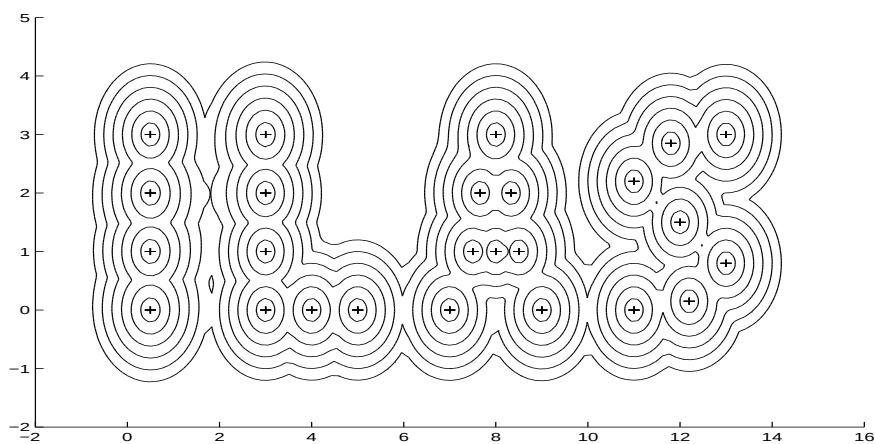




ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ & ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ

ΘΕΜΑΤΑ ΑΝΑΛΥΣΗΣ ΠΙΝΑΚΩΝ

Παναγιώτης Ι. Ψαρράκος
Καθηγητής Ε.Μ.Π.
<http://www.math.ntua.gr/~ppsarr/>



Αθήνα, Ιούνιος 2020

Αντί Προλόγου

Ένα μέρος των σημειώσεων αυτών διανεμήθηκε (σε χειρόγραφη έκδοση) για πρώτη φορά το φθινόπωρο του 2006, στα πλαίσια της διδασκαλίας του μεταπτυχιακού μαθήματος *Ανάλυση Πινάκων* του Τομέα Μαθηματικών της Σχολής Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών Ε.Μ.Π. Περιέχουν σύγχρονα θέματα Ανάλυσης Πινάκων όπως οι νόρμες πινάκων, οι παραγοντοποιήσεις LU , QR και SVD , οι κανονικοί πίνακες, το αριθμητικό πεδίο πίνακα, η θεωρία Perron-Frobenius για πίνακες με μη αρνητικά στοιχεία, οι διαταραχές ιδιοτιμών πίνακα, το ψευδοφάσμα πίνακα και οι πολυωνυμικοί πίνακες. Κάποια από τα θέματα αυτά απουσιάζουν δυστυχώς από το σύνηθες γνωστικό υπόβαθρο ενός αποφοίτου μαθηματικού τμήματος αν και βρίσκονται στο επίκεντρο της ερευνητικής δραστηριότητας στην περιοχή της Εφαρμοσμένης Γραμμικής Άλγεβρας. Ειδικότερα για την παραγοντοποίηση SVD , τη θεωρία διαταραχών ιδιοτιμών και τους πολυωνυμικούς πίνακες, μπορεί να πει κανείς με σιγουριά ότι αποτελούν ισχυρά ερευνητικά εργαλεία στην ευρύτερη περιοχή των Εφαρμοσμένων Μαθηματικών. Ταυτόχρονα, οι κανονικοί πίνακες και η θεωρία Perron-Frobenius έχουν σημαντικές εφαρμογές στη θεωρία δυναμικών συστημάτων, τη θεωρία γραφημάτων, τη δημογραφία και τα οικονομικά μαθηματικά, ενώ το αριθμητικό πεδίο πίνακα αποτελεί για δεκαετίες ένα ελκυστικό ερευνητικό θέμα. Είναι λοιπόν προφανείς οι λόγοι που οδήγησαν στην ανάγκη διδασκαλίας τόσο του προπτυχιακού όσο και του μεταπτυχιακού μαθήματος *Ανάλυσης Πινάκων*, αλλά και τα κριτήρια επιλογής της ύλης των συγκεκριμένων σημειώσεων. Ελπίζω ότι, με τον τρόπο αυτό, δίνεται η ευκαιρία σε νέους επιστήμονες να αποφύγουν (σε κάποιο βαθμό) το ρόλο του “αυτοδίδαχτου”.

Παναγιώτης Ι. Ψαρράκος

Περιεχόμενα

Συμβολισμοί	vi
1 Νόρμες Διανυσμάτων και Πινάκων	1
1.1 Νόρμες Διανυσμάτων στο $\mathbb{C}^n$	1
1.2 Ισοδυναμία Νορμών	6
1.3 Νόρμες Πινάκων στο $\mathbb{C}^{n \times n}$	11
1.4 Φασματική Ακτίνα και Νόρμες Πινάκων	16
1.5 Ασκήσεις	20
2 Παραγοντοποίηση LU	23
2.1 Εισαγωγικό Παράδειγμα	23
2.2 Στοιχειώδεις Πίνακες Gauss και Παραγοντοποίηση LU	24
2.3 Κριτήριο Παραγοντοποίησης LU και Αλγόριθμοι	26
2.4 Εφαρμογές και Σχόλια	28
2.5 Ασκήσεις	30
3 Παραγοντοποίηση QR	33
3.1 Ορθοκανονικοποίηση Gram-Schmid	33
3.2 Παραγοντοποίηση QR : Υπαρξη, Μοναδικότητα και Αλγόριθμος	34
3.3 Πίνακες Householder	35
3.4 Πλήρης Παραγοντοποίηση QR	37

3.5	Εφαρμογές και Σχόλια	39
3.6	Ασκήσεις	40
4	Παραγοντοποίηση SVD	41
4.1	Η Παραγοντοποίηση SVD	41
4.2	Άμεσες Παρατηρήσεις και Σχόλια	42
4.3	Προσεγγίσεις Χαμηλού Βαθμού (Low-Rank)	44
4.4	Ευστάθεια Τετραγωνικών Γραμμικών Συστημάτων	47
4.5	Ψευδοαντίστροφος Πίνακας	49
4.6	Ασκήσεις	51
5	Κανονικοί Πίνακες	53
5.1	Η Έννοια του Κανονικού Πίνακα	53
5.2	Αποστάσεις από την Κανονικότητα	59
5.3	Παραδείγματα	62
5.4	Ασκήσεις	64
6	Αριθμητικό Πεδίο Πίνακα	65
6.1	Ορισμός του Αριθμητικού Πεδίου	65
6.2	Βασικές Ιδιότητες του Αριθμητικού Πεδίου	66
6.3	Κυρτότητα	74
6.4	Αριθμητική Ακτίνα	77
6.5	Προσέγγιση του Αριθμητικού Πεδίου	79
6.6	Σύνορο του Αριθμητικού Πεδίου	81
6.7	Πίνακες Jordan και Ομοιότητα	83
6.8	Ασκήσεις	86
7	Θεωρία Perron-Frobenius	89
7.1	Εισαγωγικό Παράδειγμα	89
7.2	Εισαγωγή στους Μη Αρνητικούς Πίνακες	91

7.3	Θετικοί Πίνακες	95
7.4	Μη Αρνητικοί Πίνακες	101
7.5	Μη Υποβιβάσιμοι, μη Αρνητικοί Πίνακες	103
7.6	Ένα Γενικό Οριακό Θεώρημα	108
7.7	Ασκήσεις	110
8	Διαταραχές Ιδιοτιμών και Ψευδοφάσμα Πίνακα	111
8.1	Διαταραχές Ιδιοτιμών	111
8.2	Ψευδοφάσμα Πίνακα	117
8.3	Ψευδοφάσμα Κατά τη 2-Νόρμα	123
8.4	Παραδείγματα	125
8.5	Ασκήσεις	127
9	Πολυωνυμικοί Πίνακες	129
9.1	Ορισμός Πολυωνυμικού Πίνακα	129
9.2	Διαίρεση Πολυωνυμικών Πινάκων	130
9.3	Ισοδυναμίες και Κανονική Μορφή Smith	134
9.4	Standard Ζεύγη και Τριάδες	140
9.5	Αλυσίδες Jordan και Εφαρμογές	144
9.6	Ασκήσεις	151
	Βιβλιογραφία	153

Συμβολισμοί

$\mathbb{R}, \mathbb{R}^n, \mathbb{R}^{n \times m}$	Τα σύνολα των πραγματικών αριθμών, διανυσμάτων και $n \times m$ πίνακων, αντίστοιχα
$\mathbb{C}, \mathbb{C}^n, \mathbb{C}^{n \times m}$	Τα σύνολα των μιγαδικών αριθμών, διανυσμάτων και $n \times m$ πίνακων, αντίστοιχα
$\text{sign}(t)$	Το πρόσημο ενός αριθμού $t \in \mathbb{R}$
$\text{Re}(z), \text{Im}(z)$	Το πραγματικό και το φανταστικό μέρος ενός αριθμού $z \in \mathbb{C}$, αντίστοιχα
$\arg(z)$	Το όρισμα ενός αριθμού $z \in \mathbb{C}$
$\ \cdot\ _1, \ \cdot\ _2, \ \cdot\ _\infty$	Η 1-νόρμα, η 2-νόρμα και η ∞ -νόρμα διανυσμάτων, αντίστοιχα
$\ \cdot\ _1, \ \cdot\ _2, \ \cdot\ _\infty$	Η 1-νόρμα, η 2-νόρμα και η ∞ -νόρμα πινάκων, αντίστοιχα
$\ \cdot\ _F$	Η νόρμα Frobenius
$\langle \cdot, \cdot \rangle$	Το εσωτερικό γινόμενο
$\bar{A}, A^T, A^*, A^{-1}$	Ο συζυγής, ο ανάστροφος, ο αναστροφοσυζυγής και ο αντίστροφος ενός πίνακα A , αντίστοιχα
$H(A), S(A)$	Το ερμιτιανό και το αντιερμιτιανό μέρος ενός τετραγωνικού πίνακα A , αντίστοιχα
$\text{trace}(A)$	Το ίχνος ενός τετραγωνικού πίνακα A
$\det(A)$	Η ορίζουσα ενός τετραγωνικού πίνακα A
$\text{rank}(A)$	Ο βαθμός ενός πίνακα A
I_n	Ο μοναδιαίος (ταυτοτικός) πίνακας τάξης n
$\text{Range}(A)$ ή $\text{Image}(A)$	Η εικόνα ενός πίνακα A
$\text{Null}(A)$ ή $\text{Ker}(A)$	Ο πυρήνας ενός πίνακα A
$\sigma(A)$	Το φάσμα ενός τετραγωνικού πίνακα A
$\rho(A)$	Η φασματική ακτίνα ενός τετραγωνικού πίνακα A
$\kappa(A)$	Ο βαθμός (δείκτης) κατάστασης ενός τετραγωνικού πίνακα A
$F(A)$	Το αριθμητικό πεδίο ενός τετραγωνικού πίνακα A
$r(A)$	Η αριθμητική ακτίνα ενός τετραγωνικού πίνακα A
$\text{diag}\{a_1, a_2, \dots, a_k\}$	Ο διαγώνιος πίνακας με διαγώνια στοιχεία $a_1, a_2, \dots, a_k$
$\dim(V)$	Η διάσταση ενός διανυσματικού χώρου V
$V^\perp$	Το ορθογώνιο συμπλήρωμα ενός διανυσματικού υποχώρου V
$\text{span}\{x_1, x_2, \dots, x_k\}$	Η γραμμική θήκη των διανυσμάτων $x_1, x_2, \dots, x_k$
$\text{Co}(\Omega)$	Η κυρτή θήκη ενός συνόλου Ω
$\text{Int}(\Omega)$	Το εσωτερικό ενός συνόλου Ω
$\partial\Omega$	Το σύνορο ενός συνόλου Ω
$\Delta(z, \rho)$	Ο κυκλικός δίσκος με κέντρο $z \in \mathbb{C}$ και ακτίνα $\rho \geq 0$
$P(\lambda)$	Ένας πολυωνυμικός πίνακας μιγαδικής μεταβλητής λ

Κεφάλαιο 1

Νόρμες Διανυσμάτων και Πινάκων

Η έννοια της Ανάλυσης είναι συνυφασμένη με την έννοια του ορίου, ο ορισμός του οποίου απαιτεί με τη σειρά του την έννοια της απόστασης. Για το λόγο αυτό, είμαστε ουσιαστικά υποχρεωμένοι στο πρώτο κεφάλαιο, να μελετήσουμε με όποια λεπτομέρεια μας επιτρέπεται τις αποστάσεις μεταξύ διανυσμάτων στο $\mathbb{C}^n$ και πινάκων στο $\mathbb{C}^{n \times n}$.

1.1 Νόρμες Διανυσμάτων στο $\mathbb{C}^n$

Αρχικά, υπενθυμίζουμε το γνωστό ορισμό της νόρμας (του μέτρου) διανυσμάτων.

Ορισμός 1.1.1. Μία συνάρτηση $\|\cdot\| : \mathbb{C}^n \rightarrow \mathbb{R}$ ονομάζεται *νόρμα διανυσμάτων* αν για κάθε $x, y \in \mathbb{C}^n$, ικανοποιεί τα ακόλουθα:

- (i) $\|x\| \geq 0$ (μη αρνητική).
- (ii) $\|x\| = 0$ αν και μόνο αν $x = 0$.
- (iii) $\|ax\| = |a| \|x\|$ για κάθε $a \in \mathbb{C}$.
- (iv) $\|x + y\| \leq \|x\| + \|y\|$ (τριγωνική ανισότητα).

Μία συνάρτηση $\|\cdot\| : \mathbb{C}^n \rightarrow \mathbb{R}$ που ικανοποιεί τα (i), (iii) και (iv) του παραπάνω ορισμού καλείται *ημι-νόρμα διανυσμάτων*. Η ημι-νόρμα αποτελεί μία γενίκευση της έννοιας της νόρμας, η οποία επιτρέπει σε μη μηδενικά διανύσματα να έχουν μηδενικό μέτρο.

Λήμμα 1.1.1. Έστω $\|\cdot\|$ μία ημι-νόρμα διανυσμάτων στο $\mathbb{C}^n$. Τότε για κάθε ζεύγος διανυσμάτων $x, y \in \mathbb{C}^n$, ισχύει $|\|x\| - \|y\|| \leq \|x - y\|$.

Απόδειξη. Από τη σχέση $y = x + (y - x)$ και την τριγωνική ανισότητα (iv) του Ορισμού 1.1.1, έχουμε $\|y\| = \|x + (y - x)\| \leq \|x\| + \|y - x\|$. Επομένως, $\|y\| - \|x\| \leq \|y - x\|$. Ομοίως, από τη σχέση $x = y + (x - y)$ προκύπτει ότι $\|x\| - \|y\| \leq \|x - y\|$. Λόγω του (iii) του Ορισμού 1.1.1,

$$\|y - x\| = \|x - y\| \geq \pm(\|x\| - \|y\|)$$

και η απόδειξη ολοκληρώθηκε. $\square$

Ορισμός 1.1.2. Μία νόρμα $\|\cdot\|$ ονομάζεται *ορθομοναδιαία αναλλοίωτη* αν για κάθε διάνυσμα $x \in \mathbb{C}^n$ και για κάθε ορθομοναδιαίο πίνακα $U \in \mathbb{C}^{n \times n}$ (δηλαδή, $U^*U = UU^* = I_n$), ισχύει $\|Ux\| = \|x\|$.

Ορισμός 1.1.3. Μία συνάρτηση $\langle \cdot, \cdot \rangle : \mathbb{C}^n \times \mathbb{C}^n \rightarrow \mathbb{C}$ ονομάζεται *εσωτερικό γινόμενο* αν για κάθε $x, y, w \in \mathbb{C}^n$, ικανοποιεί τα ακόλουθα:

- (i) $\langle x, x \rangle \geq 0$ (μη αρνητική).
- (ii) $\langle x, x \rangle = 0$ αν και μόνο αν $x = 0$.
- (iii) $\langle x + y, w \rangle = \langle x, w \rangle + \langle y, w \rangle$ (προσθετική).
- (iv) $\langle ax, y \rangle = a \langle x, y \rangle$ για κάθε $a \in \mathbb{C}$.
- (v) $\langle x, y \rangle = \overline{\langle y, x \rangle}$.

Από τον ορισμό του εσωτερικού γινομένου, εύκολα μπορεί κανείς να επιβεβαιώσει τις παρακάτω ιδιότητες:

- (1) $\langle x, ay \rangle = \overline{\langle ay, x \rangle} = \overline{a \langle y, x \rangle} = \bar{a} \langle x, y \rangle$.
- (2) $\langle x, y + z \rangle = \overline{\langle y + z, x \rangle} = \overline{\langle y, x \rangle + \langle z, x \rangle} = \langle x, y \rangle + \langle x, z \rangle$.
- (3) $x = 0$ αν και μόνο αν $\langle x, y \rangle = 0$ για κάθε $y \in \mathbb{C}^n$.

Πόρισμα 1.1.2. Για κάθε εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$ στο $\mathbb{C}^n$, η συνάρτηση $f(x) = \sqrt{\langle x, x \rangle}$ είναι νόρμα διανυσμάτων.

Απόδειξη. Με απλές πράξεις, επαληθεύεται ότι η συνάρτηση $f(x) = \sqrt{\langle x, x \rangle}$ ικανοποιεί τον Ορισμό 1.1.1 της νόρμας διανυσμάτων. $\square$

Αν $\|\cdot\|$ είναι μία νόρμα διανυσμάτων τέτοια ώστε να ισχύει $\|x\| = \sqrt{\langle x, x \rangle}$ για κάθε $x \in \mathbb{C}^n$, για κάποιο εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$, τότε λέμε ότι η $\|\cdot\|$

επάγεται (προέρχεται) από το εσωτερικό γινόμενο $\langle \cdot, \cdot \rangle$. Στην περίπτωση αυτή, για κάθε $x, y \in \mathbb{C}^n$, ισχύει

$$\begin{aligned} \frac{1}{2} (\|x + y\|^2 + \|x - y\|^2) &= \frac{1}{2} (\langle x + y, x + y \rangle + \langle x - y, x - y \rangle) \\ &= \frac{1}{2} (2 \langle x, x \rangle + 2 \langle y, y \rangle). \end{aligned}$$

Δηλαδή, η νόρμα $\|\cdot\|$ ικανοποιεί τον κανόνα του παραλληλογράμμου:

$$\frac{1}{2} (\|x + y\|^2 + \|x - y\|^2) = \|x\|^2 + \|y\|^2, \quad \forall x, y \in \mathbb{C}^n. \quad (1.1.1)$$

Θεώρημα 1.1.3. Ο κανόνας του παραλληλογράμμου (1.1.1) αποτελεί ικανή και αναγκαία συνθήκη ώστε μία νόρμα $\|\cdot\|$ να επάγεται από ένα εσωτερικό γινόμενο.

Απόδειξη. Η αναγκαιότητα του κανόνα του παραλληλογράμμου ώστε μία νόρμα να επάγεται από εσωτερικό γινόμενο έχει ήδη αποδειχθεί παραπάνω. Αντίστροφα, αν μία νόρμα $\|\cdot\|$ ικανοποιεί τον κανόνα του παραλληλογράμμου, τότε αποδεικνύεται ότι η συνάρτηση

$$\langle x, y \rangle = \frac{1}{4} (\|x + y\|^2 - \|x - y\|^2 - i\|x + iy\|^2 + i\|x - iy\|^2)$$

είναι εσωτερικό γινόμενο από το οποίο επάγεται η νόρμα (η απόδειξη βασίζεται σε τεχνικούς υπολογισμούς και παραλείπεται). $\square$

Μία σημαντική ιδιότητα του εσωτερικού γινομένου είναι η ανισότητα Cauchy-Schwarz, γνωστή και ως ανισότητα Bunyakovsky.

Θεώρημα 1.1.4. (Cauchy-Schwarz) Έστω $\langle \cdot, \cdot \rangle$ ένα εσωτερικό γινόμενο στο $\mathbb{C}^n$ και $\|\cdot\|$ η επαγόμενη νόρμα. Τότε για κάθε $x, y \in \mathbb{C}^n$, ισχύει

$$|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle,$$

ή ισοδύναμα,

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

Η ισότητα ισχύει αν και μόνο αν τα διανύσματα x και y είναι γραμμικά εξαρτημένα.

