^* = A$, και συνεπώς ο πίνακας A είναι διαγώνιος. Τότε τα διαγώνια στοιχεία είναι οι ιδιοτιμές του τελεστή, ενώ η βάση αποτελείται από ιδιοδιανύσματα του τελεστή.

□

Παράδειγμα 5.23 Θεωρούμε τον πίνακα $A = \begin{bmatrix} 1 & 2+\mathrm{i} \\ 2-\mathrm{i} & 5 \end{bmatrix}$. Ελέγξτε οτι ο A είναι ερμιτιανός, $A^* = A$. Οι ιδιοτιμές του A είναι 0 και 6, με αντίστοιχα ιδιοδιανύσματα $(2+\mathrm{i}, -1)$

και $(2+\mathrm{i},\, 5)$. Διαιρούμε με τις νόρμες $\|(2+\mathrm{i},\, 1)\| = \sqrt{6}$ και $\|(2+\mathrm{i},\, 5)\| = \sqrt{30}$ και έχουμε τα ορθοκανονικά ιδιοδιανύσματα $u_1 = \frac{1}{\sqrt{6}}(2+\mathrm{i},\, -1)$ και $u_2 = \frac{1}{\sqrt{30}}(2+\mathrm{i},\, 5)$. Ο μοναδιαίος πίνακας

$$U = \begin{bmatrix} \frac{2+\mathrm{i}}{\sqrt{6}} & \frac{2+\mathrm{i}}{\sqrt{30}} \\ \frac{-1}{\sqrt{6}} & \frac{5}{\sqrt{30}} \end{bmatrix}$$

διαγωνιοποιεί τον A,

$$U^*AU = \begin{bmatrix} 0 & 0 \\ 0 & 6 \end{bmatrix}.$$

Πρόταση 5.17 *Για κάθε ερμιτιανό πίνακα $A \in \mathcal{M}(n, \mathbb{C})$ υπάρχει μοναδιαίος πίνακας U τέτοιος ώστε $\Lambda = U^*AU$ είναι πραγματικός διαγώνιος πίνακας.*
Για κάθε συμμετρικό πίνακα $A \in \mathcal{M}(n, \mathbb{R})$ υπάρχει ορθογώνιος πίνακας Q τέτοιος ώστε $\Lambda = Q^TAQ$ είναι πραγματικός διαγώνιος πίνακας.

Θεώρημα 5.18 (Θεώρημα Φασματικής Ανάλυσης.) *Κάθε ερμιτιανός πίνακας $A \in \mathcal{M}(n, \mathbb{C})$ με k διαφορετικές ιδιοτιμές εκφράζεται ως άθροισμα*

$$A = \lambda_1 P_1 + \cdots + \lambda_k P_k\,,$$

όπου λ_i, για $i = 1, \ldots, k$, είναι οι ιδιοτιμές και P_i είναι ο πίνακας ορθογώνιας προβολής στον ιδιόχωρο της ιδιοτιμής λ_i.

Απόδειξη. Υπενθυμίζουμε οτι ένας τρόπος να περιγράψουμε το γινόμενο δύο πινάκων, AB είναι ως άθροισμα πινάκων που προκύπτουν από το γινόμενο της i-στήλης του A με την i-γραμμή του B. Συγκεκριμένα,

$$\begin{bmatrix} a_{11} & \ldots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \ldots & a_{mn} \end{bmatrix} \begin{bmatrix} b_{11} & \ldots & b_{1k} \\ \vdots & \ddots & \vdots \\ b_{n1} & \ldots & b_{nk} \end{bmatrix} = \sum_{i=1}^{n} \begin{bmatrix} a_{1i} \\ \vdots \\ a_{mi} \end{bmatrix} \begin{bmatrix} b_{i1} & \ldots & b_{ik} \end{bmatrix}.$$

Αναλύουμε με αυτό τον τρόπο το γινόμενο $A = U(\Lambda U^*)$.