Απόδειξη. Στην περίπτωση που $y = 0$, η ανισότητα του θεωρήματος ισχύει κατά τετριμμένο τρόπο. Δεχόμαστε λοιπόν ότι $y \neq 0$ και θεωρούμε το πραγματικό δευτεροβάθμιο πολυώνυμο, ως προς $t \in \mathbb{R}$,

$$\begin{aligned} p(t) &= \langle x + ty, x + ty \rangle \\ &= \langle x, x \rangle + t \langle y, x \rangle + t \langle x, y \rangle + t^2 \langle y, y \rangle \\ &= \langle y, y \rangle t^2 + 2 \operatorname{Re}(\langle x, y \rangle) t + \langle x, x \rangle. \end{aligned}$$

Το $p(t)$ λαμβάνει αποκλειστικά μη αρνητικές τιμές και κατά συνέπεια, η διαχρίνουσα του είναι

$$4 \operatorname{Re}(\langle x, y \rangle)^2 - 4 \langle y, y \rangle \langle x, x \rangle \leq 0,$$

ή ισοδύναμα,

$$\operatorname{Re}(\langle x, y \rangle)^2 \leq \langle y, y \rangle \langle x, x \rangle.$$

Η ανισότητα αυτή ισχύει για κάθε ζεύγος $x, y \in \mathbb{C}^n$ και αντικαθιστώντας το y με το διάνυσμα $\langle x, y \rangle y$, προκύπτει ότι

$$\operatorname{Re}(\langle x, \langle x, y \rangle y \rangle)^2 \leq \langle x, x \rangle \langle y, y \rangle |\langle x, y \rangle|^2. \quad (1.1.2)$$

Όμως,

$$\begin{aligned} \operatorname{Re}(\langle x, \langle x, y \rangle y \rangle) &= \operatorname{Re}(\overline{\langle x, \langle x, y \rangle y \rangle}) = \operatorname{Re}(\langle x, y \rangle \overline{\langle x, y \rangle}) \\ &= \operatorname{Re}(|\langle x, y \rangle|^2) = |\langle x, y \rangle|^2. \end{aligned}$$

Επομένως, η ανισότητα (1.1.2) γράφεται

$$|\langle x, y \rangle|^4 \leq \langle x, x \rangle \langle y, y \rangle |\langle x, y \rangle|^2.$$

Αν $\langle x, y \rangle = 0$, τότε η ανισότητα του θεωρήματος ισχύει κατά τετριμμένο τρόπο. Αν $\langle x, y \rangle \neq 0$, τότε προφανώς $|\langle x, y \rangle|^2 \leq \langle x, x \rangle \langle y, y \rangle$.

Από το (ii) του Ορισμού 1.1.3, προκύπτει ότι το $p(t)$ έχει πραγματική (διπλή) ρίζα αν και μόνο αν $x + ty = 0$ για κάποιο t . Δηλαδή, $|\langle x, y \rangle|^2 = \langle x, x \rangle \langle y, y \rangle$ αν και μόνο αν τα διανύσματα x και y είναι παράλληλα. $\square$

Χαρακτηριστικά παραδείγματα νορμών διανυσμάτων στο $\mathbb{C}^n$ (οι οποίες χρησιμοποιούνται συχνά) είναι τα ακόλουθα:

- Η l_p -νόρμα (ή p -νόρμα), για οποιονδήποτε πραγματικό αριθμό $p \geq 1$, ορίζεται ως

$$\|x\|_p = \left\| [x_1 \ x_2 \ \cdots \ x_n]^T \right\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_n|^p)^{1/p}.$$

Κάθε p -νόρμα ($1 \leq p < +\infty$) ικανοποιεί τις προϋποθέσεις του Ορισμού 1.1.1. Ειδικότερα, η τριγωνική ανισότητα

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

είναι γνωστή και ως ανισότητα Minkowski. Επιπλέον, για τις p -νόρμες, ισχύει η ανισότητα Hölder (γενίκευση της ανισότητας Cauchy-Schwarz),

$$|\langle x, y \rangle| \leq \|x\|_p \|y\|_q, \quad \forall p, q > 1 \text{ με } \frac{1}{p} + \frac{1}{q} = 1.$$

- Η Ευκλείδεια νόρμα (ή l_2 -νόρμα, ή 2-νόρμα) ορίζεται ως

$$\|x\|_2 = \left\| [x_1 \ x_2 \ \cdots \ x_\nu]^T \right\|_2 = (|x_1|^2 + |x_2|^2 + \cdots + |x_\nu|^2)^{1/2}.$$

Αποτελεί μία ειδική περίπτωση της p -νόρμας (για $p = 2$), είναι ίσως η πιο γνωστή νόρμα διανυσμάτων, και επάγεται από το Ευκλείδιο εσωτερικό γινόμενο

$$\langle x, y \rangle = y^* x = [\bar{y}_1 \ \bar{y}_2 \ \cdots \ \bar{y}_\nu] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_\nu \end{bmatrix} = \bar{y}_1 x_1 + \bar{y}_2 x_2 + \cdots + \bar{y}_\nu x_\nu.$$

Επίσης, είναι ορθομοναδιαία αναλλοίωτη καθώς για κάθε διάνυσμα $x \in \mathbb{C}^\nu$ και για κάθε ορθομοναδιαίο πίνακα $U \in \mathbb{C}^{\nu \times \nu}$, ισχύει

$$\|Ux\|_2^2 = (Ux)^* Ux = x^* U^* Ux = x^* x = \|x\|_2^2.$$

- Η αθροιστική νόρμα (ή l_1 -νόρμα, ή 1-νόρμα) ορίζεται ως

$$\|x\|_1 = \left\| [x_1 \ x_2 \ \cdots \ x_\nu]^T \right\|_1 = |x_1| + |x_2| + \cdots + |x_\nu|$$

και είναι γνωστή και ως *νόρμα του Μανχάταν*, καθώς συνδέεται με την απόσταση που διανύει ένα αυτοκίνητο κινούμενο πάνω στο ορθογώνιο πλέγμα δρόμων του Μανχάταν.

- Η μεγιστική νόρμα (ή $\max$ -νόρμα, ή ∞ -νόρμα) ορίζεται ως

$$\|x\|_\infty = \left\| [x_1 \ x_2 \ \cdots \ x_\nu]^T \right\|_\infty = \max \{|x_1|, |x_2|, \dots, |x_\nu|\}$$

και είναι η πιο εύκολα υπολογίσιμη νόρμα (αναφερόμενοι πάντα στο $\mathbb{C}^\nu$).

Παρατήρηση 1.1.1. Η ∞ -νόρμα και η p -νόρμα συνδέονται με τη σχέση

$$\|x\|_\infty = \lim_{p \rightarrow +\infty} \|x\|_p, \quad \forall x \in \mathbb{C}^\nu.$$

Πράγματι, αν για ένα διάνυσμα $x = [x_1 \ x_2 \ \cdots \ x_\nu]^T \in \mathbb{C}^\nu$, θεωρήσουμε την p -νόρμα

$$\|x\|_p = (|x_1|^p + |x_2|^p + \cdots + |x_\nu|^p)^{1/p}$$

και την ∞ -νόρμα

$$\|x\|_\infty = \max \{|x_1|, |x_2|, \dots, |x_\nu|\},$$

τότε

$$\frac{\|x\|_p}{\|x\|_\infty} = \left(\left| \frac{x_1}{\|x\|_\infty} \right|^p + \left| \frac{x_2}{\|x\|_\infty} \right|^p + \cdots + \left| \frac{x_\nu}{\|x\|_\infty} \right|^p \right)^{1/p} \in [1, \nu^{1/p}].$$

Επειδή $\lim_{p \rightarrow +\infty} \nu^{1/p} = \lim_{p \rightarrow +\infty} 1 = 1$, από το κριτήριο παρεμβολής, έχουμε

$$\lim_{p \rightarrow +\infty} \frac{\|x\|_p}{\|x\|_\infty} = 1.$$

Παρατήρηση 1.1.2. Πέρα των νορμών που ορίζονται στο $\mathbb{C}^n$, μπορούμε να ορίσουμε νόρμες και σε διανυσματικούς χώρους άπειρης διάστασης. Αν, για παράδειγμα θεωρήσουμε το διανυσματικό χώρο $C[a, b]$ των συνεχών στο κλειστό διάστημα $[a, b]$ πραγματικών συναρτήσεων $f(t)$, τότε μπορούμε να ορίσουμε

$$\text{την } l_p\text{-νόρμα } \|f\|_p = \left(\int_a^b |f(t)|^p dt \right)^{1/p} \quad (p \geq 1),$$

$$\text{την } l_2\text{-νόρμα } \|f\|_2 = \left(\int_a^b f(t)^2 dt \right)^{1/2},$$

$$\text{την } l_1\text{-νόρμα } \|f\|_1 = \int_a^b |f(t)| dt,$$

$$\text{και την } \infty\text{-νόρμα } \|f\|_\infty = \max\{f(t) : t \in [a, b]\}.$$

1.2 Ισοδυναμία Νορμών

Είναι σαφές από τα παραδείγματα που προηγήθηκαν πως υπάρχουν πολλές διαφορετικές συναρτήσεις $\|\cdot\| : \mathbb{C}^n \rightarrow \mathbb{R}$ που ικανοποιούν τον ορισμό της νόρμας. Αυτό είναι πολύ σημαντικό, καθώς μία νόρμα μπορεί να είναι περισσότερο εύχρηστη από κάποια άλλη για ένα συγκεκριμένο πρόβλημα. Σε πλήθος εφαρμογών, εμφανίζεται το φαινόμενο η νόρμα στην οποία βασίζεται η θεωρία και η νόρμα που είναι πιο εύκολο να υπολογιστεί σε μία δεδομένη κατάσταση να μην ταυτίζονται. Επομένως, είναι σημαντικό να γνωρίζουμε τη σχέση που μπορεί να συνδέει δύο διαφορετικές νόρμες. Όμως, στο διανυσματικό χώρο $\mathbb{C}^n$ που μελετάμε (καθώς και σε όλους τους διανυσματικούς χώρους πεπερασμένης διάστασης), όλες οι νόρμες είναι “ισοδύναμες”, όπως θα διαπιστώσουμε στη συνέχεια.

Οι διανυσματικές νόρμες μπορούν να χρησιμοποιηθούν για να “μετρήσουμε” τη σύγκλιση μιας ακολουθίας διανυσμάτων, με την έννοια ότι η επιλογή της νόρμας επηρεάζει την ταχύτητα σύγκλισης. Για μία νόρμα διανυσμάτων $\|\cdot\|$, μία ακολουθία διανυσμάτων $x^{(k)} \in \mathbb{C}^n$ ($k = 1, 2, \dots$) κι ένα $x \in \mathbb{C}^n$, λέμε ότι η ακολουθία $x^{(k)}$ συγκλίνει στο x ως προς τη νόρμα $\|\cdot\|$ αν ισχύει η σχέση $\lim \|x^{(k)} - x\| = 0$. Επομένως, κατά τη μελέτη σύγκλισης μιας ακολουθίας απαιτείται (προς το παρόν) να είναι ξεκάθαρο ποια νόρμα χρησιμοποιείται. Ένα πολύ ενδιαφέρον ερώτημα που δημιουργείται είναι το αν μία ακολουθία διανυσμάτων μπορεί να συγκλίνει ως προς μία νόρμα και να μη συγκλίνει ως προς μία άλλη, δηλαδή το αν το όριο επηρεάζεται από τη νόρμα που επιλέγεται.

Για να απαντήσουμε το συγκεκριμένο ερώτημα, θα χρειαστούμε ένα γενικό λήμμα σχετικά με τις ιδιότητες συνέχειας των νορμών.

Λήμμα 1.2.1. Έστω μία νόρμα $\|\cdot\|$ διανυσμάτων στο $\mathbb{C}^n$ και n διανύσματα $x^{(1)}, x^{(2)}, \dots, x^{(n)} \in \mathbb{C}^n$. Τότε η συνάρτηση $g : \mathbb{C}^n \rightarrow \mathbb{R}$ με

$$g(z_1, z_2, \dots, z_n) = \|z_1 x^{(1)} + z_2 x^{(2)} + \dots + z_n x^{(n)}\|$$

είναι ομοιόμορφα συνεχής.

Απόδειξη. Αρκεί να αποδείξουμε ότι η συνάρτηση g ικανοποιεί τη συνθήκη Lipschitz. Θεωρούμε δύο τυχαία διανύσματα $u = u_1 x^{(1)} + u_2 x^{(2)} + \dots + u_n x^{(n)}$ και $v = v_1 x^{(1)} + v_2 x^{(2)} + \dots + v_n x^{(n)}$ του $\mathbb{C}^n$. Χρησιμοποιώντας τις βασικές ιδιότητες της νόρμας $\|\cdot\|$, βλέπουμε ότι

$$\begin{aligned} & |g(u_1, u_2, \dots, u_n) - g(v_1, v_2, \dots, v_n)| \\ &= \left| \|u\| - \|v\| \right| \leq \|u - v\| \\ &= \|(u_1 - v_1)x^{(1)} + (u_2 - v_2)x^{(2)} + \dots + (u_n - v_n)x^{(n)}\| \\ &\leq |u_1 - v_1| \|x^{(1)}\| + |u_2 - v_2| \|x^{(2)}\| + \dots + |u_n - v_n| \|x^{(n)}\|. \end{aligned}$$

Επομένως,

$$|g(u_1, u_2, \dots, u_n) - g(v_1, v_2, \dots, v_n)| \leq c \max\{|u_i - v_i| : i = 1, 2, \dots, n\},$$

όπου $c = n \max\{\|x^{(i)}\| : i = 1, 2, \dots, n\}$ η απαιτούμενη σταθερά. $\square$

Πόρισμα 1.2.2. Κάθε νόρμα $\|\cdot\|$ στο $\mathbb{C}^n$ είναι ομοιόμορφα συνεχής συνάρτηση των στοιχείων του διανύσματος.

Απόδειξη. Η απόδειξη προκύπτει άμεσα από το Λήμμα 1.2.1 αν θέσουμε $n = n$ και θεωρήσουμε ότι τα διανύσματα $x^{(1)}, x^{(2)}, \dots, x^{(n)} \in \mathbb{C}^n$ είναι τα στοιχεία της κανονικής βάσης του $\mathbb{C}^n$. $\square$

Θεώρημα 1.2.3. Έστω $\|\cdot\|_a$ και $\|\cdot\|_b$ δύο τυχαίες νόρμες στο $\mathbb{C}^n$. Τότε υπάρχουν θετικοί (πεπερασμένοι και σταθεροί) αριθμοί $\mu, M \in \mathbb{R}$ τέτοιοι ώστε

$$\mu \|x\|_a \leq \|x\|_b \leq M \|x\|_a, \quad \forall x \in \mathbb{C}^n.$$

Απόδειξη. Θεωρούμε την l_1 -νόρμα

$$\|x\|_1 = \left\| [x_1 \ x_2 \ \dots \ x_n]^T \right\|_1 = |x_1| + |x_2| + \dots + |x_n|$$

και μία τυχαία νόρμα $\|\cdot\|$ στο $\mathbb{C}^n$. Η μοναδιαία σφαίρα του $\mathbb{C}^n$ ως προς τη l_1 -νόρμα,

$$\mathcal{B}_{\|\cdot\|_1}(0, 1) = \{x \in \mathbb{C}^n : \|x\|_1 = 1\},$$

είναι ένα κλειστό και φραγμένο υποσύνολο του $\mathbb{C}^n$. Επομένως, από το Θεώρημα του Weierstrass¹ έπεται ότι η συνεχής συνάρτηση $\|\cdot\| : \mathbb{C}^n \rightarrow [0, +\infty)$ λαμβάνει μία ελάχιστη θετική τιμή μ και μία μέγιστη θετική τιμή $M < +\infty$. Δηλαδή, για κάθε διάνυσμα $y \in \mathbb{C}^n$, έχουμε

$$\left\| \frac{y}{\|y\|_1} \right\|_1 = \frac{\|y\|_1}{\|y\|_1} = 1 \quad \text{και} \quad \mu \leq \left\| \frac{y}{\|y\|_1} \right\| \leq M < +\infty.$$

Συνεπώς, για κάθε διάνυσμα $y \in \mathbb{C}^n$, ισχύει

$$\mu \|y\|_1 \leq \|y\| \leq M \|y\|_1.$$

Εφαρμόζοντας, τη σχέση αυτή για $\|\cdot\| = \|\cdot\|_a$ και για $\|\cdot\| = \|\cdot\|_b$, έχουμε

$$\mu_a \|y\|_1 \leq \|y\|_a \leq M_a \|y\|_1$$

και

$$\mu_b \|y\|_1 \leq \|y\|_b \leq M_b \|y\|_1,$$

απ' όπου καταλήγουμε στη σχέση

$$\frac{1}{M_a} \|y\|_a \leq \|y\|_1 \leq \frac{1}{\mu_b} \|y\|_b \leq \frac{M_b}{\mu_b} \|y\|_1 \leq \frac{M_b}{\mu_a \mu_b} \|y\|_a,$$

ολοκληρώνοντας την απόδειξη. □

Ειδικότερα για τις νόρμες $\|\cdot\|_1$, $\|\cdot\|_2$ και $\|\cdot\|_\infty$, έχουμε το ακόλουθο αποτέλεσμα.

Πρόταση 1.2.4. Για κάθε διάνυσμα $x \in \mathbb{C}^n$, ισχύουν οι ανισότητες

$$\|x\|_2 \leq \|x\|_1 \leq \sqrt{n} \|x\|_2,$$

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

και

$$\|x\|_\infty \leq \|x\|_2 \leq \sqrt{n} \|x\|_\infty.$$

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. □

Μια άμεση συνέπεια του Θεωρήματος 1.2.3 είναι το γεγονός ότι η σύγκλιση μιας ακολουθίας διανυσμάτων ως προς μία νόρμα σε ένα μιγαδικό χώρο πεπερασμένης διάστασης είναι *ανεξάρτητη* από τη νόρμα που χρησιμοποιείται.

¹Μία συνεχής πραγματική συνάρτηση ορισμένη σε ένα κλειστό και φραγμένο σύνολο $\Omega \subset \mathbb{C}^n$ είναι φραγμένη και επιτυγχάνει τις τιμές του ελαχίστου άνω φράγματος (supremum) και του μεγίστου κάτω φράγματος (infimum) επί του Ω .

Πόρισμα 1.2.5. Έστω $\|\cdot\|_a$ και $\|\cdot\|_b$ δύο τυχαίες νόρμες στο $\mathbb{C}^n$ και $x^{(k)} \in \mathbb{C}^n$ ($k = 1, 2, \dots$) μία ακολουθία διανυσμάτων. Τότε η ακολουθία $x^{(k)}$ συγκλίνει σε ένα διάνυσμα x ως προς τη νόρμα $\|\cdot\|_a$ αν και μόνο αν συγκλίνει στο x ως προς τη νόρμα $\|\cdot\|_b$.

Απόδειξη. Από το Πόρισμα 1.2.3, υπάρχουν θετικοί αριθμοί $M, \mu \in \mathbb{R}$ τέτοιοι ώστε

$$\mu \|x^{(k)} - x\|_a \leq \|x^{(k)} - x\|_b \leq M \|x^{(k)} - x\|_a, \quad k = 1, 2, \dots$$

και η απόδειξη είναι προφανής. $\square$

Ορισμός 1.2.1. Δύο νόρμες $\|\cdot\|_a$ και $\|\cdot\|_b$ ονομάζονται *ισοδύναμες* αν, οποιαδήποτε ακολουθία διανυσμάτων $x^{(k)} \in \mathbb{C}^n$ ($k = 1, 2, \dots$) συγκλίνει σε ένα διάνυσμα $x \in \mathbb{C}^n$ ως προς τη νόρμα $\|\cdot\|_a$ αν και μόνο αν συγκλίνει στο ίδιο διάνυσμα ως προς τη νόρμα $\|\cdot\|_b$.

Το Πόρισμα 1.2.5 μας εξασφαλίζει τα δύο επόμενα αποτελέσματα.

Πόρισμα 1.2.6. Στο διανυσματικό χώρο $\mathbb{C}^n$, όλες οι νόρμες είναι ισοδύναμες.

Πόρισμα 1.2.7. Έστω $\|\cdot\|$ μία νόρμα στο $\mathbb{C}^n$ και $x^{(k)} \in \mathbb{C}^n$ ($k = 1, 2, \dots$) μία ακολουθία διανυσμάτων. Τότε η ακολουθία $x^{(k)}$ συγκλίνει σε ένα διάνυσμα $x \in \mathbb{C}^n$ ως προς τη νόρμα $\|\cdot\|$ αν και μόνο αν συγκλίνει στο x κατά στοιχείο, δηλαδή $x_i^{(k)} \rightarrow x_i$ για κάθε $i = 1, 2, \dots, n$.

Απόδειξη. Αρκεί να εργαστούμε με την l_1 -νόρμα. $\square$

Θα λέμε λοιπόν ότι μία ακολουθία $x^{(k)} \in \mathbb{C}^n$ ($k = 1, 2, \dots$) συγκλίνει σε ένα διάνυσμα $x \in \mathbb{C}^n$ αν η ακολουθία συγκλίνει στο x ως προς μία οποιαδήποτε νόρμα.

Παρατήρηση 1.2.1. Σε ένα διανυσματικό χώρο άπειρης διάστασης, δύο νόρμες μπορεί να μην είναι ισοδύναμες. Για παράδειγμα, ας θεωρήσουμε στο διανυσματικό χώρο $C[0, 1]$ των συνεχών στο κλειστό διάστημα $[0, 1]$ πραγματικών συναρτήσεων, την ακολουθία

$$f_k(t) = \begin{cases} 0, & 0 \leq t < 1/k \\ 2(k^{3/2}t - k^{1/2}), & 1/k \leq t < 3/(2k) \\ 2(-k^{3/2}t + 2k^{1/2}), & 3/(2k) \leq t < 2/k \\ 0, & 2/k \leq t \leq 1 \end{cases} \quad (k = 1, 2, \dots)$$

και τις νόρμες $\|\cdot\|_1$, $\|\cdot\|_2$ και $\|\cdot\|_\infty$ που ορίσαμε στην Παρατήρηση 1.2.1. Τότε έχουμε

$$\|f_k(t)\|_1 = \frac{1}{2\sqrt{k}} \rightarrow 0, \quad \|f_k(t)\|_2 = \frac{1}{\sqrt{3}} \rightarrow \frac{1}{\sqrt{3}} \quad \text{και} \quad \|f_k(t)\|_\infty = \sqrt{k} \rightarrow +\infty.$$

Όλες οι γνωστές l_p -νόρμες έχουν την ιδιότητα να εξαρτώνται αποκλειστικά από τα μέτρα των στοιχείων των διανυσμάτων. Επιπλέον, αποτελούν αύξουσες συναρτήσεις των μέτρων των στοιχείων. Οι δύο αυτές ιδιότητες αποδεικνύονται ισοδύναμες μεταξύ τους. Για ένα διάνυσμα $x = [x_i] \in \mathbb{C}^\nu$, γράφουμε $|x| = [|x_i|]$, ενώ για δύο τυχαία διανύσματα $x = [x_i], y = [y_i] \in \mathbb{C}^\nu$, γράφουμε $|x| \leq |y|$ αν $|x_i| \leq |y_i|$ για κάθε $i = 1, 2, \dots, \nu$.

Ορισμός 1.2.2. Μία νόρμα διανυσμάτων $\|\cdot\|$ καλείται *μονότονη* αν η σχέση $|x| \leq |y|$ συνεπάγεται ότι $\|x\| \leq \|y\|$, ενώ καλείται *απόλυτη* αν $\|x\| = \||x|\|$ για κάθε $x \in \mathbb{C}^\nu$.

Θεώρημα 1.2.8. Μία νόρμα $\|\cdot\|$ στον διανυσματικό χώρο $\mathbb{C}^\nu$ είναι μονότονη αν και μόνο αν είναι απόλυτη.

Απόδειξη. Έστω ότι η νόρμα $\|\cdot\|$ είναι μονότονη. Θεωρούμε ένα τυχαίο διάνυσμα $x \in \mathbb{C}^\nu$ και το $y = |x|$. Τότε

$$|x| \leq |y| \Rightarrow \|x\| \leq \|y\| \quad \text{και} \quad |y| \leq |x| \Rightarrow \|y\| \leq \|x\|.$$

Επομένως, $\|x\| = \|y\| = \||x|\|$, δηλαδή η νόρμα $\|\cdot\|$ είναι απόλυτη.

Για το αντίστροφο, υποθέτουμε ότι η νόρμα $\|\cdot\|$ είναι απόλυτη. Έστω ένα διάνυσμα $x \in \mathbb{C}^\nu$, ένας φυσικός αριθμός $k \in \{1, 2, \dots, \nu\}$ κι ένας πραγματικός αριθμός $a \in [0, 1]$. Τότε ισχύει

$$\begin{aligned} & \left\| [x_1 \ x_2 \ \cdots \ a x_k \ \cdots \ x_\nu]^T \right\| \\ &= \left\| \frac{1-a}{2} [x_1 \ x_2 \ \cdots \ -x_k \ \cdots \ x_\nu]^T + \frac{1-a}{2} x + a x \right\| \\ &\leq \frac{1-a}{2} \left\| [x_1 \ x_2 \ \cdots \ -x_k \ \cdots \ x_\nu]^T \right\| + \frac{1-a}{2} \|x\| + a \|x\| \\ &= \frac{1-a}{2} \|x\| + \frac{1-a}{2} \|x\| + a \|x\| \\ &= \|x\|. \end{aligned}$$

Επαναλαμβάνοντας τη σχέση αυτή για όλα τα στοιχεία του x , καταλήγουμε στο ότι για κάθε επιλογή πραγματικών αριθμών $a_1, a_2, \dots, a_\nu \in [0, 1]$, ισχύει

$$\left\| [a_1 x_1 \ a_2 x_2 \ \cdots \ a_\nu x_\nu]^T \right\| \leq \left\| [x_1 \ x_2 \ \cdots \ x_\nu]^T \right\| = \|x\|.$$

Αν λοιπόν θεωρήσουμε δύο τυχαία διανύσματα $x, y \in \mathbb{C}^\nu$ με $|x| \leq |y|$, τότε υπάρχουν πραγματικοί αριθμοί $a_1, a_2, \dots, a_\nu \in [0, 1]$ και $\theta_1, \theta_2, \dots, \theta_\nu \in [0, 2\pi]$ τέτοιοι ώστε $[x_1 \ x_2 \ \cdots \ x_\nu]^T = [a_1 e^{i\theta_1} y_1 \ a_2 e^{i\theta_2} y_2 \ \cdots \ a_\nu e^{i\theta_\nu} y_\nu]^T$. Κατά συνέ-

πεια, έχουμε

$$\begin{aligned} \left\| [x_1 \ x_2 \ \cdots \ x_\nu]^T \right\| &= \left\| [a_1 e^{i\theta_1} y_1 \ a_2 e^{i\theta_2} y_2 \ \cdots \ a_\nu e^{i\theta_\nu} y_\nu]^T \right\| \\ &= \left\| [a_1 y_1 \ a_2 y_2 \ \cdots \ a_\nu y_\nu]^T \right\| \\ &\leq \left\| [y_1 \ y_2 \ \cdots \ y_\nu]^T \right\|. \end{aligned}$$

Δηλαδή, η νόρμα $\|\cdot\|$ είναι μονότονη. $\square$

1.3 Νόρμες Πινάκων στο $\mathbb{C}^{\nu \times \nu}$

Γνωρίζουμε ότι το σύνολο $\mathbb{C}^{\nu \times \nu}$ των $\nu \times \nu$ μιγαδικών πινάκων είναι ένας διανυσματικός χώρος διάστασης ν^2 , ο οποίος είναι ισόμορφος με το διανυσματικό χώρο $\mathbb{C}^{\nu^2}$. Μπορούμε λοιπόν να ορίσουμε νόρμες πινάκων με τρόπο ανάλογο με αυτόν που ορίσαμε τις νόρμες διανυσμάτων στην Παράγραφο 1.1.

Ορισμός 1.3.1. Μία συνάρτηση $\|\cdot\| : \mathbb{C}^{\nu \times \nu} \rightarrow \mathbb{R}$ ονομάζεται *νόρμα πινάκων* αν για κάθε $A, B \in \mathbb{C}^{\nu \times \nu}$, ικανοποιεί τα ακόλουθα:

- (i) $\|A\| \geq 0$ (μη αρνητική).
- (ii) $\|A\| = 0$ αν και μόνο αν $A = 0$.
- (iii) $\|a A\| = |a| \|A\|$ για κάθε $a \in \mathbb{C}$.
- (iv) $\|A + B\| \leq \|A\| + \|B\|$ (τριγωνική ανισότητα).
- (v) $\|A B\| \leq \|A\| \|B\|$ (υπο-πολλαπλασιαστική ιδιότητα).

Ειδικότερα, για κάθε νόρμα πινάκων $\|\cdot\|$ και για το μοναδιαίο πίνακα I_ν , ισχύει

$$\|I_\nu\| = \|I_\nu^2\| \leq \|I_\nu\|^2 \Rightarrow \|I_\nu\| \geq 1.$$

Επομένως, για κάθε αντιστρέψιμο πίνακα $A \in \mathbb{C}^{\nu \times \nu}$,

$$1 \leq \|I_\nu\| = \|A A^{-1}\| \leq \|A\| \|A^{-1}\| \Rightarrow \|A^{-1}\| \geq \frac{1}{\|A\|}.$$

Κάποιες από τις νόρμες διανυσμάτων που συναντήσαμε στην προηγούμενη παράγραφο αποτελούν νόρμες πινάκων όταν εφαρμόζονται στο διανυσματικό χώρο $\mathbb{C}^{\nu \times \nu}$, ενώ κάποιες άλλες όχι. Τα πιο γνωστά παραδείγματα είναι οι l_p -νόρμες, για $p = 1, 2, \infty$.

- Η l_1 -νόρμα ενός πίνακα $A = [a_{ij}] \in \mathbb{C}^{\nu \times \nu}$ ορίζεται ως

$$\|A\|_{l_1} = \sum_{i,j=1}^{\nu} |a_{ij}|.$$

- Η νόρμα *Frobenius* (ή l_2 -νόρμα) ενός πίνακα $A = [a_{ij}] \in \mathbb{C}^{\nu \times \nu}$ ορίζεται ως

$$\|A\|_F = \left(\sum_{i,j=1}^{\nu} |a_{ij}|^2 \right)^{1/2} = \sqrt{\text{trace}(A^* A)},$$

όπου με $\text{trace}(\cdot)$ συμβολίζουμε το ίχνος πίνακα.

- Η l_∞ -νόρμα ενός πίνακα $A = [a_{ij}] \in \mathbb{C}^{\nu \times \nu}$ ορίζεται ως

$$\|A\|_{l_\infty} = \max\{|a_{ij}| : i, j = 1, 2, \dots, \nu\}.$$

Με απλές πράξεις, μπορεί κανείς να επαληθεύσει ότι οι συναρτήσεις $\|\circ\|_{l_1}$ και $\|\circ\|_F$ αποτελούν πράγματι νόρμες πινάκων. Αντίθετα, η $\|\circ\|_{l_\infty}$, αν και ικανοποιεί τα (i)–(iv) του Ορισμού 1.3.1, δεν ικανοποιεί το (v) του ορισμού και δεν είναι νόρμα πινάκων. Για παράδειγμα,

$$\left\| \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\|_{l_\infty}^2 = \left\| \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \right\|_{l_\infty} = 2 \geq 1 = 1^2 = \left\| \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \right\|_{l_\infty}^2.$$

Αν για έναν πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, συμβολίσουμε με $a_1, a_2, \dots, a_\nu \in \mathbb{C}^\nu$ τις στήλες του A , τότε ο A γράφεται $A = [a_1 \ a_2 \ \dots \ a_\nu]$. Έτσι μπορεί κανείς να παρατηρήσει ότι

$$\|A\|_F^2 = \|a_1\|_2^2 + \|a_2\|_2^2 + \dots + \|a_\nu\|_2^2.$$

Δεδομένου ότι η l_2 -νόρμα διανυσμάτων είναι ορθομοναδιαία αναλλοίωτη, για κάθε ορθομοναδιαίο πίνακα $U \in \mathbb{C}^{\nu \times \nu}$, ισχύει

$$\begin{aligned} \|U A\|_F^2 &= \|[U a_1 \ U a_2 \ \dots \ U a_\nu]\|_F^2 \\ &= \|U a_1\|_2^2 + \|U a_2\|_2^2 + \dots + \|U a_\nu\|_2^2 \\ &= \|a_1\|_2^2 + \|a_2\|_2^2 + \dots + \|a_\nu\|_2^2 \\ &= \|A\|_F^2. \end{aligned}$$

Επίσης, επειδή $\|A^*\|_F = \|A\|_F$ για κάθε $A \in \mathbb{C}^{\nu \times \nu}$, έπεται ότι για κάθε ζεύγος ορθομοναδιαίων πινάκων $U, V \in \mathbb{C}^{\nu \times \nu}$, ισχύει

$$\|U A V\|_F = \|A V\|_F = \|(A V)^*\|_F = \|V^* A^*\|_F = \|A^*\|_F = \|A\|_F.$$

Αυτό σημαίνει ότι η νόρμα Frobenius στο $\mathbb{C}^{\nu \times \nu}$ είναι ορθομοναδιαία αναλλοίωτη.

Ιδιαίτερα σημαντικό ρόλο στις νόρμες πινάκων παίζουν οι επαγόμενες νόρμες, ή αλλιώς φυσικές νόρμες.

Ορισμός 1.3.2. Έστω $\|\cdot\|$ μία νόρμα διανυσμάτων στο $\mathbb{C}^\nu$. Η επαγόμενη από την $\|\cdot\|$ νόρμα στο $\mathbb{C}^{\nu \times \nu}$ ορίζεται ως

$$\|A\| = \max_{\|x\|=1} \|Ax\| = \max_{\|x\| \leq 1} \|Ax\| = \max_{x \neq 0} \frac{\|Ax\|}{\|x\|}.$$

Θεώρημα 1.3.1. Έστω $\|\cdot\|$ μία νόρμα διανυσμάτων στο $\mathbb{C}^\nu$ και $\|\circ\|$ η επαγόμενη από την $\|\cdot\|$ νόρμα στο $\mathbb{C}^{\nu \times \nu}$. Τότε ισχύουν τα ακόλουθα:

- (i) $\|H\|$ είναι νόρμα πινάκων.
- (ii) $\|Ax\| \leq \|A\| \|x\|$ για κάθε $A \in \mathbb{C}^{\nu \times \nu}$ και $x \in \mathbb{C}^\nu$.
- (iii) $\|I_\nu\| = 1$.

Απόδειξη. (i) Για να αποδείξουμε ότι η συνάρτηση $\|\circ\|$ είναι νόρμα πινάκων, αρκεί να επαληθεύσουμε τον Ορισμό 1.3.1. Τα (i)–(iv) του ορισμού προκύπτουν άμεσα από τη σχέση

$$\|A\| = \max_{\|x\|=1} \|Ax\| \quad (A \in \mathbb{C}^{\nu \times \nu}).$$

Για το (v) του Ορισμού 1.3.1, θεωρούμε δύο πίνακες $A, B \in \mathbb{C}^{\nu \times \nu}$ και παρατηρούμε ότι

$$\begin{aligned} \|AB\| &= \max_{x \neq 0} \frac{\|ABx\|}{\|x\|} = \max_{Bx \neq 0} \left(\frac{\|A(Bx)\|}{\|Bx\|} \frac{\|Bx\|}{\|x\|} \right) \\ &\leq \left(\max_{Bx \neq 0} \frac{\|A(Bx)\|}{\|Bx\|} \right) \left(\max_{x \neq 0} \frac{\|Bx\|}{\|x\|} \right) = \|A\| \|B\|. \end{aligned}$$

(ii) Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Τότε για κάθε μη μηδενικό διάνυσμα $x \in \mathbb{C}^\nu$, ισχύει

$$\frac{\|Ax\|}{\|x\|} \leq \|A\|,$$

ή ισοδύναμα,

$$\|Ax\| \leq \|A\| \|x\|.$$

Η τελευταία σχέση ισχύει κατά τετριμμένο τρόπο για $x = 0$.

- (iii) Προφανώς, έχουμε ότι $\|I_\nu\| = \max_{\|x\|=1} \|I_\nu x\| = \max_{\|x\|=1} \|x\| = 1$. □

Το θεώρημα αυτό παρέχει την αναγκαία συνθήκη $\|I_\nu\| = 1$ για μία νόρμα πινάκων στο $\mathbb{C}^{\nu \times \nu}$ ώστε να επάγεται από κάποια νόρμα διανυσμάτων στο $\mathbb{C}^\nu$, η οποία όμως δυστυχώς δεν είναι και ικανή συνθήκη.

Πόρισμα 1.3.2. Έστω $\|\circ\|$ μία επαγόμενη (φυσική) νόρμα στο $\mathbb{C}^{\nu \times \nu}$.

- (i) Για κάθε πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, ισχύει $\|A^k\| \leq \|A\|^k$, $k = 1, 2, \dots$
- (ii) Για κάθε αντιστρέψιμο πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, ισχύει $\|A\|^{-1} \leq \|A^{-1}\|$.

Στη συνέχεια, αναφέρουμε τα σημαντικότερα παραδείγματα νορμών πινάκων που επάγονται από τις γνωστές l_p -νόρμες και οι οποίες μπορούν να υπολογιστούν χωρίς να απαιτείται η χρήση του Ορισμού 1.3.2. Σε κάθε περίπτωση, θεωρούμε έναν τετραγωνικό πίνακα $A = [a_{ij}] \in \mathbb{C}^{\nu \times \nu}$.

- Η νόρμα πινάκων μεγίστου αθροίσματος κατά στήλη στο $\mathbb{C}^{\nu \times \nu}$ ορίζεται ως

$$\|A\|_1 = \max_{1 \leq j \leq \nu} \sum_{i=1}^{\nu} |a_{ij}|.$$

Η νόρμα $\|\circ\|_1$ επάγεται από τη διανυσματική l_1 -νόρμα. Πράγματι, αν θεωρήσουμε τον πίνακα A γραμμένο ως προς τις στήλες, $A = [a_1 \ a_2 \ \dots \ a_\nu]$, τότε για κάθε διάνυσμα $x = [x_i] \in \mathbb{C}^\nu$, ισχύει

$$\begin{aligned} \|Ax\|_1 &= \|x_1 a_1 + x_2 a_2 + \dots + x_\nu a_\nu\|_1 \\ &\leq |x_1| \|a_1\|_1 + |x_2| \|a_2\|_1 + \dots + |x_\nu| \|a_\nu\|_1 \\ &\leq (|x_1| + |x_2| + \dots + |x_\nu|) \max_{1 \leq k \leq \nu} \|a_k\|_1 \\ &= \|x\|_1 \|A\|_1. \end{aligned}$$

Επομένως,

$$\max_{\|x\|_1=1} \|Ax\|_1 \leq \|A\|_1.$$

Αν τώρα επιλέξουμε $x = e_k$ το διάνυσμα της κανονικής βάσης που αντιστοιχεί στην στήλη a_k του A με τη μεγαλύτερη l_1 -νόρμα, τότε παρατηρούμε ότι $\|A e_k\|_1 = \|a_k\|_1 = \|A\|_1$. Δηλαδή,

$$\max_{\|x\|_1=1} \|Ax\|_1 = \|A\|_1.$$

Συνεπώς, η νόρμα $\|\circ\|_1$ επάγεται από τη διανυσματική l_1 -νόρμα και αποτελεί νόρμα πινάκων.

- Η νόρμα πινάκων μεγίστου αθροίσματος κατά γραμμή στο $\mathbb{C}^{\nu \times \nu}$ ορίζεται ως

$$\|A\|_\infty = \max_{1 \leq i \leq \nu} \sum_{j=1}^{\nu} |a_{ij}|.$$

Η νόρμα $\|\circ\|_\infty$ επάγεται από τη διανυσματική l_∞ -νόρμα, δηλαδή

$$\|A\|_\infty = \max_{\|x\|_\infty=1} \|Ax\|_\infty.$$

Η απόδειξη αυτής της παρατήρησης ακολουθεί τα βήματα της απόδειξης του προηγούμενου παραδείγματος και αφήνεται στον αναγνώστη ως άσκηση. Προφανώς, και η νόρμα $\|\circ\|_\infty$ (ως επαγόμενη) αποτελεί νόρμα πινάκων.

- Η φασματική (τελεστική) νόρμα πινάκων στο $\mathbb{C}^{\nu \times \nu}$ ορίζεται ως

$$\|A\|_2 = \max \left\{ \sqrt{\lambda} : \lambda \in \sigma(A^*A) \right\},$$

όπου με $\sigma(A)$ συμβολίζουμε το φάσμα ενός πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, δηλαδή το σύνολο των ιδιοτιμών του A . Παρατηρούμε ότι αν $\lambda \in \sigma(A^*A)$ και $x \in \mathbb{C}^\nu$ ένα (μη μηδενικό) ιδιοδιάνυσμα του A^*A που αντιστοιχεί στην ιδιοτιμή λ , τότε

$$(Ax)^*(Ax) = x^*A^*Ax = x^*\lambda x = \lambda(x^*x) \Rightarrow \|Ax\|_2^2 = \lambda\|x\|_2^2.$$

Επομένως, οι ιδιοτιμές του πίνακα A^*A είναι μη αρνητικές κι έτσι μπορούν να οριστούν οι (μη αρνητικές) τετραγωνικές ρίζες τους. Επιπλέον, η φασματική νόρμα $\|\circ\|_2$ είναι νόρμα πινάκων η οποία επάγεται από την Ευκλείδεια νόρμα $\|\cdot\|_2$, δηλαδή

$$\|A\|_2 = \max_{\|x\|_2=1} \|Ax\|_2,$$

και είναι ορθομοναδιαία αναλλοίωτη, δηλαδή $\|UAV\|_2 = \|A\|_2$ για κάθε ζεύγος ορθομοναδιαίων πινάκων $U, V \in \mathbb{C}^{\nu \times \nu}$. Οι αποδείξεις αυτών των ιδιοτήτων της φασματικής νόρμας $\|\circ\|_2$ αφήνονται στον αναγνώστη ως ασκήσεις.