$$\begin{bmatrix} u_{11} & \ldots & u_{1n} \\ \vdots & \ddots & \vdots \\ u_{n1} & \ldots & u_{nn} \end{bmatrix} \begin{bmatrix} \lambda_1 \bar{u}_{11} & \ldots & \lambda_1 \bar{u}_{n1} \\ \vdots & \ddots & \vdots \\ \lambda_n \bar{u}_{1n} & \ldots & \lambda_n \bar{u}_{nn} \end{bmatrix} =$$

$$\lambda_1 \begin{bmatrix} u_{11} \\ \vdots \\ u_{n1} \end{bmatrix} \begin{bmatrix} \bar{u}_{11} & \ldots & \bar{u}_{n1} \end{bmatrix} + \cdots + \lambda_n \begin{bmatrix} u_{1n} \\ \vdots \\ u_{nn} \end{bmatrix} \begin{bmatrix} \bar{u}_{1n} & \ldots & \bar{u}_{nn} \end{bmatrix}.$$

Παρατηρούμε οτι

$$\begin{bmatrix} u_{1i} \\ \vdots \\ u_{ni} \end{bmatrix} \begin{bmatrix} \bar{u}_{1i} & \ldots & \bar{u}_{ni} \end{bmatrix}$$

είναι ο πίνακας ορθογώνιας προβολής στον υπόχωρο που παράγεται από το ιδιοδιάνυσμα $(u_{1i}, \ldots, u_{ni})$.

Εάν η ιδιοτιμή λ_j έχει πολλαπλότητα k και ορθοκανονικά ιδιοδιανύσματα $w_1, \ldots, w_k$, τότε ο πίνακας ορθογώνιας προβολής στον ιδιόχωρο της λ_j είναι το άθροισμα των ορθογωνίων προβολών σε κάθε ένα από τα $w_1, \ldots, w_k$,

$$P_j = w_1 w_1^* + \cdots + w_k w_k^* .$$

□

Παράδειγμα 5.24 Θεωρούμε τον πίνακα A του Παραδείγματος 5.23. Ο A έχει ιδιοτιμές 0 και 6, με αντίστοιχα ορθοκανονικά ιδιοδιανύσματα $u_1 = \frac{1}{\sqrt{6}}(2+\mathrm{i}, -1)$ και $u_2 = \frac{1}{\sqrt{30}}(2+\mathrm{i}, 5)$. Οι πίνακες προβολής στους ιδιόχωρους είναι $u_1 u_1^*$ και $u_2 u_2^*$. Η φασματική ανάλυση του πίνακα A είναι

$$A = 0 u_1 u_1^* + 6 u_2 u_2^* = 6 \begin{bmatrix} \frac{2+\mathrm{i}}{\sqrt{30}} \\ \frac{5}{\sqrt{30}} \end{bmatrix} \begin{bmatrix} \frac{2-\mathrm{i}}{\sqrt{30}} & \frac{5}{\sqrt{30}} \end{bmatrix} .$$

5.11 Κανονικοί τελεστές

Οι ερμιτιανοί δεν είναι οι μόνοι τελεστές που διαγωνιοποιούνται από ορθοκανονικά ιδιοδιανύσματα. Για τελεστές σε μιγαδικούς διανυσματικούς χώρους, μπορούμε να διατυπώσουμε ένα απλό κριτήριο που χαρακτηρίζει τους ορθογώνια διαγωνιοποιήσιμους τελεστές.

Ορισμός 5.13. Θεωρούμε διανυσματικό χώρο V με εσωτερικό γινόμενο. Ένας τελεστής $L : V \longrightarrow V$ λέγεται **κανονικός** εάν

$$L \circ L^* = L^* \circ L\,.$$

Εάν A είναι ο πίνακας του τελεστή L ως προς μία ορθοκανονική βάση του V, τότε ο πίνακας του τελεστή L^* είναι ο συζυγής πίνακας A^*, και ο τελεστής L είναι κανονικός εάν και μόνον εάν $AA^* = A^*A$.