Ακολουθώντας τα βήματα της απόδειξης του Λήμματος 1.2.1, μπορεί κανείς να επαληθεύσει το επόμενο αποτέλεσμα.

Πρόταση 1.3.3. *Μία νόρμα πινάκων $\|\circ\| : \mathbb{C}^{\nu \times \nu} \rightarrow [0, +\infty)$ είναι ομοιόμορφα συνεχής συνάρτηση των στοιχείων του πίνακα.*

Η ισοδυναμία των νορμών πινάκων στο $\mathbb{C}^{\nu \times \nu}$ ορίζεται και αποδεικνύεται με τον ίδιο ακριβώς τρόπο που ορίζεται και αποδεικνύεται η ισοδυναμία των νορμών διανυσμάτων στο $\mathbb{C}^\nu$. Επιπλέον, για τις νόρμες πινάκων $\|\circ\|_1$, $\|\circ\|_2$, $\|\circ\|_\infty$, $\|\circ\|_F$ και $\|\circ\|_{l_1}$, έχουμε το ακόλουθο αποτέλεσμα.

Πρόταση 1.3.4. *Για κάθε πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, ισχύουν οι ανισότητες*

$$\frac{1}{\sqrt{\nu}} \|A\|_2 \leq (\|A\|_1 \text{ ή } \|A\|_\infty) \leq \sqrt{\nu} \|A\|_2,$$

$$\frac{1}{\sqrt{\nu}} \|A\|_F \leq (\|A\|_1 \text{ ή } \|A\|_\infty) \leq \sqrt{\nu} \|A\|_F,$$

$$\frac{1}{\nu} \|A\|_\infty \leq \|A\|_1 \leq \nu \|A\|_\infty,$$

$$\frac{1}{\sqrt{\nu}} \|A\|_F \leq \|A\|_2 \leq \|A\|_F$$

και

$$(\|A\|_1 \text{ ή } \|A\|_\infty \text{ ή } \|A\|_F) \leq \|A\|_{l_1} \leq (\nu \|A\|_1 \text{ ή } \nu \|A\|_\infty \text{ ή } \nu \|A\|_F).$$

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

1.4 Φασματική Ακτίνα και Νόρμες Πινάκων

Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ με φάσμα $\sigma(A) = \{\lambda \in \mathbb{C} : \det(\lambda I_\nu - A) = 0\}$. Η φασματική ακτίνα του A ορίζεται ως

$$\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}.$$

Θεώρημα 1.4.1. Έστω $\|\cdot\|$ μία νόρμα πινάκων στο $\mathbb{C}^{\nu \times \nu}$. Τότε για κάθε $A \in \mathbb{C}^{\nu \times \nu}$, ισχύει $\rho(A) \leq \|A\|$.

Απόδειξη. Για κάθε ιδιοτιμή λ του πίνακα A , ισχύει $|\lambda| \leq \rho(A)$. Επιπλέον, υπάρχει τουλάχιστον μία ιδιοτιμή $\lambda_0 \in \sigma(A)$ τέτοια ώστε $|\lambda_0| = \rho(A)$. Θεωρούμε ακόμη ένα ιδιοδιάνυσμα $x_0 \in \mathbb{C}^\nu$ του A που αντιστοιχεί στην ιδιοτιμή λ_0 και τον $\nu \times \nu$ πίνακα $X_0 = [x_0 \ x_0 \ \cdots \ x_0]$ (δηλαδή, με όλες τις στήλες του ίσες με x_0). Τότε έχουμε

$$A X_0 = \lambda_0 X_0 \text{ και } |\lambda_0| \|X_0\| = \|\lambda_0 X_0\| = \|A X_0\| \leq \|A\| \|X_0\|.$$

Επομένως, $\rho(A) = |\lambda_0| \leq \|A\|$. $\square$

Η φασματική ακτίνα δεν αποτελεί νόρμα πινάκων. Είναι αξιοσημείωτο ότι η σχέση $\rho(A) = 0$ δεν συνεπάγεται ότι $A = 0$, ενώ μπορεί κανείς να βρει δύο πίνακες $A, B \in \mathbb{C}^{\nu \times \nu}$ τέτοιους ώστε $\rho(A + B) > \rho(A) + \rho(B)$. Παρ' όλα αυτά, η σχέση $\rho(A) \leq \|A\|$ (για οποιαδήποτε νόρμα πινάκων) του Θεωρήματος 1.4.1 εξασφαλίζει ότι η συνάρτηση $\rho(\cdot) : \mathbb{C}^{\nu \times \nu} \rightarrow [0, +\infty)$ είναι ομοιόμορφα συνεχής συνάρτηση των στοιχείων του πίνακα.

Από το σημείο αυτό και για τη συνέχεια, μας είναι απαραίτητο το γνωστό Λήμμα του Schur.

Λήμμα 1.4.2. (Schur) Έστω ένας τυχαίος πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Τότε υπάρχουν ένας ορθομοναδιαίος πίνακας $U \in \mathbb{C}^{\nu \times \nu}$ κι ένας άνω τριγωνικός πίνακας $T \in \mathbb{C}^{\nu \times \nu}$ με διαγώνια στοιχεία τις ιδιοτιμές του A (λαμβάνοντας υπόψη και τις πολλαπλότητες), τέτοιοι ώστε

$$A = U T U^*.$$

Απόδειξη. Έστω $\lambda_1, \lambda_2, \dots, \lambda_\nu$ οι (όχι κατ' ανάγκη διακεκριμένες) ιδιοτιμές του A και $x_1 \in \mathbb{C}^\nu$ ένα ιδιοδιάνυσμα του A , με $x_1^* x_1 = 1$, το οποίο αντιστοιχεί στην ιδιοτιμή λ_1 του A . Θεωρούμε επίσης μία ορθοκανονική βάση $\{x_1, w_2, w_3, \dots, w_\nu\}$ του $\mathbb{C}^\nu$ και τον αντίστοιχο ορθομοναδιαίο πίνακα

$$W_1 = [x_1 \ w_2 \ w_3 \ \cdots \ w_\nu] = \begin{bmatrix} x_1 & \hat{W}_1 \end{bmatrix} \in \mathbb{C}^{\nu \times \nu},$$

όπου $\hat{W}_1 = [w_2 \ w_3 \ \cdots \ w_\nu] \in \mathbb{C}^{\nu \times (\nu-1)}$. Αφού $w_2^* x_1 = w_3^* x_1 = \cdots = w_\nu^* x_1 = 0$, έπεται ότι $\hat{W}_1^* x_1 = 0$. Κατά συνέπεια, με απλές πράξεις βλέπουμε ότι

$$\begin{aligned} W_1^* A W_1 &= \begin{bmatrix} x_1^* \\ \hat{W}_1^* \end{bmatrix} A \begin{bmatrix} x_1 & \hat{W}_1 \end{bmatrix} = \begin{bmatrix} x_1^* \\ \hat{W}_1^* \end{bmatrix} \begin{bmatrix} A x_1 & A \hat{W}_1 \end{bmatrix} \\ &= \begin{bmatrix} x_1^* \lambda_1 x_1 & x_1^* A \hat{W}_1 \\ \hat{W}_1^* \lambda_1 x_1 & \hat{W}_1^* A \hat{W}_1 \end{bmatrix} = \begin{bmatrix} \lambda_1 & x_1^* A \hat{W}_1 \\ 0 & \hat{W}_1^* A \hat{W}_1 \end{bmatrix}. \end{aligned}$$

Ο $(\nu-1) \times (\nu-1)$ πίνακας $A_2 = \hat{W}_1^* A \hat{W}_1$ έχει ιδιοτιμές ακριβώς τις $\lambda_2, \lambda_3, \dots, \lambda_\nu$. Ομοίως με παραπάνω, μπορούμε να κατασκευάσουμε έναν $(\nu-1) \times (\nu-1)$ ορθομοναδιαίο πίνακα

$$W_2 = \begin{bmatrix} x_2 & \hat{W}_2 \end{bmatrix} \in \mathbb{C}^{(\nu-1) \times (\nu-1)} \quad (\text{όπου } \hat{W}_2 \in \mathbb{C}^{(\nu-1) \times (\nu-2)}),$$

τέτοιον ώστε

$$W_2^* A_2 W_2 = \begin{bmatrix} \lambda_2 & x_2^* A_2 \hat{W}_2 \\ 0 & \hat{W}_2^* A_2 \hat{W}_2 \end{bmatrix}.$$

Επομένως,

$$\begin{bmatrix} 1 & 0 \\ 0 & W_2 \end{bmatrix}^* W_1^* A W_1 \begin{bmatrix} 1 & 0 \\ 0 & W_2 \end{bmatrix} = \begin{bmatrix} \lambda_1 & * & * \\ 0 & \lambda_2 & x_2^* A_2 \hat{W}_2 \\ 0 & 0 & \hat{W}_2^* A_2 \hat{W}_2 \end{bmatrix},$$

όπου ο $(\nu-2) \times (\nu-2)$ πίνακας $A_3 = \hat{W}_2^* A_2 \hat{W}_2$ έχει ιδιοτιμές ακριβώς τις $\lambda_3, \dots, \lambda_\nu$. Επαναλαμβάνοντας τα ίδια βήματα, μπορούμε να κατασκευάσουμε συνολικά $\nu-1$ ορθομοναδιαίους πίνακες

$$W_1 \in \mathbb{C}^{\nu \times \nu}, \ W_2 \in \mathbb{C}^{(\nu-1) \times (\nu-1)}, \ W_3 \in \mathbb{C}^{(\nu-2) \times (\nu-2)}, \ \dots, \ W_{\nu-1} \in \mathbb{C}^{2 \times 2}$$

τέτοιους ώστε ο πίνακας

$$U = W_1 \begin{bmatrix} 1 & 0 \\ 0 & W_2 \end{bmatrix} \begin{bmatrix} I_2 & 0 \\ 0 & W_3 \end{bmatrix} \cdots \begin{bmatrix} I_{\nu-2} & 0 \\ 0 & W_{\nu-1} \end{bmatrix}$$

να ικανοποιεί τη σχέση

$$U^* A U = \begin{bmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & \lambda_2 & * & \cdots & * \\ 0 & 0 & \lambda_3 & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & \lambda_\nu \end{bmatrix}.$$

Ο πίνακας U είναι προφανώς ορθομοναδιαίος και η απόδειξη ολοκληρώθηκε. $\square$

Στο θεώρημα που ακολουθεί αποδεικνύεται ότι η φασματική ακτίνα $\rho(A)$ είναι το μέγιστο κάτω φράγμα (infimum) των νορμών πινάκων.

Θεώρημα 1.4.3. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ κι ένας πραγματικός αριθμός $\varepsilon > 0$. Τότε υπάρχει μία νόρμα πινάκων $\|\circ\|$ τέτοια ώστε $\|A\| \leq \rho(A) + \varepsilon$.

Απόδειξη. Από το Λήμμα 1.4.2 του Schur, υπάρχουν ένας ορθομοναδιαίος πίνακας $U \in \mathbb{C}^{\nu \times \nu}$ και ένας άνω τριγωνικός πίνακας $T \in \mathbb{C}^{\nu \times \nu}$ τέτοιοι ώστε $A = UTU^*$. Τα διαγώνια στοιχεία του T , $\lambda_1, \lambda_2, \dots, \lambda_\nu$, είναι οι ιδιοτιμές του πίνακα A , λαμβάνοντας υπόψη και τις πολλαπλότητες.

Θεωρούμε ένα διαγώνιο πίνακα $D_t = \text{diag}\{t, t^2, \dots, t^\nu\}$ για κάποιον πραγματικό αριθμό $t \neq 0$. Ο αντίστροφος του D_t είναι ο $D_t^{-1} = \text{diag}\{t^{-1}, t^{-2}, \dots, t^{-\nu}\}$. Υπολογίζουμε τον πίνακα

$$D_t T D_t^{-1} = \begin{bmatrix} \lambda_1 & t^{-1} * & t^{-2} * & \dots & t^{-\nu+1} * \\ 0 & \lambda_2 & t^{-1} * & \dots & t^{-\nu+2} * \\ 0 & 0 & \lambda_3 & \dots & t^{-\nu+3} * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & \lambda_\nu \end{bmatrix},$$

όπου με “*” συμβολίζουμε (μιγαδικούς) αριθμούς ανεξάρτητους του t . Για $t > 0$ αρκετά μεγάλο, το άθροισμα των μέτρων όλων των μη διαγωνίων στοιχείων του $D_t T D_t^{-1}$ είναι μικρότερο ή ίσο του $\varepsilon > 0$ και συνεπώς, $\|D_t T D_t^{-1}\|_1 \leq \rho(A) + \varepsilon$. Ας ορίσουμε τώρα τη συνάρτηση $\|\circ\|_{D_t^{-1}} : \mathbb{C}^{\nu \times \nu} \rightarrow [0, +\infty)$, με

$$\|M\|_{D_t^{-1}} = \|D_t U^* M U D_t^{-1}\|_1 = \|(U D_t^{-1})^{-1} M U D_t^{-1}\|_1, \quad \forall M \in \mathbb{C}^{\nu \times \nu}.$$

Μπορούμε να επαληθεύσουμε ότι, αφού ο πίνακας $U D_t^{-1}$ είναι αντιστρέψιμος, η συνάρτηση $\|\circ\|_{D_t^{-1}}$, είναι μία νόρμα πινάκων (βλέπε την Άσκηση 1.5.5 παρακάτω), για την οποία προφανώς ισχύει ότι $\|A\|_{D_t^{-1}} = \|D_t T D_t^{-1}\|_1 \leq \rho(A) + \varepsilon$. $\square$

Ολοκληρώνοντας το πρώτο κεφάλαιο, παρουσιάζουμε κάποιες ενδιαφέρουσες (κυρίως από μεριάς εφαρμογών σε επαναληπτικές αριθμητικές μεθόδους) ιδιότητες των νορμών πινάκων.

Θεώρημα 1.4.4. Έστω ένα πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Τότε $\lim A^k = 0$ αν και μόνο αν $\rho(A) < 1$.

Απόδειξη. Αν $\lim A^k = 0$, τότε είναι προφανές ότι για κάθε ιδιοτιμή $\lambda \in \sigma(A)$, ισχύει $\lim \lambda^k = 0$. Αυτό σημαίνει ότι όλες οι ιδιοτιμές του A έχουν μέτρο μικρότερο του 1 κι επομένως, $\rho(A) < 1$.

Για το αντίστροφο, υποθέτουμε ότι $\rho(A) < 1$. Τότε, από το Θεώρημα 1.4.3, υπάρχει μία νόρμα πινάκων $\|\circ\|$ τέτοια ώστε $\|A\| < 1$. Αφού για κάθε φυσικό αριθμό k , ισχύει $\|A^k\| \leq \|A\|^k$, έπεται ότι $\lim \|A^k\| \leq \lim \|A\|^k = 0$. $\square$

Θεώρημα 1.4.5. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Για κάθε νόρμα πινάκων $\|\circ\|$ στο $\mathbb{C}^{\nu \times \nu}$, ισχύει

$$\rho(A) = \lim \|A^k\|^{1/k}.$$

Απόδειξη. Από το Θεώρημα 1.4.1, γνωρίζουμε ότι $\rho(A) \leq \|A\|$. Επομένως, για κάθε $k = 1, 2, \dots$,

$$\rho(A)^k = \rho(A^k) \leq \|A^k\| \Rightarrow \rho(A) \leq \|A^k\|^{1/k}.$$

Για ένα τυχαίο $\varepsilon > 0$, θεωρούμε τον πίνακα $A_\varepsilon = (\rho(A) + \varepsilon)^{-1}A$, ο οποίος έχει φασματική ακτίνα $\rho(A_\varepsilon) = (\rho(A) + \varepsilon)^{-1}\rho(A) < 1$, και από το Θεώρημα 1.4.4, $\lim \|A_\varepsilon^k\| = 0$. Λόγω της σύγκλισης αυτής, υπάρχει φυσικός αριθμός N_ε τέτοιος ώστε

$$\|A_\varepsilon^k\| < 1, \quad \forall k \geq N_\varepsilon,$$

ή ισοδύναμα,

$$\|A^k\|^{1/k} < \rho(A) + \varepsilon, \quad \forall k \geq N_\varepsilon,$$

Η σχέση αυτή ισχύει για κάθε επιλογή του $\varepsilon > 0$, και δεδομένου ότι $\rho(A) \leq \|A^k\|^{1/k}$, καταλήγουμε στο ότι $\rho(A) = \lim \|A^k\|^{1/k}$. $\square$

Θεώρημα 1.4.6. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Αν υπάρχει νόρμα πινάκων $\|\circ\|$ στο $\mathbb{C}^{\nu \times \nu}$ τέτοια ώστε $\|I_\nu - A\| < 1$, τότε ο A είναι αντιστρέψιμος και

$$A^{-1} = \sum_{k=0}^{+\infty} (I_\nu - A)^k.$$

Απόδειξη. Ο χώρος $\mathbb{C}^{\nu \times \nu}$ είναι πλήρης (ως χώρος πεπερασμένης διάστασης). Προκειμένου λοιπόν να αποδείξουμε ότι η σειρά $\sum_{k=0}^{+\infty} (I_\nu - A)^k$ συγκλίνει, αρκεί να δείξουμε ότι η ακολουθία των μερικών αθροισμάτων $S_n = \sum_{k=0}^n (I_\nu - A)^k$ είναι Cauchy (βασική). Πράγματι, αφού

$$\sum_{k=0}^{+\infty} \|I_\nu - A\|^k = \frac{1}{1 - \|I_\nu - A\|},$$

γνωρίζουμε ότι για τυχαίο $\varepsilon > 0$, υπάρχει φυσικός αριθμός n_ε τέτοιος ώστε

$$\left\| \sum_{k=n_\varepsilon}^{+\infty} (I_\nu - A)^k \right\| \leq \sum_{k=n_\varepsilon}^{+\infty} \|(I_\nu - A)^k\| \leq \sum_{k=n_\varepsilon}^{+\infty} \|I_\nu - A\|^k \leq \varepsilon.$$

Συνεπώς, για οποιονδήποτε φυσικό αριθμό $m > n_\varepsilon$, έχουμε ότι

$$\|S_m - S_{n_\varepsilon}\| = \left\| \sum_{k=n_\varepsilon+1}^m (I_\nu - A)^k \right\| \leq \sum_{k=n_\varepsilon+1}^m \|(I_\nu - A)^k\| \leq \sum_{k=n_\varepsilon+1}^m \|I_\nu - A\|^k \leq \varepsilon.$$

Επομένως, η ακολουθία των μερικών αθροισμάτων $S_n = \sum_{k=0}^n (I_\nu - A)^k$ είναι Cauchy και η σειρά $\sum_{k=0}^{+\infty} (I_\nu - A)^k$ συγχλίνει.

Τέλος, παρατηρούμε ότι, για $n \rightarrow +\infty$, η ακολουθία πινάκων

$$I_\nu - (I_\nu - A)^{n+1} = [I_\nu - (I_\nu - A)] \sum_{k=0}^n (I_\nu - A)^k = A \sum_{k=0}^n (I_\nu - A)^k$$

συγχλίνει στον ταυτοτικό πίνακα I_ν (βλέπε το Θεώρημα 1.4.4). $\square$

Αν θεωρήσουμε την υπόθεση του θεωρήματος αυτού για τη νόρμα πινάκων $\|\circ\|_1$, τότε οδηγούμαστε στο ακόλουθο πόρισμα.

Πόρισμα 1.4.7. Έστω ένα πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Αν για τα στοιχεία της (κυρίας) διαγωνίου του A ισχύει

$$|a_{ii}| > |a_{i1}| + \cdots + |a_{i,i-1}| + |a_{i,i+1}| + \cdots + |a_{i,\nu}|, \quad \forall i = 1, 2, \dots, \nu,$$

τότε ο πίνακας A είναι αντιστρέψιμος.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Ένας πίνακας που ικανοποιεί την υπόθεση του παραπάνω πορίσματος λέμε ότι έχει ισχυρή διαγώνια κυριαρχία.

1.5 Ασκήσεις

1. Να αποδείξετε την Πρόταση 1.2.4.
2. Να αποδείξετε ότι οι $\|\circ\|_1$ και $\|\circ\|_F$ ικανοποιούν τον Ορισμό 1.3.1 της νόρμας πινάκων.
3. Να αποδείξετε ότι η συνάρτηση $\|\circ\| = \nu \|\circ\|_\infty$ είναι νόρμα πινάκων στο $\mathbb{C}^{\nu \times \nu}$.
4. Να αποδείξετε ότι η φασματική νόρμα $\|\circ\|_2$ στο $\mathbb{C}^{\nu \times \nu}$ επάγεται από την Ευκλείδεια νόρμα διανυσμάτων $\|\cdot\|_2$ στο $\mathbb{C}^\nu$ και είναι ορθομοναδιαία αναλ-λοϊωτη.
5. Έστω $\|\circ\|$ μία νόρμα πινάκων στο $\mathbb{C}^{\nu \times \nu}$ και $S \in \mathbb{C}^{\nu \times \nu}$ ένας αντιστρέψιμος πίνακας. Να αποδείξετε ότι η συνάρτηση $\|A\|_S = \|S^{-1}AS\|$ αποτελεί νόρμα πινάκων.

6. Να αποδείξετε ότι για κάθε ιδιοτιμή λ ενός αντιστρέψιμου πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, ισχύει $\|A^{-1}\|_p^{-1} \leq |\lambda| \leq \|A\|_p$.
7. Να αποδείξετε ότι για κάθε $A \in \mathbb{C}^{\nu \times \mu}$, ισχύει $\|A\|_2^2 \leq \|A\|_\infty \|A\|_1$.
8. Να αποδείξετε ότι η νόρμα Frobenius επάγεται από εσωτερικό γινόμενο πινάκων, ενώ η 2-νόρμα δεν επάγεται από εσωτερικό γινόμενο πινάκων.
9. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Να αποδείξετε ότι για κάθε ερμιτιανό πίνακα $H \in \mathbb{C}^{\nu \times \nu}$, ισχύει $\|A - \frac{A+A^*}{2}\|_F \leq \|A - H\|_F$.
10. Να αποδείξετε ότι για κάθε $A, B \in \mathbb{C}^{\nu \times \nu}$, ισχύει $\|AB\|_F \leq \|A\|_2 \|B\|_F$ και $\|AB\|_F \leq \|A\|_F \|B\|_2$.
11. Να αποδείξετε ότι για κάθε πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, ισχύει

$$\|A\|_2 = \max\{|y^* A x| : x, y \in \mathbb{C}^\nu, \|x\|_2 = \|y\|_2 = 1\}.$$
12. Να αποδείξετε την Πρόταση 1.3.4.
13. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ με φασματική ακτίνα $\rho(A) < 1$. Να αποδείξετε ότι η σειρά $I_\nu + A + A^2 + \dots$ συγκλίνει στον πίνακα $(I_\nu - A)^{-1}$.
14. Να αποδείξετε το Πόρισμα 1.4.7.

Κεφάλαιο 2

Παραγοντοποίηση LU

2.1 Εισαγωγικό Παράδειγμα

Είναι γνωστό ότι η επίλυση ενός τριγωνικού γραμμικού συστήματος είναι μία εύκολη διαδικασία. Στην παρατήρηση αυτή βασίζεται και η μέθοδος απαλοιφής του Gauss, η οποία έχει ως βασικό στόχο τη μετατροπή ενός δεδομένου γραμμικού συστήματος σε ένα ισοδύναμο τριγωνικό σύστημα. Η μετατροπή αυτή επιτυγχάνεται με κατάλληλες πράξεις γραμμών (εξισώσεων), όπως στο παρακάτω σύστημα:

$$\left. \begin{array}{rrcr} x & + y & + z & = 6 \\ 2x & - y & - z & = -3 \\ 4x & - 2y & + z & = 3 \end{array} \right\} \begin{array}{l} \Gamma_2 \rightarrow \Gamma_2 - 2\Gamma_1 \\ \Gamma_3 \rightarrow \Gamma_3 - 4\Gamma_1 \end{array}$$

$$\left. \begin{array}{rrcr} x & + y & + z & = 6 \\ & - 3y & - 3z & = -15 \\ & - 6y & - 3z & = -21 \end{array} \right\} \Gamma_3 \rightarrow \Gamma_3 - 2\Gamma_2$$

$$\left. \begin{array}{rrcr} x & + y & + z & = 6 \\ & - 3y & - 3z & = -15 \\ & & 3z & = 9 \end{array} \right\}.$$

Το τελικό σύστημα είναι ισοδύναμο με το αρχικό, και ως τριγωνικό, επιλύεται πολύ εύκολα:

$$\begin{aligned} 3z &= 9 \Rightarrow z = 3, \\ -3y - 9 &= -15 \Rightarrow -3y = 6 \Rightarrow y = 2, \\ x + 2 + 3 &= 6 \Rightarrow x = 1. \end{aligned}$$

Σημειώνουμε ότι το παράδειγμα μας έχει επιλέγει έτσι ώστε: (i) δεν εμφανίζεται μηδενικό διαγώνιο στοιχείο, και συνεπώς, δεν απαιτείται αντιμετάθεση γραμμών,

(ii) δεν πολλαπλασιάζουμε (ή διαιρούμε) με κάποιον αριθμό, εξίσωση του συστήματος ώστε να την απλοποιήσουμε, και (iii) ο πίνακας του (αρχικού) συστήματος,

$$\begin{bmatrix} 1 & 1 & 1 \\ 2 & -1 & -1 \\ 4 & -2 & 1 \end{bmatrix},$$

είναι αντιστρέψιμος. Στις παραγράφους που ακολουθούν θα δικαιολογηθεί πλήρως η επιλογή μας αυτή.

2.2 Στοιχειώδεις Πίνακες Gauss και Παραγοντοποίηση LU

Έστω ένα $\nu \times \nu$ γραμμικό σύστημα

$$\left. \begin{array}{cccc} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1\nu}x_\nu & = & \beta_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2\nu}x_\nu & = & \beta_2 \\ \vdots & & \vdots \\ a_{\nu 1}x_1 + a_{\nu 2}x_2 + \cdots + a_{\nu\nu}x_\nu & = & \beta_\nu \end{array} \right\},$$

ή όπως γράφεται με χρήση πινάκων,

$$\underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1\nu} \\ a_{21} & a_{22} & \cdots & a_{2\nu} \\ \vdots & \vdots & \ddots & \vdots \\ a_{\nu 1} & a_{\nu 2} & \cdots & a_{\nu\nu} \end{bmatrix}}_A \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_\nu \end{bmatrix}}_x = \underbrace{\begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_\nu \end{bmatrix}}_\beta.$$

Αν ο πίνακας A του συστήματος είναι αντιστρέψιμος, τότε η μετατροπή του σε άνω τριγωνικό, σύμφωνα με την απαλοιφή Gauss, αποτελείται από $\nu - 1$ βήματα όπου στο k -οστό βήμα ($k = 1, 2, \dots, \nu - 1$) μηδενίζονται τα στοιχεία της k -στήλης του πίνακα κάτω από τη διαγώνιο. Δηλαδή, έχουμε τη μετατροπή

$$\begin{bmatrix} a_{1k} \\ a_{2k}^{(1)} \\ \vdots \\ a_{kk}^{(k-1)} \\ a_{k+1,k}^{(k-1)} \\ \vdots \\ a_{\nu k}^{(k-1)} \end{bmatrix} \rightarrow \begin{bmatrix} a_{1k} \\ a_{2k}^{(1)} \\ \vdots \\ a_{kk}^{(k-1)} \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Οι εκθέτες (εντός παρενθέσεων) που παρουσιάζονται στα στοιχεία κάτω από το πρώτο δηλώνουν ότι τα στοιχεία αυτά έχουν ήδη μεταβληθεί κατά τα προηγούμενα $k - 1$ βήματα της απαλοιφής Gauss. Η μετατροπή αυτή ολοκληρώνεται όταν για κάθε $i = k + 1, k + 2, \dots, \nu$, προσθέσουμε στη i -γραμμή την k -γραμμή πολλαπλασιασμένη με τον συντελεστή $r_i^{(k-1)} = -a_{ik}^{(k-1)}/a_{kk}^{(k-1)}$. Με χρήση γινόμενου πινάκων, γράφουμε

$$\begin{array}{c} \underbrace{k-\text{στήλη}} \\ \left[\begin{array}{ccccccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & r_{k+1}^{(k-1)} & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{\nu}^{(k-1)} & 0 & \cdots & 1 \end{array} \right] \left[\begin{array}{c} a_{1k} \\ a_{2k}^{(1)} \\ \vdots \\ a_{kk}^{(k-1)} \\ a_{k+1,k}^{(k-1)} \\ \vdots \\ a_{\nu k}^{(k-1)} \end{array} \right] = \left[\begin{array}{c} a_{1k} \\ a_{2k}^{(1)} \\ \vdots \\ a_{kk}^{(k-1)} \\ 0 \\ \vdots \\ 0 \end{array} \right]. \end{array}$$

Προφανώς, όπως και στο εισαγωγικό παράδειγμα, υποθέτουμε ότι κατά την απαλοιφή Gauss, δεν μηδενίζονται διαγώνια στοιχεία του πίνακα του συστήματος. Ο $\nu \times \nu$ πίνακας

$$M_k = \begin{array}{c} \underbrace{k-\text{στήλη}} \\ \left[\begin{array}{ccccccc} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & r_{k+1}^{(k-1)} & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & r_{\nu}^{(k-1)} & 0 & \cdots & 1 \end{array} \right] \end{array}$$

καλείται *στοιχειώδης πίνακας Gauss*. Καθώς για τη μετατροπή του A σε άνω τριγωνικό πίνακα, χρησιμοποιούμε την πρώτη γραμμή και στη συνέχεια “κινούμε” προς τα κάτω, κατασκευάζουμε διαδοχικά $\nu - 1$ στοιχειώδεις πίνακες Gauss, $M_1, M_2, \dots, M_{\nu-1}$. Το γινόμενο αυτών των πινάκων,

$$M = M_{\nu-1} M_{\nu-2} \cdots M_2 M_1,$$

είναι ένας $\nu \times \nu$ κάτω τριγωνικός πίνακας με όλα τα διαγώνια στοιχεία του ίσα με 1. Επιπλέον, ο πίνακας

$$U = M A = M_{\nu-1} M_{\nu-2} \cdots M_2 M_1 A$$

είναι άνω τριγωνικός, σύμφωνα με την απαλοιφή Gauss.

Εύκολα μπορεί κανείς να δει ότι κάθε πίνακας M_k είναι αντιστρέψιμος και

$$M_k^{-1} = \begin{bmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & -r_{k+1}^{(k-1)} & 1 & \cdots & 0 \\ \vdots & \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & -r_{\nu}^{(k-1)} & 0 & \cdots & 1 \end{bmatrix}.$$

Κατά συνέπεια ο πίνακας M είναι αντιστρέψιμος και ο $L = M^{-1}$ είναι κάτω τριγωνικός με όλα τα διαγώνια στοιχεία του ίσα με 1. Τελικά λοιπόν ο (αντιστρέψιμος) πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ του συστήματος γράφεται

$$A = LU,$$

όπου ο πίνακας L είναι κάτω τριγωνικός με όλα τα διαγώνια στοιχεία του ίσα με 1, και ο πίνακας U είναι άνω τριγωνικός με μη μηδενικά διαγώνια στοιχεία γνωστά ως οδηγία στοιχεία (*pivots*) της απαλοιφής Gauss.

Στο εισαγωγικό παράδειγμα της προηγούμενης παραγράφου, μπορεί κανείς να επαληθεύσει ότι

$$\begin{aligned} M_1 &= \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix}, \quad M_2 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}, \\ M &= M_2 M_1 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ -4 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ -2 & 1 & 0 \\ 0 & -2 & 1 \end{bmatrix}, \\ L &= M^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ 4 & 2 & 1 \end{bmatrix} \quad \text{και} \quad U = \begin{bmatrix} 1 & 1 & 1 \\ 0 & -3 & -3 \\ 0 & 0 & 3 \end{bmatrix}. \end{aligned}$$

Προφανώς ισχύει

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

2.3 Κριτήριο Παραγοντοποίησης LU και Αλγόριθμοι

Η παραγοντοποίηση LU ενός τετραγωνικού πίνακα A είναι δυνατή όταν δεν παρουσιάζονται μηδενικά οδηγία στοιχεία. Η συνθήκη αυτή είναι ισοδύναμη με την

παρουσία μη μηδενικών κυρίων υποοριζουσών του A , όπως αποδεικνύεται και στο θεώρημα που ακολουθεί.

Θεώρημα 2.3.1. Ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ έχει παραγοντοποίηση LU όταν για κάθε $k = 1, 2, \dots, \nu-1$, η κύρια υποορίζουσα $\det(A(1:k, 1:k))$, που σχηματίζεται από τις k πρώτες γραμμές και τις k πρώτες στήλες του A , είναι μη μηδενική. Επιπλέον, αν η παραγοντοποίηση LU υπάρχει και ο A είναι αντιστρέψιμος, τότε η παραγοντοποίηση LU είναι μοναδική.

Απόδειξη. Ας υποθέσουμε ότι έχουν ολοκληρωθεί τα πρώτα $k-1$ βήματα της απαλοιφής Gauss, δηλαδή ότι ο πίνακας A έχει μετασχηματισθεί στον πίνακα

$$A^{(k-1)} = M_{k-1} \cdots M_1 A$$

με $a_{kk}^{(k-1)}$ το k -οστό οδηγό στοιχείο του. Καθώς κατά την απαλοιφή Gauss (που εμείς εφαρμόζουμε) δεν αντιμετωπίζουμε γραμμές ούτε πολλαπλασιάζουμε γραμμές με αριθμούς, από τις ιδιότητες των οριζουσών έχουμε ότι

$$\begin{aligned} \det(A(1:k, 1:k)) &= \det(A^{(k-1)}(1:k, 1:k)) \\ &= a_{11}^{(1)} a_{22}^{(1)} \cdots a_{k-1,k-1}^{(k-2)} a_{kk}^{(k-1)}. \end{aligned}$$

Αν λοιπόν $\det(A(1:k, 1:k)) \neq 0$, τότε $a_{kk}^{(k-1)} \neq 0$ και η απαλοιφή Gauss προς κατασκευή της παραγοντοποίησης LU μπορεί να συνεχιστεί.

Για τη μοναδικότητα της παραγοντοποίησης LU , υποθέτουμε ότι

$$A = L_1 U_1 = L_2 U_2.$$

Εφόσον ο A είναι αντιστρέψιμος, οι πίνακες L_1, L_2, U_1, U_2 είναι επίσης αντιστρέψιμοι. Κατά συνέπεια,

$$L_2^{-1} L_1 = U_2 U_1^{-1},$$

όπου ο πίνακας $L_2^{-1} L_1$ είναι κάτω τριγωνικός με όλα τα διαγώνια στοιχεία του ίσα με 1 και ο $U_2 U_1^{-1}$ είναι άνω τριγωνικός. Επομένως, καταλήγουμε στο ότι $L_2^{-1} L_1 = U_2 U_1^{-1} = I_\nu$, ή ισοδύναμα, στο ότι $L_1 = L_2$ και $U_1 = U_2$. $\square$

Οι συνθήκες του παραπάνω θεωρήματος είναι ικανές αλλά όχι αναγκαίες. Για παράδειγμα,

$$\begin{bmatrix} 0 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 0 & 1 \end{bmatrix}.$$

Στη συνέχεια, παρουσιάζουμε δύο απλούς αλγόριθμους με εντολές τύπου Matlab, έναν για την άνω τριγωνοποίηση ενός $\nu \times \nu$ πραγματικού πίνακα A μέσω της απαλοιφής Gauss και έναν για την παραγοντοποίηση LU .

ΑΛΓΟΡΙΘΜΟΣ 1.1 (Ανω τριγωνοποίηση)

 $k = 1$

```

while    ( $A(k, k) \neq 0$ ) & ( $k \leq \nu - 1$ )
     $r(k + 1 : \nu) = -A(k + 1 : \nu, k)/A(k, k)$ 
     $A(k + 1 : \nu, :) = A(k + 1 : \nu, :) + r(k + 1 : \nu)' \cdot A(k, :)$ 
     $k = k + 1$ 
end

```

ΑΛΓΟΡΙΘΜΟΣ 1.2 (Παραγοντοποίηση LU)
 $L = I_\nu, U = A, k = 1$

```

while    ( $A(k, k) \neq 0$ ) & ( $k \leq \nu - 1$ )
     $r(k + 1 : \nu) = -U(k + 1 : \nu, k)/U(k, k)$ 
     $U(k + 1 : \nu, :) = U(k + 1 : \nu, :) + r(k + 1 : \nu)' \cdot U(k, :)$ 
     $M = I_\nu$ 
     $M(k + 1 : \nu, k) = -r(k + 1 : \nu)'$ 
     $L = L \cdot M$ 
     $k = k + 1$ 
end

```

Ο “τόνος” που εμφανίζεται στο $r(k + 1 : \nu)$ δηλώνει ότι το Matlab λαμβάνει το συγκεκριμένο διάνυσμα ως στήλη (δηλαδή, δηλώνει αναστροφή γραμμής).

2.4 Εφαρμογές και Σχόλια

Αν ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ έχει παραγοντοποιηθεί στη μορφή LU , $A = LU$, τότε κάθε γραμμικό σύστημα

$$Ax = \beta \Leftrightarrow LUx = \beta$$

επιλύεται εύκολα σε δύο βήματα. Αρχικά λύνουμε το κάτω τριγωνικό σύστημα $Ly = \beta$, και στη συνέχεια, το άνω τριγωνικό σύστημα $Ux = y$.

Εύκολα αντιλαμβάνεται κανείς ότι οι υπολογιστικές απαιτήσεις της παραγοντοποίησης LU και της απαλοιφής Gauss ταυτίζονται. Το κέρδος της παραγοντοποίησης LU είναι σημαντικό όταν πρέπει να λύσουμε ταυτόχρονα περισσότερα από ένα

συστήματα με τον ίδιο πίνακα A . Για παράδειγμα, όταν θέλουμε να υπολογίσουμε τον A^{-1} λύνοντας την εξίσωση

$$AX = I_\nu \Leftrightarrow [AX_1 \ AX_2 \ \cdots \ AX_\nu] = [e_1 \ e_2 \ \cdots \ e_\nu],$$

όπου $X_1, X_2, \dots, X_\nu$ είναι οι στήλες του X και $e_1, e_2, \dots, e_\nu$ τα διανύσματα της κανονικής βάσης (δηλαδή, οι στήλες του I_ν).

Αν θεωρήσουμε τον διαγώνιο πίνακα $D = \text{diag}\{u_{11}, u_{22}, \dots, u_{\nu\nu}\}$, με διαγώνια στοιχεία τα (μη μηδενικά) οδηγία στοιχεία της απαλοιφής Gauss, τότε

$$A = LU = LDD^{-1}U = LDM^*,$$

όπου ο πίνακας $M^* = D^{-1}U$ είναι άνω τριγωνικός. Δηλαδή, κάθε αντιστρέψιμος πίνακας A που έχει παραγοντοποίηση LU , παραγοντοποιείται και στη μορφή $A = LDM^*$, όπου οι πίνακες L και M είναι κάτω τριγωνικοί με όλα τα διαγώνια στοιχεία τους ίσα με 1 και ο πίνακας D είναι διαγώνιος με $\det(D) = \det(A)$. Από το Θεώρημα 2.3.1, οι πίνακες L, M, D είναι μοναδικοί.

Ας υποθέσουμε τώρα ότι ο $A \in \mathbb{C}^{\nu \times \nu}$ είναι ερμιτιανός. Τότε ο πίνακας

$$M^{-1}A(M^*)^{-1} = M^{-1}LD$$

είναι ερμιτιανός και τριγωνικός, άρα διαγώνιος. Αφού ο διαγώνιος πίνακας D είναι αντιστρέψιμος, και ο πίνακας $M^{-1}L$ είναι διαγώνιος. Παρατηρούμε ακόμη ότι οι πίνακες M^{-1} και L είναι κάτω τριγωνικοί με όλα τα διαγώνια στοιχεία τους ίσα με 1. Συνεπώς, $M^{-1}L = I_\nu$, ή ισοδύναμα, $M = L$. Δηλαδή, κάθε αντιστρέψιμος ερμιτιανός πίνακας A γράφεται

$$A = LDL^* \quad (\text{παραγοντοποίηση Cholesky}),$$

όπου ο L είναι κάτω τριγωνικός με τα όλα τα διαγώνια στοιχεία του ίσα με 1 και ο D είναι διαγώνιος με $\det(D) = \det(A)$.