Παράδειγμα 5.25 Θεωρούμε τον πίνακα $A = \begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}$ και τον τελεστή $T_A : \mathbb{C}^2 \longrightarrow \mathbb{C}^2$. Τότε $A^* = \begin{bmatrix} 2 & 3 \\ -3 & 2 \end{bmatrix}$ και

$$AA^* = \begin{bmatrix} 13 & 0 \\ 0 & 13 \end{bmatrix} = A^*A\,,$$

άρα T_A είναι κανονικός τελεστής στο $\mathbb{C}^2$ με το κανονικό εσωτερικό γινόμενο. Οι ιδιοτιμές του T_A είναι $2+3\mathrm{i}$, $2-3\mathrm{i}$, και τα ιδιοδιανύσματα $\frac{1}{\sqrt{2}}(\mathrm{i},\, 1)$, $\frac{1}{\sqrt{2}}(-\mathrm{i},\, 1)$ αποτελούν ορθοκανονική βάση του $\mathbb{C}^2$. Ως προς αυτή τη βάση, ο πίνακας του T_A είναι ο

$$\frac{1}{2}\begin{bmatrix} -\mathrm{i} & 1 \\ \mathrm{i} & 1 \end{bmatrix}\begin{bmatrix} 2 & -3 \\ 3 & 2 \end{bmatrix}\begin{bmatrix} \mathrm{i} & -\mathrm{i} \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 2+3\mathrm{i} & 0 \\ 0 & 2-3\mathrm{i} \end{bmatrix}.$$

Λήμμα 5.19 *Εάν V είναι διανυσματικός χώρος με εσωτερικό γινόμενο, L είναι κανονικός τελεστής στον V και v είναι ιδιοδιάνυσμα για την ιδιοτιμή λ, τότε v είναι επίσης ιδιοδιάνυσμα του τελεστή L^* για την ιδιοτιμή $\bar{\lambda}$.*

Απόδειξη. Θεωρούμε τον τελεστή $L - \lambda\mathbf{I}$, και έχουμε

$$\begin{aligned}(L-\lambda\mathbf{I})\circ(L-\lambda\mathbf{I})^* &= (L-\lambda\mathbf{I})\circ(L^*-\bar{\lambda}\mathbf{I}) = L\circ L^* - \lambda L^* - \bar{\lambda}L + |\lambda|^2\mathbf{I}\,,\\ (L-\lambda\mathbf{I})^*\circ(L-\lambda\mathbf{I}) &= (L^*-\bar{\lambda}\mathbf{I})\circ(L-\lambda\mathbf{I}) = L^*\circ L - \lambda L^* - \bar{\lambda}L + |\lambda|^2\mathbf{I}\,,\end{aligned}$$

και αφού L είναι κανονικός, $L - \lambda\mathbf{I}$ είναι επίσης κανονικός.

Για το ιδιοδιάνυσμα v ισχύει $(L - \lambda\mathbf{I})(v) = 0$. Άρα

$$\begin{aligned}0 = \langle (L-\lambda\mathbf{I})(v),\, (L-\lambda\mathbf{I})(v)\rangle &= \langle v,\, (L-\lambda\mathbf{I})^*\circ(L-\lambda\mathbf{I})(v)\rangle\\ &= \langle v,\, (L-\lambda\mathbf{I})\circ(L-\lambda\mathbf{I})^*(v)\rangle\\ &= \langle (L-\lambda\mathbf{I})^*(v),\, (L-\lambda\mathbf{I})^*(v)\rangle\,,\end{aligned}$$

δηλαδή $(L^* - \bar{\lambda}\mathbf{I})(v) = 0$ και v είναι ιδιοδιάνυσμα του L^* για την ιδιοτιμή $\bar{\lambda}$.

□

Λήμμα 5.20 *Τα ιδιοδιανύσματα ενός κανονικού τελεστή για διαφορετικές ιδιοτιμές είναι ορθογώνια.*

Απόδειξη. Εάν λ_1, λ_2 είναι δύο διαφορετικές ιδιοτιμές του τελεστή L, και v_1, v_2 είναι αντίστοιχα ιδιοδιανύσματα,

$$\begin{aligned}\lambda_1\langle v_1, v_2\rangle &= \langle \lambda_1 v_1, v_2\rangle \\ &= \langle L(v_1), v_2\rangle \\ &= \langle v_1, L^*(v_2)\rangle \\ &= \langle v_1, \bar{\lambda}_2 v_2\rangle = \lambda_2\langle v_1, v_2\rangle\,.\end{aligned}$$

Αφού $\lambda_1 \neq \lambda_2$, συμπεραίνουμε οτι $\langle v_1, v_2\rangle = 0$.