Γνωρίζουμε ότι όταν κατά την απαλοιφή Gauss εμφανίζονται μηδενικά διαγώνια στοιχεία (μη αποδεκτά για οδηγία στοιχεία), μπορούμε να συνεχίσουμε τη διαδικασία αντιμετωπίζοντας δύο γραμμές του συστήματος. Ακόμη, χωρίς να δημιουργείται πρόβλημα στην απαλοιφή Gauss, μπορούμε να θεωρήσουμε ότι πραγματοποιούμε όλες τις αντιμεταθέσεις γραμμών από την αρχή, πριν την διαδικασία της απαλοιφής. Αν συμβολίσουμε με P τον πίνακα μετάθεσης που προκύπτει από την εφαρμογή των απαιτούμενων αντιμεταθέσεων επί του μοναδιαίου πίνακα I_ν , τότε PA είναι ο πίνακας που προκύπτει αν οι συγκεκριμένες αντιμεταθέσεις εφαρμοστούν επί του A . Για παράδειγμα, η αντιμετάθεση πρώτης και δεύτερης γραμμής γράφεται

$$\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 & 2 \\ 3 & 4 & 5 \\ 6 & 7 & 8 \end{bmatrix} = \begin{bmatrix} 3 & 4 & 5 \\ 0 & 1 & 2 \\ 6 & 7 & 8 \end{bmatrix}.$$

Αν ο πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ είναι αντιστρέψιμος, τότε είναι γνωστό ότι η απαλοιφή Gauss, με χρήση και αντιμεταθέσεων γραμμών, οδηγεί πάντα σε τριγωνοποίηση του A . Αυτό σημαίνει ότι αν ο A είναι αντιστρέψιμος, τότε υπάρχει πίνακας μετάθεσης P τέτοιος ώστε ο PA να έχει παραγοντοποίηση LU , δηλαδή

$$PA = LU.$$

Για παράδειγμα,

$$\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 3/4 & 1 & 0 & 0 \\ 1/2 & -2/7 & 1 & 0 \\ 1/4 & -3/7 & 5/8 & 1 \end{bmatrix} \begin{bmatrix} 8 & 7 & 9 & 5 \\ 0 & 7/4 & 5/4 & 21/4 \\ 0 & 0 & -8/7 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Αν $P \neq I_\nu$, δηλαδή αν απαιτηθούν αντιμεταθέσεις γραμμών, τότε λέμε ότι έχουμε μερική οδήγηση στην απαλοιφή Gauss.

Αν ο βαθμός ενός $A \in \mathbb{C}^{\nu \times \nu}$ είναι $\kappa < \nu$ (δηλαδή, ο A δεν είναι αντιστρέψιμος), τότε για κάποιον πίνακα μετάθεσης P , ο πίνακας PA έχει παραγοντοποίηση LU , $PA = LU$, όπου οι $\nu - \kappa$ τελευταίες γραμμές του U μπορούν να είναι μη-

δενικές. Για παράδειγμα, αν $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{bmatrix}$, τότε ο βαθμός του πίνακα είναι

$\text{rank}(A) = 2$ και ισχύει

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 1/4 & 1 & 0 & 0 \\ 3/4 & 1/3 & 1 & 0 \\ 1/2 & 2/3 & 0 & 1 \end{bmatrix} \begin{bmatrix} 4 & 5 & 6 & 7 \\ 0 & 3/4 & 3/2 & 9/4 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

2.5 Ασκήσεις

1. Χρησιμοποιώντας στοιχειώδεις πίνακες Gauss (χωρίς οδήγηση) παραγοντο-

ποιήστε στη μορφή LU τον πίνακα $A = \begin{bmatrix} 2 & 1 & 1 & 0 \\ 4 & 3 & 3 & 1 \\ 8 & 7 & 9 & 5 \\ 6 & 7 & 9 & 8 \end{bmatrix}$.

2. Χρησιμοποιώντας την παραγοντοποίηση LU του πίνακα A της προηγούμενης άσκησης και λύνοντας την εξίσωση

$$AX = I_\nu \Leftrightarrow A[X_1 \ X_2 \ \cdots \ X_\nu] = [e_1 \ e_2 \ \cdots \ e_\nu],$$

υπολογίστε τον A^{-1} .

3. Κατασκευάστε παραλλαγή του Αλγορίθμου 1.2 για παραγοντοποίηση στη μορφή LU ενός $\nu \times \mu$ ($\nu \neq \mu$) πίνακα A . Επαληθεύστε τον αλγόριθμο σας στους πίνακες $\begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \\ 4 & 5 & 7 \end{bmatrix}$ και $\begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 7 \end{bmatrix}$.
4. Υποθέτοντας ότι ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ έχει παραγοντοποίηση LU , να αποδείξετε ότι το σύστημα $Ax = \beta$ επιλύεται με παραγοντοποίηση LU του επαυξημένου πίνακα $[A \ \beta]$.
5. Εφαρμόζοντας την απαλοιφή Gauss στον πίνακα $A \in \mathbb{C}^{\nu \times \nu}$ με πράξεις στηλών (αντί γραμμών) να παραγοντοποιήσετε τον A στη μορφή $A = LU$, όπου ο L είναι κάτω τριγωνικός και ο U είναι άνω τριγωνικός με όλα τα διαγώνια στοιχεία του ίσα με 1.
6. Κατασκευάστε παραλλαγή της απαλοιφής Gauss η οποία να εισάγει μηδενικά στις στήλες του $A \in \mathbb{C}^{\nu \times \nu}$ ξεκινώντας από τη τελευταία στήλη και προχωρώντας προς τα αριστερά. Τι μορφή παραγοντοποίησης δίνει η μέθοδος αυτή;

Κεφάλαιο 3

Παραγοντοποίηση QR

3.1 Ορθοκανονικοποίηση Gram-Schmid

Θεωρούμε έναν πίνακα $A \in \mathbb{C}^{\nu \times \mu}$ ($\nu \geq \mu$), του οποίου οι στήλες $a_1, a_2, \dots, a_\mu$ είναι γραμμικά ανεξάρτητες, δηλαδή ο βαθμός του A είναι $\text{rank}(A) = \mu$. Θέλουμε να κατασκευάσουμε ένα ορθοκανονικό σύστημα διανυσμάτων $q_1, q_2, \dots, q_\mu$ τέτοιο ώστε

$$\text{span}\{a_1, a_2, \dots, a_\mu\} = \text{span}\{q_1, q_2, \dots, q_\mu\}.$$

Η μεθοδολογία που θα περιγράψουμε είναι γνωστή ως *ορθοκανονικοποίηση Gram-Schmidt* και βασίζεται στην εξής παρατήρηση: αν $x, y \in \mathbb{C}^\nu$ είναι δύο γραμμικά ανεξάρτητα διανύσματα και συμβολίσουμε με $u = \frac{y^* x}{y^* y} y$ την προβολή του x επί του y , τότε το διάνυσμα $x - u$ είναι κάθετο στο y , αφού

$$y^*(x - u) = y^* x - \frac{y^* x}{y^* y} y^* y = 0.$$

Αρχικά, κανονικοποιούμε το διάνυσμα a_1 γράφοντας $q_1 = a_1 / \|a_1\|_2$. Έπειτα κατασκευάζουμε ένα ορθογώνιο στο q_1 διάνυσμα $w_2 = a_2 - (q_1^* a_2) q_1$ και το κανονικοποιούμε γράφοντας $q_2 = w_2 / \|w_2\|_2$. Στο τρίτο βήμα της μεθόδου κατασκευάζουμε ένα ορθογώνιο στο $\text{span}\{q_1, q_2\}$ διάνυσμα $w_3 = a_3 - (q_1^* a_3) q_1 - (q_2^* a_3) q_2$ και το κανονικοποιούμε γράφοντας $q_3 = w_3 / \|w_3\|_2$. Συνεχίζοντας με τον τρόπο αυτό, λαμβάνουμε τα διανύσματα

$$\begin{aligned} q_1 &= \frac{a_1}{r_{11}}, \\ q_2 &= \frac{a_2 - r_{12} q_1}{r_{22}}, \end{aligned}$$

$$\begin{aligned}
q_3 &= \frac{a_3 - r_{13} q_1 - r_{23} q_2}{r_{33}}, \\
&\vdots \\
q_\mu &= \frac{a_\mu - \sum_{i=1}^{\mu-1} r_{i\mu} q_i}{r_{\mu\mu}},
\end{aligned}$$

όπου

$$r_{ij} = q_i^* a_j \quad (i \neq j) \quad \text{και} \quad r_{jj} = \left\| a_j - \sum_{i=1}^{j-1} r_{ij} q_i \right\|_2.$$

Ορίζουμε τον $\nu \times \mu$ πίνακα $Q = [q_1 \ q_2 \ \cdots \ q_\mu]$ που οι στήλες του σχηματίζουν ένα ορθοκανονικό σύστημα και τον $\mu \times \mu$ άνω τριγωνικό πίνακα

$$R = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1\mu} \\ 0 & r_{22} & \cdots & r_{2\mu} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & r_{\mu\mu} \end{bmatrix}$$

που έχει στοιχεία τους συντελεστές της ορθοκανονικοποίησης Gram-Schmidt. Τότε η παραγοντοποίηση

$$A = QR \Leftrightarrow [a_1 \ a_2 \ \cdots \ a_\mu] = [q_1 \ q_2 \ \cdots \ q_\mu] \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1\mu} \\ 0 & r_{22} & \cdots & r_{2\mu} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & r_{\mu\mu} \end{bmatrix}$$

είναι άμεση και καλείται (απλή) παραγοντοποίηση QR του πίνακα A .

3.2 Παραγοντοποίηση QR : Υπαρξη, Μοναδικότητα και Αλγόριθμος

Όπως είδαμε στην προηγούμενη παράγραφο, η (απλή) παραγοντοποίηση QR προκύπτει άμεσα από την ορθοκανονικοποίηση Gram-Schmidt. Στην πρόταση που ακολουθεί περιγράφεται ένα κριτήριο ύπαρξης και μοναδικότητας της συγκεκριμένης παραγοντοποίησης.

Πρόταση 3.2.1. Κάθε πίνακας $A \in \mathbb{C}^{\nu \times \mu}$ ($\nu \geq \mu$) βαθμού $\text{rank}(A) = \mu$ έχει μοναδική παραγοντοποίηση QR .

Απόδειξη. Προφανώς, ο πίνακας A έχει παραγοντοποίηση QR αν και μόνο αν η ορθοκανονικοποίηση Gram Schmidt ολοκληρώνεται με επιτυχία. Η ορθοκανονικοποίηση δεν ολοκληρώνεται με επιτυχία μόνο στην περίπτωση που εμφανιστεί ένα

διάνυσμα $w_j = a_j - \sum_{i=1}^{j-1} r_{ij} q_i$ μηδενικό. Κάτι τέτοιο όμως είναι αδύνατο αφού

$$\text{rank}(A) = \dim(\text{span}\{a_1, a_2, \dots, a_\mu\}) = \mu.$$

Η μοναδικότητα της παραγοντοποίησης QR είναι επίσης προφανής διότι κατά την ορθοκανονικοποίηση Gram-Schmidt, αφού $\text{rank}(A) = \mu$, τα διανύσματα q_i και οι συντελεστές r_{ij} ορίζονται κατά μοναδικό τρόπο. $\square$

Με βάση την ορθοκανονικοποίηση Gram-Schmidt, για $A \in \mathbb{R}^{\nu \times \mu}$, μπορούμε άμεσα να σχεδιάσουμε έναν αλγόριθμο κατασκευής των πινάκων Q και R .

ΑΛΓΟΡΙΘΜΟΣ 2.1 (Παραγοντοποίηση QR)

$$R(1, 1) = \|A(:, 1)\|_2, Q(:, 1) = A(:, 1)/R(1, 1)$$

```

for    k = 2 : μ
    R(1 : k - 1, k) = Q(1 : ν, 1 : k - 1)' · A(1 : ν, k)
    z = A(1 : ν, k) - Q(1 : ν, 1 : k - 1) · R(1 : k - 1, k)
    R(k, k) = \|z\|_2
    Q(1 : ν, k) = z/R(k, k)
end

```

Σημειώνεται ότι στο k -οστό βήμα του αλγορίθμου σχηματίζονται οι k -οστές στήλες του Q και του R , ενώ οι “τόνοι” δηλώνουν πάντα αναστροφή.

3.3 Πίνακες Householder

Έστω ένα μη μηδενικό διάνυσμα $u \in \mathbb{C}^\nu$. Ο $\nu \times \nu$ μιγαδικός πίνακας

$$P = I_\nu - \frac{2}{u^* u} u u^*$$

καλείται πίνακας (αντανάκλασης) *Householder*, ενώ το διάνυσμα u λέγεται διάνυσμα *Householder* που αντιστοιχεί στον P . Για κάθε διάνυσμα $w \in \mathbb{C}^\nu$, το διάνυσμα

$$P w = w - \frac{2}{u^* u} u u^* w = w - \frac{2 u^* w}{u^* u} u$$

είναι η αντανάκλαση του w ως προς το υπερεπίπεδο $\text{span}\{u\}^\perp$ (δηλαδή, το ορθογώνιο συμπλήρωμα του $\text{span}\{u\}$). Ο πίνακας Householder P είναι προφανώς ερμιτιανός και ορθομοναδιαίος (δηλαδή, $P = P^*$ και $P^* P = P P^* = I_\nu$). Είναι επίσης αξιοσημείωτο ότι αποτελεί μία διαταραχή του μοναδιαίου πίνακα I_ν , βαθμού

$$\text{rank}\left(\frac{2}{u^* u} u u^*\right) = \text{rank}(u u^*) = 1.$$

Ας υποθέσουμε τώρα πως θέλουμε το Pw να είναι πολλαπλάσιο του διανύσματος e_1 της κανονικής βάσης. Τότε $Pw \in \text{span}\{e_1\}$ και $u \in \text{span}\{w, e_1\}$ (τα διανύσματα u , w και e_1 είναι συνεπίπεδα). Γράφουμε αυθαίρετα $u = w + a e_1$ ($a \in \mathbb{C}$) και παρατηρούμε ότι

$$u^*w = w^*w + \bar{a} e_1^T w = w^*w + \bar{a} w_1,$$

όπου w_1 είναι το πρώτο στοιχείο του w το οποίο, χωρίς βλάβη της γενικότητας, μπορεί να θεωρηθεί πραγματικός αριθμός, ενώ

$$u^*u = (w + a e_1)^*(w + a e_1) = w^*w + 2 \text{Re}(a) w_1 + |a|^2.$$

Επομένως,

$$Pw = \left(I_\nu - \frac{2}{u^*u} u u^* \right) w = \left(1 - 2 \frac{w^*w + \bar{a} w_1}{w^*w + 2 w_1 \text{Re}(a) + |a|^2} \right) w - 2a \frac{u^*w}{u^*u} e_1.$$

Ο συντελεστής του w μηδενίζεται αν θέτουμε $a = e^{i \arg(w_1)} \|w\|_2$ και τότε

$$u = w + e^{i \arg(w_1)} \|w\|_2 e_1 \Rightarrow Pw = \left(I_\nu - 2 \frac{u^*u}{u^*u} \right) w = -e^{i \arg(w_1)} \|w\|_2 e_1.$$

Δηλαδή, για δεδομένο διάνυσμα w , κατασκευάσαμε το διάνυσμα Housholder u και τον αντίστοιχο πίνακα Housholder $P = I_\nu - \frac{2}{u^*u} u u^*$, έτσι ώστε το διάνυσμα Pw να ανήκει στο $\text{span}\{e_1\}$. Για παράδειγμα, αν

$$w = \begin{bmatrix} 3 \\ 1 \\ 5 \\ 1 \end{bmatrix} \quad \text{με} \quad \|w\|_2 = 6 \quad \text{και} \quad w_1 = 3,$$

τότε για το διάνυσμα Housholder

$$u = w + e^{i0} \|w\|_2 e_1 = w + \|w\|_2 e_1 = \begin{bmatrix} 9 \\ 1 \\ 5 \\ 1 \end{bmatrix},$$

ο αντίστοιχος πίνακας Housholder

$$P = I_\nu - 2 \frac{u u^T}{u^T u} = \frac{1}{54} \begin{bmatrix} -27 & -9 & -45 & -9 \\ -9 & 53 & -5 & -1 \\ -45 & -5 & 29 & -5 \\ -9 & -1 & -5 & 53 \end{bmatrix}$$

ικανοποιεί τη σχέση

$$Pw = \begin{bmatrix} -6 \\ 0 \\ 0 \\ 0 \end{bmatrix} \in \text{span}\{e_1\}.$$

3.4 Πλήρης Παραγοντοποίηση QR

Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \mu}$ ($\nu \geq \mu$) βαθμού $\text{rank}(A) = \mu$, δηλαδή οι στήλες του A , $a_1, a_2, \dots, a_\mu$, είναι γραμμικά ανεξάρτητες. Αν P_1 είναι ένας πίνακας Housholder τέτοιος ώστε $P_1 a_1 \in \text{span}\{e_1\} \subset \mathbb{C}^\nu$, τότε

$$\begin{aligned} P_1 A &= P_1 [a_1 \ a_2 \ \cdots \ a_\mu] \\ &= [P_1 a_1 \ P_1 a_2 \ \cdots \ P_1 a_\mu] \\ &= \begin{bmatrix} \begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \end{bmatrix} & P_1 a_2 & \cdots & P_1 a_\mu \end{bmatrix}. \end{aligned}$$

Λόγω της γραμμικής ανεξαρτησίας των στηλών του A , $P_1 a_2 \notin \text{span}\{e_1\}$, δηλαδή η στήλη $P_1 a_2$ έχει μη μηδενικά στοιχεία κάτω από το πρώτο της στοιχείο. Αν ορίσουμε $w_2 \in \mathbb{C}^{\nu-1}$ το μη μηδενικό διάνυσμα που προκύπτει με τη διαγραφή του πρώτου στοιχείου του διανυσματος στήλη $P_1 a_2$, τότε υπάρχει $(\nu-1) \times (\nu-1)$ πίνακας Housholder P_2 τέτοιος ώστε το $P_2(P_1 a_2)$ να είναι πολλαπλάσιο του $e_1 \in \mathbb{C}^{\nu-1}$. Συνεπώς, έχουμε

$$\begin{bmatrix} 1 & 0 \\ 0 & P_2 \end{bmatrix} P_1 a_2 = \begin{bmatrix} * \\ P_2 w_2 \end{bmatrix} = \begin{bmatrix} * \\ * \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

και

$$\begin{bmatrix} 1 & 0 \\ 0 & P_2 \end{bmatrix} P_1 A = \begin{bmatrix} * & * & * & \cdots & * \\ 0 & * & * & \cdots & * \\ 0 & 0 & * & \cdots & * \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & * & \cdots & * \end{bmatrix}.$$

Είναι φανερό ότι ο πίνακας $\begin{bmatrix} 1 & 0 \\ 0 & P_2 \end{bmatrix}$ είναι ορθομοναδιαίος. Χωρίς βλάβη της γενικότητας, μπορούμε να θεωρήσουμε ότι $\nu > \mu$. Επαναλαμβάνοντας την ίδια διαδικασία, μπορούμε να κατασκευάσουμε Housholder πίνακες $P_3 \in \mathbb{C}^{(\nu-2) \times (\nu-2)}$, $P_4 \in \mathbb{C}^{(\nu-3) \times (\nu-3)}$, $\dots$, $P_\mu \in \mathbb{C}^{(\nu-\mu+1) \times (\nu-\mu+1)}$ και $\nu \times \nu$ ορθομοναδιαίους πίνακες

$$\begin{bmatrix} I_2 & 0 \\ 0 & P_3 \end{bmatrix}, \begin{bmatrix} I_3 & 0 \\ 0 & P_4 \end{bmatrix}, \dots, \begin{bmatrix} I_{\mu-1} & 0 \\ 0 & P_\mu \end{bmatrix}$$

τέτοιους ώστε

$$\left(\begin{bmatrix} I_{\mu-1} & 0 \\ 0 & P_\mu \end{bmatrix} \cdots \begin{bmatrix} I_2 & 0 \\ 0 & P_3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & P_2 \end{bmatrix} P_1 \right) A = \begin{bmatrix} * & * & \dots & * & * \\ 0 & * & \dots & * & * \\ 0 & 0 & \dots & * & * \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & * \\ 0 & 0 & \dots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 0 & 0 \end{bmatrix}. \quad (3.4.1)$$

Θεωρώντας τον ορθομοναδιαίο πίνακα

$$\tilde{Q} = \begin{bmatrix} I_{\mu-1} & 0 \\ 0 & P_\mu \end{bmatrix} \cdots \begin{bmatrix} I_2 & 0 \\ 0 & P_3 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & P_2 \end{bmatrix} P_1$$

και τον πίνακα $\hat{R}$ στο δεξιό μέλος της (3.4.1), ο οποίος είναι “άνω τριγωνικός” με $\nu - \mu$ μηδενικές γραμμές στο κάτω μέρος, έχουμε

$$\tilde{Q} A = \hat{R} \Leftrightarrow A = \hat{Q} \hat{R}, \quad \text{όπου } \hat{Q} = \tilde{Q}^*.$$

Δηλαδή, με τη σταδιακή τριγωνποίηση του A που περιγράψαμε, οι πίνακες Housholder δίνουν μία παραγοντοποίηση του πίνακα A “τύπου QR ”. Συγκεκριμένα, έχουμε την παραγοντοποίηση

$$A = \hat{Q} \hat{R},$$

όπου εδώ ο $\hat{Q}$ είναι $\nu \times \nu$ ορθομοναδιαίος και ο $\hat{R}$ είναι $\nu \times \mu$ “άνω τριγωνικός”. Η παραγοντοποίηση αυτή καλείται *πλήρης παραγοντοποίηση QR* του A και έχει συγκεκριμένα πλεονεκτήματα. Για παράδειγμα, αν ο πίνακας $A \in \mathbb{C}^{\nu \times \mu}$ ($\nu > \mu$) είναι βαθμού $r < \mu$, τότε ο $\hat{R}$ έχει μηδενικές ακριβώς τις $\nu - \mu$ τελευταίες γραμμές του και $\mu - r$ μηδενικά στην κύρια διαγώνιο. Έτσι, καθώς υπάρχουν ευσταθείς και αποτελεσματικές μεθοδολογίες για την πλήρη παραγοντοποίηση QR ενός πίνακα, η παραγοντοποίηση αυτή αποτελεί ένα διαδεδομένο τρόπο υπολογισμού του βαθμού πίνακα.