□

Θεώρημα 5.21 *Θεωρούμε διανυσματικό χώρο V πεπερασμένης διάστασης $\dim V = n$ πάνω από το $\mathbb{C}$, με εσωτερικό γινόμενο, και τελεστή $L : V \longrightarrow V$. Ο τελεστής L είναι κανονικός εάν και μόνον εάν ο L είναι μοναδιαία διαγωνιοποιήσιμος, δηλαδή υπάρχει ορθοκανονική βάση του V που αποτελείται από ιδιοδιανύσματα του L. Ως προς αυτή τη βάση, ο πίνακας του L είναι διαγώνιος.*

Απόδειξη. Υποθέτουμε οτι υπάρχει ορθοκανονική βάση του V από ιδιοδιανύσματα του L. Ο πίνακας του L ως προς αυτή τη βάση είναι

$$\Lambda = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix},$$

και ο πίνακας του συζυγούς τελεστή L^* ως προς την ίδια βάση είναι

$$\Lambda^* = \bar{\Lambda} = \begin{bmatrix} \bar{\lambda}_1 & & 0 \\ & \ddots & \\ 0 & & \bar{\lambda}_n \end{bmatrix}.$$

Αφού $\Lambda\Lambda^* = \Lambda^*\Lambda$, έπεται οτι $L \circ L^* = L^* \circ L$ και ο L είναι κανονικός.

Αντίστροφα, υποθέτουμε οτι ο τελεστής L είναι κανονικός. Ο L έχει μία ιδιοτιμή $\lambda_1 \in \mathbb{C}$, με ιδιοδιάνυσμα v_1. Θεωρούμε τον υπόχωρο $W_1 = \langle v_1\rangle^\perp$ των διανυσμάτων που είναι ορθογώνια προς το v_1, και θέτουμε $V_1 = W_1$. Εάν $w \in V_1$, τότε

$$\langle v_1, L(w)\rangle = \langle L^*(v_1), w\rangle = \bar{\lambda}_1\langle v_1, w\rangle = 0\,.$$

Άρα $L(w) \in V_1$, και V_1 είναι αναλλοίωτος υπόχωρος του L. Άρα υπάρχει ένα ιδιοδιάνυσμα v_2 του L που ανήκει στον υπόχωρο V_1. Με τον ίδιο τρόπο δείχνουμε οτι ο υπόχωρος $W_2 = \langle v_2\rangle^\perp$

είναι αναλλοίωτος από τον L. Θέτουμε $V_2 = V_1 \cap W_2$, και παρατηρούμε οτι V_2 είναι υπόχωρος του V_1 και είναι αναλλοίωτος από τον L.

Υποθέτουμε οτι για $k < n$ έχουμε βρεί ιδιοδιανύσματα $v_1, v_2, \ldots, v_k$ του L τέτοια ώστε $\langle v_i, v_j \rangle = 0$ για $i \neq j$, και $V_k = \langle v_1 \rangle^{\perp} \cap \cdots \langle v_k \rangle^{\perp}$ είναι υπόχωρος του V αναλλοίωτος από τον L. Τότε υπάρχει ιδιοδιάνυσμα v_{k+1} του L στον V_k, και ο υπόχωρος $V_{k+1} = V_k \cap \langle v_{k+1} \rangle^{\perp}$ είναι αναλλοίωτος από τον L.

Αφού $\dim V < \infty$, καταλήγουμε σε βάση $\{v_1, \ldots, v_n\}$ του V, που αποτελείται από ιδιοδιανύσματα του L. Θέτουμε $q_i = \frac{1}{\langle v_i, v_i \rangle^{1/2}} v_i$, και $\{q_1, \ldots, q_n\}$ είναι ορθοκανονική βάση του V από ιδιοδιανύσματα του L.

$\square$