Προφανώς, στην περίπτωση των τετραγωνικών πινάκων η (απλή) παραγοντοποίηση QR και η πλήρης παραγοντοποίηση QR ταυτίζονται. Πρέπει ακόμη να σημειώσουμε ότι, για $\nu > \mu$, η ορθοκανονικοποίηση Gram-Schmidt μπορεί να οδηγήσει σε πλήρη παραγοντοποίηση QR , αρκεί να επεκτείνουμε το ορθοκανονικό σύστημα διανυσμάτων $q_1, q_2, \dots, q_\mu$ σε ορθοκανονική βάση του $\mathbb{C}^\nu$ και να προσθέσουμε $\nu - \mu$ μηδενικές γραμμές στο κάτω μέρος του πίνακα R . Στην περίπτωση αυτή βέβαια χάνεται η μοναδικότητα της Πρότασης 3.2.1.

3.5 Εφαρμογές και Σχόλια

Αν ένας αντιστρέψιμος πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ έχει παραγοντοποιηθεί στη μορφή QR , $A = QR$, τότε κάθε σύστημα $Ax = \beta \Leftrightarrow QRx = \beta$ επιλύεται εύκολα σε δύο βήματα. Αρχικά, θεωρούμε το σύστημα $Qy = \beta$ που έχει λύση $y = Q^*\beta$, και στη συνέχεια λύνουμε το τριγωνικό σύστημα $Rx = y$. Επιπλέον, ο αντίστροφος του A , $A^{-1} = (QR)^{-1} = R^{-1}Q^*$, είναι άμεσα υπολογίσιμος αφού ο R^{-1} προκύπτει από την επίλυση της εξίσωσης

$$RR^{-1} = I_\nu \Leftrightarrow \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1\nu} \\ 0 & r_{22} & \cdots & r_{2\nu} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & r_{\nu\nu} \end{bmatrix} \begin{bmatrix} 1/r_{11} & x_{12} & \cdots & x_{1\nu} \\ 0 & 1/r_{22} & \cdots & x_{2\nu} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \cdots & 1/r_{\nu\nu} \end{bmatrix} = I_\nu$$

$$\Leftrightarrow \begin{cases} r_{11}x_{12} + \frac{r_{12}}{r_{22}} = 0 & \Leftrightarrow x_{12} = -\frac{r_{12}}{r_{11}r_{22}} \\ r_{22}x_{23} + \frac{r_{23}}{r_{33}} = 0 & \Leftrightarrow x_{23} = -\frac{r_{23}}{r_{22}r_{33}} \\ \vdots & \vdots \\ r_{\nu-1,\nu-1}x_{\nu-1,\nu} + \frac{r_{\nu,\nu-1}}{r_{\nu\nu}} = 0 & \Leftrightarrow x_{\nu-1,\nu} = -\frac{r_{\nu-1,\nu}}{r_{\nu-1,\nu-1}r_{\nu\nu}} \\ r_{11}x_{13} + r_{12}x_{23} + \frac{r_{13}}{r_{33}} = 0 & \Leftrightarrow x_{13} = -\frac{r_{13}}{r_{11}r_{33}} + \frac{r_{12}r_{23}}{r_{11}r_{22}r_{33}} \\ \vdots & \vdots \end{cases}.$$

Το πρόβλημα των ελαχίστων τετραγώνων διατυπώνεται (για τη 2-νόρμα) ως εξής: δεδομένου ενός πίνακα $A \in \mathbb{C}^{\nu \times \mu}$ ($\nu \geq \mu$) και ενός διανύσματος $\beta \in \mathbb{C}^\nu$, να υπολογιστεί ένα $x \in \mathbb{C}^\mu$ τέτοιο ώστε η απόσταση $\|\beta - Ax\|_2$ να είναι η ελάχιστη δυνατή. Όταν $\text{rank}(A) = \mu$, μία δημοφιλής μέθοδος επίλυσης του προβλήματος αυτού είναι μέσω της (απλής) παραγοντοποίησης QR του πίνακα A , $A = QR$. Συγκεκριμένα, αν $q_1, q_2, \dots, q_\mu$ είναι οι στήλες του Q , τότε μπορεί κανείς να επαληθεύσει ότι ο πίνακας $P = QQ^*$ είναι πίνακας ορθογώνιας προβολής στον υπόχωρο $\text{span}\{q_1, q_2, \dots, q_\mu\}$, δηλαδή

$$P^2 = P, \quad P^* = P \quad \text{και} \quad Pq = q \quad \text{για κάθε} \quad q \in \text{span}\{q_1, q_2, \dots, q_\mu\}.$$

Συνεπώς, το διάνυσμα $y = P\beta$ ανήκει στον υπόχωρο που παράγουν οι στήλες του A , δηλαδή στην εικόνα του A , $\text{Range}(A)$. Αυτό σημαίνει ότι το σύστημα $Ax = y$ έχει ακριβή λύση, και γράφοντας

$$QRx = QQ^*\beta,$$

έχουμε

$$Rx = Q^*\beta.$$

Αφού ο $\mu \times \mu$ άνω τριγωνικός πίνακας R είναι αντιστρέψιμος, η λύση του παραπάνω προβλήματος ελαχίστων τετραγώνων είναι $x = R^{-1}Q^*\beta$.

3.6 Ασκήσεις

1. Έστω οι πίνακες $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$ και $B = \begin{bmatrix} 1 & 2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$. Κατασκευάστε μία απλή και μία πλήρη παραγοντοποίηση QR για καθέναν από τους δύο αυτούς πίνακες.
2. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \mu}$ ($\nu \geq \mu$) με απλή παραγοντοποίηση QR , $A = Q R$. Να αποδείξετε ότι $\text{rank}(A) = \mu$ αν και μόνο αν όλα τα διαγώνια στοιχεία του R είναι μη μηδενικά.
3. Με τη μέθοδο παραγοντοποίησης QR , να λυθεί το γραμμικό σύστημα

$$\begin{cases} x + y - z = 0 \\ x = 1 \\ x - 2z = 3 \\ x + y + z = -2 \end{cases}.$$

4. Υπολογίστε τις ιδιοτιμές και την ορίζουσα ενός $\nu \times \nu$ πίνακα Housholder $P = I_\nu - \frac{2}{u^* u} u u^*$.
5. Έστω $x, y \in \mathbb{C}^\nu$ μη μηδενικά διανύσματα. Κατασκευάστε έναν πίνακα Housholder P τέτοιον ώστε το διάνυσμα Px να είναι πολλαπλάσιο του y .
6. Με χρήση πινάκων Housholder, να αποδείξετε ότι για τυχαία διανύσματα $x, y \in \mathbb{C}^\nu$, ισχύει $\det(I_\nu + x y^*) = 1 + y^* x$.
7. Έστω οι $\nu \times \nu$ πίνακες A και $D = \text{diag}\{d_1, d_2, \dots, d_\nu\}$. Πως μπορεί κανείς να κατασκευάσει ορθομοναδιαίο πίνακα Q τέτοιον ώστε $Q^* A - D Q^* = R$, όπου ο R είναι ένας άνω τριγωνικός πίνακας;
8. Κατασκευάστε έναν αλγόριθμο για την πλήρη παραγοντοποίηση QR ενός $\nu \times \mu$ ($\nu \geq \mu$) μιγαδικού πίνακα με χρήση πινάκων Housholder.

Κεφάλαιο 4

Παραγοντοποίηση SVD

4.1 Η Παραγοντοποίηση SVD

Η σημαντικότερη, ως προς τις εφαρμογές, παραγοντοποίηση πινάκων είναι η γνωστή παραγοντοποίηση ιδιζουσών τιμών (*singular value decomposition*), την οποία και θα καλούμε παραγοντοποίηση SVD .

Θεώρημα 4.1.1. Αν A είναι ένας $\nu \times \mu$ μιγαδικός πίνακας, τότε υπάρχουν ορθομοναδιαίοι πίνακες $U \in \mathbb{C}^{\nu \times \nu}$ και $V \in \mathbb{C}^{\mu \times \mu}$ τέτοιοι ώστε

$$A = U \operatorname{diag}\{s_1, s_2, \dots, s_{\min\{\nu, \mu\}}\} V^*$$

όπου $\operatorname{diag}\{s_1, s_2, \dots, s_{\min\{\nu, \mu\}}\} \in \mathbb{C}^{\nu \times \mu}$ και $s_1 \geq s_2 \geq \dots \geq s_{\min\{\nu, \mu\}} \geq 0$.

Απόδειξη. Αρχικά, θεωρούμε μοναδιαία διανύσματα $x \in \mathbb{C}^\mu$ και $y \in \mathbb{C}^\nu$ τέτοια ώστε $Ax = \|A\|_2 y$ (η απόδειξη της ύπαρξης βασίζεται στον ορισμό της επαγόμενης νόρμας πινάκων και αφήνεται ως άσκηση). Τότε υπάρχουν πίνακες $V_2 \in \mathbb{C}^{\mu \times (\mu-1)}$ και $U_2 \in \mathbb{C}^{\nu \times (\nu-1)}$ τέτοιοι ώστε οι $V = [x \ V_2] \in \mathbb{C}^{\mu \times \mu}$ και $U = [y \ U_2] \in \mathbb{C}^{\nu \times \nu}$ να είναι ορθομοναδιαίοι. Εύκολα μπορεί κανείς να δει ότι

$$U^* A V = \begin{bmatrix} y^* A x & w^* \\ 0 & B \end{bmatrix} = \begin{bmatrix} \|A\|_2 & w^* \\ 0 & B \end{bmatrix}.$$

Παρατηρούμε ακόμη ότι

$$\begin{aligned} (\|A\|_2^2 + w^* w)^2 &\leq \left\| \begin{bmatrix} \|A\|_2 & w^* \\ 0 & B \end{bmatrix} \begin{bmatrix} \|A\|_2 \\ w \end{bmatrix} \right\|_2^2 \\ &\leq \left\| \begin{bmatrix} \|A\|_2 & w^* \\ 0 & B \end{bmatrix} \right\|_2^2 (\|A\|_2^2 + \|w\|_2^2), \end{aligned}$$

από το οποίο προκύπτει ότι

$$\left\| \begin{bmatrix} \|A\|_2 & w^* \\ 0 & B \end{bmatrix} \right\|_2^2 \geq \|A\|_2^2 + \|w\|_2^2.$$

Όμως, επειδή οι πίνακες U και V είναι ορθομοναδιαίοι, έχουμε ότι

$$\left\| \begin{bmatrix} \|A\|_2 & w^* \\ 0 & B \end{bmatrix} \right\|_2^2 = \|A\|_2^2,$$

και για το λόγο αυτό πρέπει $w = 0$. Δηλαδή, ισχύει

$$U^* A V = \begin{bmatrix} \|A\|_2 & 0 \\ 0 & B \end{bmatrix}.$$

Η απόδειξη ολοκληρώνεται με επαγωγή. □

Ο διαγώνιος πίνακας $\text{diag}\{s_1, s_2, \dots, s_{\min\{\nu, \mu\}}\}$ συμβολίζεται με Σ και οι τιμές

$$s_1 \geq s_2 \geq \dots \geq s_{\min\{\nu, \mu\}} \geq 0$$

καλούνται *ιδιάζουσες τιμές* του A . Αν γράψουμε $U = [u_1 \ u_2 \ \dots \ u_\nu]$ και $V = [v_1 \ v_2 \ \dots \ v_\mu]$, τότε τα διανύσματα στήλες u_i και v_i ($1 \leq i \leq \min\{\nu, \mu\}$) λέγονται *αριστερά και δεξιά ιδιάζοντα διανύσματα* του A που αντιστοιχούν στην ιδιάζουσα τιμή s_i . Από τις σχέσεις $AV = U\Sigma$ και $A^*U = V\Sigma^T$ προκύπτει ότι

$$A v_i = s_i u_i \text{ και } A^* u_i = s_i v_i, \quad \forall i = 1, 2, \dots, \min\{\nu, \mu\}.$$

Ένα απλό παράδειγμα παραγοντοποίησης SVD είναι

$$A = \begin{bmatrix} 0.96 & 1.72 \\ 2.28 & 0.96 \end{bmatrix} = U \Sigma V^* = \begin{bmatrix} 0.6 & -0.8 \\ 0.8 & 0.6 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0.8 & 0.6 \\ 0.6 & -0.8 \end{bmatrix}^*.$$

4.2 Άμεσες Παρατηρήσεις και Σχόλια

Η παραγοντοποίηση SVD δίνει πολλές πληροφορίες για τη δομή ενός πίνακα. Αν θεωρήσουμε ότι η ιδιάζουσα τιμή s_r είναι η ελάχιστη μη μηδενική ιδιάζουσα τιμή ενός $\nu \times \mu$ πίνακα A , δηλαδή αν

$$s_1 \geq s_2 \geq \dots \geq s_r > s_{r+1} = \dots = s_{\min\{\nu, \mu\}} = 0,$$

τότε

$$\text{rank}(A) = r,$$

ο πυρήνας του A είναι

$$(\text{Ker}(A) =) \text{Null}(A) = \text{span}\{v_{r+1}, v_{r+2}, \dots, v_\mu\}$$

και η εικόνα του A είναι

$$(\text{Image}(A) =) \text{Range}(A) = \text{span}\{u_1, u_2, \dots, u_r\}.$$

Από την απόδειξη του Θεωρήματος 4.1.1 είναι φανερό ότι $s_1 = s_1(A) = \|A\|_2$, ενώ μπορεί κανείς να δει ότι η (ορθομοναδιαία αναλλοίωτη) νόρμα Frobenius του A , $\|A\|_F = \sqrt{\sum_{i,j} |a_{ij}|^2}$, γράφεται

$$\|A\|_F = \sqrt{s_1^2 + s_2^2 + \dots + s_{\min\{\nu, \mu\}}^2}.$$

Επιπλέον, για $\nu \geq \mu$, ισχύει

$$\min_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2} = s_\mu,$$

ενώ για $\nu = \mu$,

$$|\det(A)| = s_1 s_2 \dots s_\nu.$$

Από τις σχέσεις

$$AA^* = U \text{diag}\{s_1^2, s_2^2, \dots, s_{\min\{\nu, \mu\}}^2\} U^*$$

και

$$A^*A = V \text{diag}\{s_1^2, s_2^2, \dots, s_{\min\{\nu, \mu\}}^2\} V^*,$$

συμπεραίνουμε ότι:

1. Οι μη μηδενικές ιδιάζουσες τιμές του A είναι οι τετραγωνικές ρίζες των μη μηδενικών ιδιοτιμών των (θετικά ημιορισμένων) ερμιτιανών πινάκων AA^* και A^*A .
2. Τα αριστερά ιδιάζοντα διανύσματα του A είναι τα ιδιοδιανύσματα του AA^* .
3. Τα δεξιά ιδιάζοντα διανύσματα του A είναι τα ιδιοδιανύσματα του A^*A .

Κάθε $\nu \times \nu$ μιγαδικός πίνακας A γράφεται

$$A = U \Sigma V^* = U \Sigma U^* U V^* = (U \Sigma U^*) (U V^*),$$

όπου ο πίνακας $P = U \Sigma U^*$ είναι ερμιτιανός και θετικά ημιορισμένος, ενώ ο $W = U V^*$ είναι ορθομοναδιαίος. Δηλαδή, κάθε $\nu \times \nu$ πίνακας A γράφεται ως γινόμενο $A = PW$ ενός θετικά ημιορισμένου πίνακα P με $\text{rank}(P) = \text{rank}(A)$ και ενός ορθομοναδιαίου πίνακα W . Η παραγοντοποίηση αυτή είναι γνωστή ως *πολική παραγοντοποίηση* (*polar decomposition*) του πίνακα A και αποτελεί ανάλογο της πολικής γραφής ενός μιγαδικού αριθμού $z = |z|e^{i\theta}$, όπου το μέτρο $|z| \geq 0$ αντιστοιχεί στον θετικά ημιορισμένο πίνακα P και το $e^{i\theta}$ στον ορθομοναδιαίο πίνακα W .

4.3 Προσεγγίσεις Χαμηλού Βαθμού (Low-Rank)

Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \mu}$ βαθμού $\text{rank}(A) = r$, με παραγοντοποίηση $SV D$

$$\begin{aligned} A &= U \text{diag}\{s_1, s_2, \dots, s_{\min\{\nu, \mu\}}\} V^* \\ &= [u_1 \ u_2 \ \dots \ u_\nu] \begin{bmatrix} s_1 & 0 & \dots & 0 & \dots \\ 0 & s_2 & \dots & 0 & \dots \\ \vdots & \vdots & \ddots & \vdots & \\ 0 & 0 & \dots & s_r & \dots \\ \vdots & \vdots & & \vdots & 0_{(\nu-r) \times (\mu-r)} \end{bmatrix} \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_\mu^* \end{bmatrix}. \end{aligned}$$

Με απλές πράξεις βλέπουμε ότι

$$\begin{aligned} A &= \sum_{i=1}^r [u_1 \ u_2 \ \dots \ u_\nu] \text{diag}\{0, \dots, 0, s_i, 0, \dots, 0\} \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_\mu^* \end{bmatrix} \\ &= \sum_{i=1}^r s_i u_i v_i^*. \end{aligned}$$

Δηλαδή, ο A γράφεται ως άθροισμα r πινάκων βαθμού 1. Η συγκεκριμένη γραφή οδηγεί σε προσέγγιση του πίνακα A από πίνακες που έχουν βαθμό μικρότερο του $r = \text{rank}(A)$.

Θεώρημα 4.3.1. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \mu}$ βαθμού $\text{rank}(A) = r$, με παραγοντοποίηση $SV D$, $A = U \text{diag}\{s_1, s_2, \dots, s_r, 0, \dots, 0\} V^*$ ($s_1 \geq s_2 \geq \dots \geq s_r > 0$). Για τυχαίο φυσικό αριθμό $k < r$, θεωρούμε τον πίνακα

$$A_k = \sum_{i=1}^k s_i u_i v_i^*.$$

Τότε ισχύει ότι $\text{rank}(A_k) = k$ και $\min_{\text{rank}(B)=k} \|A - B\|_2 = \|A - A_k\|_2 = s_{k+1}$.

Απόδειξη. Χωρίς βλάβη της γενικότητας, θεωρούμε ότι $\nu \leq \mu$. Από τον ορισμό του A_k προκύπτει άμεσα ότι

$$\text{rank}(A_k) = k,$$

$$U^*(A - A_k)V = \text{diag}\{0, 0, \dots, 0, s_{k+1}, s_{k+2}, \dots, s_{\min\{\nu, \mu\}}\}$$

και

$$\|A - A_k\|_2 = s_{k+1}.$$

Είναι λοιπόν φανερό ότι

$$\min_{\text{rank}(B)=k} \|A - B\|_2 \leq \|A - A_k\|_2 = s_{k+1}.$$

Ας θεωρήσουμε τώρα έναν $\nu \times \mu$ πίνακα B βαθμού $\text{rank}(B) = k$ και μία ορθοκανονική βάση του πυρήνα $\text{Null}(B)$, $\{x_1, x_2, \dots, x_{\mu-k}\}$. Λόγω των διαστάσεων τους, οι υπόχωροι $\text{span}\{x_1, x_2, \dots, x_{\mu-k}\} = \text{Null}(B)$ και $\text{span}\{v_1, v_2, \dots, v_{k+1}\}$ έχουν μη τετριμμένη τομή. Δηλαδή, υπάρχει μοναδιαίο διάνυσμα

$$y \in \text{Null}(B) \cap \text{span}\{v_1, v_2, \dots, v_{k+1}\}.$$

Αφού $By = 0$ και

$$Ay = \sum_{i=1}^{k+1} s_i u_i v_i^* y = \sum_{i=1}^{k+1} s_i (v_i^* y) u_i,$$

έπεται ότι

$$\begin{aligned} \|A - B\|_2^2 &\geq \|(A - B)y\|_2^2 = \|Ay\|_2^2 = \left(\sum_{i=1}^{k+1} s_i (v_i^* y) u_i \right)^* \left(\sum_{i=1}^{k+1} s_i (v_i^* y) u_i \right) \\ &= \sum_{i=1}^{k+1} s_i^2 |v_i^* y|^2 \geq s_{k+1}^2 \sum_{i=1}^{k+1} |v_i^* y|^2 = s_{k+1}^2 \|y\|_2^2 = s_{k+1}^2. \end{aligned}$$

Άρα $\min_{\text{rank}(B)=k} \|A - B\|_2 = s_{k+1}$ και η απόδειξη ολοκληρώθηκε. $\square$

Πόρισμα 4.3.2. (Απόσταση από τη μη αντιστρεψιμότητα) Έστω ένας αντιστρέψιμος πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ με παραγοντοποίηση SVD ,

$$A = U \text{diag}\{s_1, s_2, \dots, s_\nu\} V^* \quad (s_1 \geq s_2 \geq \dots \geq s_\nu > 0).$$

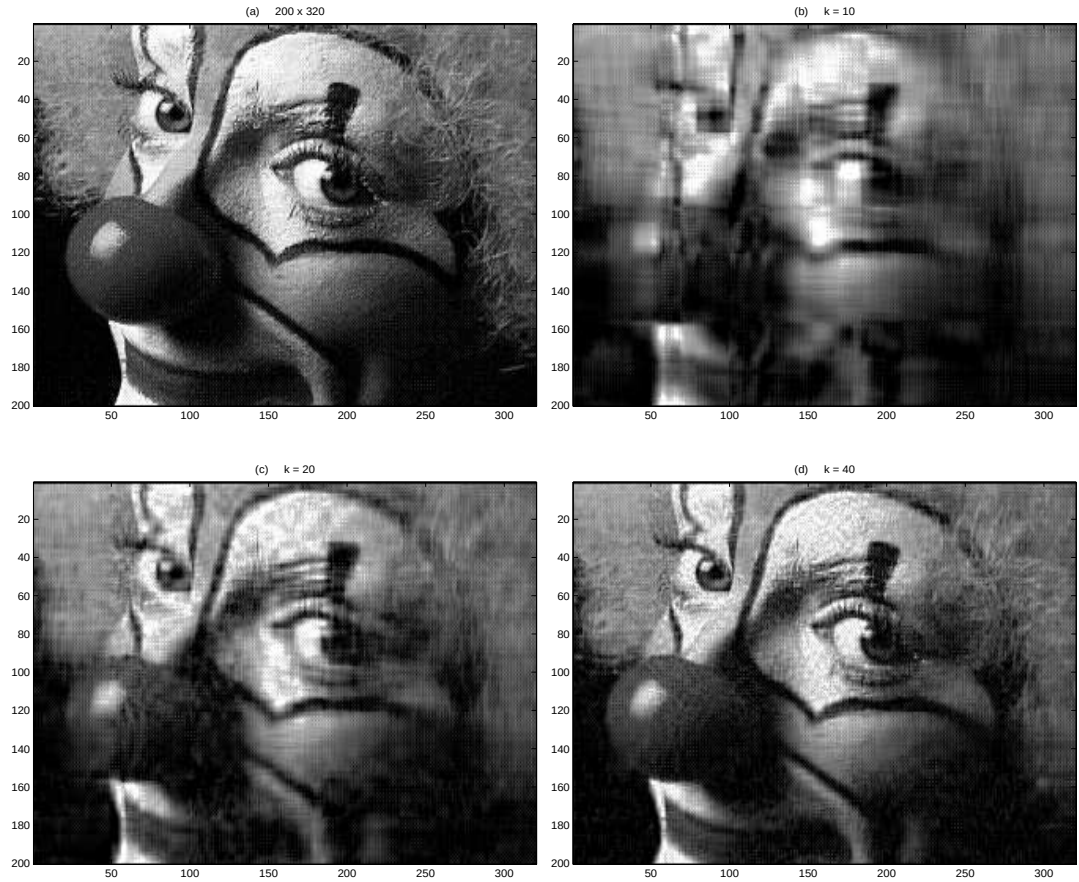
Τότε

$$\min\{\|\Delta\|_2 : \Delta \in \mathbb{C}^{\nu \times \mu}, \det(A + \Delta) = 0\} = s_\nu$$

και ένας πλησιέστερος στον A μη αντιστρέψιμος πίνακας (όχι κατ' ανάγκη μοναδικός) είναι ο

$$\begin{aligned} A_{\nu-1} &= U \text{diag}\{s_1, s_2, \dots, s_{\nu-1}, 0\} V^* \\ &= s_1 u_1 v_1^* + s_2 u_2 v_2^* + \dots + s_{\nu-1} u_{\nu-1} v_{\nu-1}^*. \end{aligned}$$

Η γνώση ενός $\nu \times \mu$ πίνακα A απαιτεί τη διαχείριση $\nu \mu$ στοιχείων, ενώ η γνώση της προσέγγισης του A , $A_k = s_1 u_1 v_1^* + s_2 u_2 v_2^* + \dots + s_k u_k v_k^*$, απαιτεί τη διαχείριση $(\nu + \mu + 1)k$ στοιχείων. Επομένως, αν $k < (\nu \mu)/(\nu + \mu + 1)$, τότε η προσέγγιση A_k απαιτεί τη γνώση και διαχείριση λιγότερων στοιχείων από τα



Σχήμα 4.1: Προσεγγίσεις εικόνας με χρήση της παραγοντοποίησης SVD .

στοιχεία του ίδιου του πίνακα A . Ειδικά στην περίπτωση που ο πίνακας A έχει πολλές “μικρές” ιδιάζουσες τιμές, μπορούμε να επιλέξουμε k πολύ μικρότερο του $(\nu + \mu) / (\nu + \mu + 1)$ ώστε η προσέγγιση A_k να είναι ικανοποιητική αλλά και ιδιαίτερα οικονομική ως προς τη διαχείριση στοιχείων. Κάτι τέτοιο είναι σύνηθες στη μεταφορά ψηφιακής εικόνας σε μεγάλες αποστάσεις (δορυφόροι, διαστημικές αποστολές), όπου τα στοιχεία του πίνακα είναι δείκτες φωτεινότητας. Για παράδειγμα, η εικόνα (a) του Σχήματος 4.1 αντιστοιχεί σε έναν 200×320 πίνακα και είναι μια γνωστή εικόνα του Matlab, ενώ οι εικόνες (b), (c) και (d) είναι προσεγγίσεις της που προκύπτουν με τη χρήση $k = 10$, $k = 20$ και $k = 40$ ιδιάζουσών τιμών, αντίστοιχα, και παράγονται με τις ακόλουθες εντολές του Matlab:

```
load clown.mat; [U,S,V] = svd(X); colormap('gray');
image(U(:,1:k) * S(1:k,1:k) * V(:,1:k)')
```

Είναι φανερό ότι, για $k = 40$ έχουμε μία πολύ ικανοποιητική προσέγγιση που απαιτεί τη γνώση και διαχείριση 20840 στοιχείων αντί των 64000 στοιχείων του αρχικού πίνακα.

Στη συνέχεια, περιοριζόμαστε στην περίπτωση των τετραγωνικών πινάκων και θεωρούμε ένα $\nu \times \nu$ μιγαδικό πίνακα A και έναν αριθμό $\lambda \in \mathbb{C}$. Αν ο $\lambda I_\nu - A$ έχει παραγοντοποίηση SVD της μορφής

$$\lambda I_\nu - A = U \Sigma V^* = U \operatorname{diag} \{s_1(\lambda), s_2(\lambda), \dots, s_\nu(\lambda)\} V^*,$$

με ιδιάζουσες τιμές $s_1(\lambda) \geq s_2(\lambda) \geq \dots \geq s_\nu(\lambda) \geq 0$ (εξαρτόμενες από το λ), τότε από το γεγονός ότι ο βαθμός $\operatorname{rank}(\lambda I_\nu - A)$ ταυτίζεται με το πλήθος των μη μηδενικών ιδιάζουσων τιμών του $\lambda I_\nu - A$, έπεται το παρακάτω αποτέλεσμα.

Πόρισμα 4.3.3. Ο αριθμός $\lambda \in \mathbb{C}$ είναι ιδιοτιμή του πίνακα A γεωμετρικής πολλαπλότητας γ αν και μόνο αν

$$s_1(\lambda) \geq s_2(\lambda) \geq \dots \geq s_{\nu-\gamma}(\lambda) > s_{\nu-\gamma+1}(\lambda) = \dots = s_\nu(\lambda) = 0.$$

Το πόρισμα αυτό και το Θεώρημα 4.3.1 οδηγούν στον ορισμό και τον υπολογισμό της απόστασης του A από το σύνολο των $\nu \times \nu$ πινάκων που έχουν έναν τυχαίο αριθμό $\lambda \in \mathbb{C}$ ως ιδιοτιμή γεωμετρικής πολλαπλότητας (τουλάχιστον) γ ,

$$\begin{aligned} \mathcal{E}_A(\lambda, \gamma) &= \min \{ \|\Delta\|_2 : \Delta \in \mathbb{C}^{\nu \times \nu}, \text{ με } \lambda \text{ ιδιοτιμή του } A + \Delta \\ &\quad \text{γεωμετρικής πολλαπλότητας } \gamma \} \\ &= s_{\nu-\gamma+1}(\lambda). \end{aligned}$$

Πλησιέστερος στον A πίνακας με τον λ ως ιδιοτιμή γεωμετρικής πολλαπλότητας γ είναι ο

$$A + \Delta_0 = \lambda I_\nu - U \operatorname{diag} \{s_1, \dots, s_{\nu-\gamma}, 0, \dots, 0\} V^*,$$

όπου $\|\Delta_0\|_2 = s_{\nu-\gamma+1} = \mathcal{E}_A(\lambda, \gamma)$.

4.4 Ευστάθεια Τετραγωνικών Γραμμικών Συστημάτων

Έστω ένας $\nu \times \nu$ αντιστρέψιμος πίνακας A με παραγοντοποίηση SVD

$$A = U \Sigma V^* = U \operatorname{diag} \{s_1, s_2, \dots, s_\nu\} V^* = \sum_{i=1}^{\nu} s_i u_i v_i^*.$$

Τότε ισχύει

$$(U \operatorname{diag} \{s_1, s_2, \dots, s_\nu\} V^*) (V \operatorname{diag} \{s_1^{-1}, s_2^{-1}, \dots, s_\nu^{-1}\} U^*) = I_\nu,$$

δηλαδή

$$A^{-1} = V \operatorname{diag} \left\{ \frac{1}{s_1}, \frac{1}{s_2}, \dots, \frac{1}{s_\nu} \right\} U^*,$$

άρα και $\|A^{-1}\|_2 = 1/s_\nu$. Συνεπώς, το σύστημα $Ax = \beta$ έχει λύση

$$x = A^{-1}\beta = \sum_{i=1}^{\nu} \left(\frac{u_i^* \beta}{s_i} \right) v_i.$$

Η γραφή αυτή δείχνει ότι μικρές διαταραχές στα στοιχεία του A ή του β μπορούν να προκαλέσουν σχετικά μεγάλες διαταραχές στη λύση x αν η ιδιάζουσα τιμή s_ν είναι “μικρή”, δηλαδή αν ο A είναι “κοντά” σε μη αντιστρέψιμο πίνακα.

Για να μελετήσουμε την ευστάθεια του συστήματος $Ax = \beta$, θεωρούμε το διαταραγμένο και παραμετροποιημένο σύστημα

$$(A + \varepsilon F)x(\varepsilon) = \beta + \varepsilon f, \quad x(0) = x, \quad (4.4.1)$$

με $F \in \mathbb{C}^{\nu \times \nu}$ και $f \in \mathbb{C}^\nu$ δεδομένα και ε μία πραγματική μεταβλητή. Αφού ο πίνακας A είναι αντιστρέψιμος, υπάρχει περιοχή του 0 όπου ορίζεται η συνεχής διανυσματική συνάρτηση

$$x(\varepsilon) = (A + \varepsilon F)^{-1}(\beta + \varepsilon f).$$

Επιπλέον, παραγωγίζοντας κατά μέλη τη σχέση (4.4.1) καταλήγουμε στη διανυσματική παράγωγο

$$\dot{x}(0) = A^{-1}(f - Fx)$$

και το ανάπτυγμα Taylor πρώτης τάξης της $x(\varepsilon)$,

$$\begin{aligned} x(\varepsilon) &= x + \varepsilon \dot{x}(0) + O(\varepsilon^2) \\ &= x + \varepsilon A^{-1}(f - Fx) + O(\varepsilon^2). \end{aligned}$$

Επομένως,

$$\frac{\|x(\varepsilon) - x\|_2}{\|x\|_2} \leq |\varepsilon| \|A^{-1}\|_2 \left(\frac{\|f\|_2}{\|x\|_2} + \|F\|_2 \right) + O(\varepsilon^2).$$

Στο σημείο αυτό πρέπει να ορίσουμε το *βαθμό (δείκτη) κατάστασης* (για την 2-νόρμα) του πίνακα A ,

$$\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{s_1(A)}{s_\nu(A)} \geq 1,$$

δεχόμενοι κατά σύμβαση ότι $\kappa_2(A) = \infty$ όταν ο A δεν είναι αντιστρέψιμος. Χρησιμοποιώντας την ανισότητα $\|\beta\|_2 \leq \|A\|_2 \|x\|_2$, βλέπουμε ότι

$$\frac{\|x(\varepsilon) - x\|_2}{\|x\|_2} \leq \kappa_2(A) \left(|\varepsilon| \frac{\|F\|_2}{\|A\|_2} + |\varepsilon| \frac{\|f\|_2}{\|\beta\|_2} \right) + O(\varepsilon^2),$$

όπου $|\varepsilon| \frac{\|F\|_2}{\|A\|_2}$ και $|\varepsilon| \frac{\|f\|_2}{\|\beta\|_2}$ είναι τα σχετικά σφάλματα στον πίνακα A και στο διάνυσμα β . Δηλαδή, το σχετικό σφάλμα της λύσης x μπορεί να γίνει ίσο με

$k_2(A)$ φορές το συνολικό σχετικό σφάλμα των A και β . Αυτό σημαίνει ότι ο βαθμός κατάστασης $k_2(A)$ “ποσοτικοποιεί” την ευαισθησία του συστήματος $Ax = \beta$. Οι πίνακες με σχετικά μικρό βαθμό κατάστασης ονομάζονται *πίνακες καλής κατάστασης* (*well conditioned*).

Ο βαθμός κατάστασης του A ορίζεται για κάθε νόρμα με τον ίδιο τρόπο. Έτσι έχουμε $\kappa_1(A) = \|A\|_1 \|A^{-1}\|_1$, $\kappa_\infty(A) = \|A\|_\infty \|A^{-1}\|_\infty$ κ.ο.κ. Επιπλέον, από την Πρόταση 1.3.4 προκύπτουν οι ανισότητες

$$\frac{1}{\nu} \kappa_1(A) \leq \kappa_2(A) \leq \nu \kappa_1(A),$$

$$\frac{1}{\nu} \kappa_\infty(A) \leq \kappa_2(A) \leq \nu \kappa_\infty(A)$$

και

$$\frac{1}{\nu} \kappa_1(A) \leq \kappa_\infty(A) \leq \nu \kappa_1(A).$$

Τέλος, χωρίς απόδειξη, έχουμε το ακόλουθο αποτέλεσμα.

Θεώρημα 4.4.1. Έστω ένα $\nu \times \nu$ γραμμικό σύστημα

$$Ax = \beta \quad (A \in \mathbb{C}^{\nu \times \nu}, \beta \in \mathbb{C}^\nu)$$

και μία διαταραχή του,

$$(A + \Delta A)y = \beta + \Delta\beta,$$

όπου x και y είναι οι αντίστοιχες λύσεις. Έστω ακόμη $\|\cdot\|$ μία οποιαδήποτε νόρμα πίνακα που επάγεται από νόρμα διανύσματος, με αντίστοιχο βαθμό κατάστασης $\kappa(A)$. Αν $\|\Delta A\| \leq \varepsilon \|A\|$ και $\|\Delta\beta\| \leq \varepsilon \|\beta\|$, τότε ισχύει

$$\frac{\|y - x\|}{\|x\|} \leq \frac{2\varepsilon \kappa(A)}{1 - \varepsilon \kappa(A)}.$$

4.5 Ψευδοαντίστροφος Πίνακας

Όπως είδαμε στην προηγούμενη παράγραφο, αν ο πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ είναι αντιστρέψιμος με παραγοντοποίηση SVD της μορφής $A = U \operatorname{diag} \{s_1, s_2, \dots, s_\nu\} V^*$ ($s_1 \geq s_2 \geq \dots \geq s_\nu > 0$), τότε

$$A^{-1} = V \operatorname{diag} \{s_1^{-1}, s_2^{-1}, \dots, s_\nu^{-1}\} U^*$$

και $\|A^{-1}\|_2 = 1/s_\nu$. Στην περίπτωση που ο A είναι ένας $\nu \times \mu$ πίνακας βαθμού $\text{rank}(A) = r$ και με παραγοντοποίηση SVD

$$A = U \begin{bmatrix} s_1 & 0 & \cdots & 0 & 0 & \cdots \\ 0 & s_2 & \cdots & 0 & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ 0 & 0 & \cdots & s_r & 0 & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots \\ \vdots & \vdots & & \vdots & \vdots & \ddots \end{bmatrix}_{(\nu \times \mu)} V^*,$$

ορίζουμε τον $\mu \times \nu$ (*Moore-Penrose*) ψευδοαντίστροφο του A ,

$$A^\dagger = V \begin{bmatrix} s_1^{-1} & 0 & \cdots & 0 & 0 & \cdots \\ 0 & s_2^{-1} & \cdots & 0 & 0 & \cdots \\ \vdots & \vdots & \ddots & \vdots & \vdots & \\ 0 & 0 & \cdots & s_r^{-1} & 0 & \cdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots \\ \vdots & \vdots & & \vdots & \vdots & \ddots \end{bmatrix}_{(\mu \times \nu)} U^*.$$

Ο ψευδοαντίστροφος $A^\dagger$ ικανοποιεί τις ακόλουθες σχέσεις (βλέπε την Άσκηση 4.6.4 παρακάτω):

1. $A A^\dagger A = A$,
2. $A^\dagger A A^\dagger = A^\dagger$,
3. οι πίνακες $A A^\dagger$ και $A^\dagger A$ είναι ερμιτιανοί.

Αν ο πίνακας A είναι $\nu \times \mu$ ($\nu > \mu$), με βαθμό $\text{rank}(A) = \mu$, τότε το πρόβλημα των ελαχίστων τετραγώνων

$$\min_{x \in \mathbb{C}_\mu} \|A x - \beta\|_2$$

έχει, όπως είδαμε στην Παράγραφο 3.5, λύση

$$x = R^{-1} Q^* \beta,$$

όπου Q και R είναι οι πίνακες της παραγοντοποίησης QR . Εύκολα μπορεί κανείς να επαληθεύσει ότι

$$\begin{aligned} x = R^{-1} Q^* \beta &\Leftrightarrow Q R x = Q Q^* \beta \\ &\Leftrightarrow (R^* Q^*) (Q R) x = R^* Q^* \beta \\ &\Leftrightarrow A^* A x = A^* \beta \\ &\Leftrightarrow x = (A^* A)^{-1} A^* \beta \\ &\Leftrightarrow x = A^\dagger \beta. \end{aligned}$$

4.6 Ασκήσεις

1. Έστω ένας $\nu \times \mu$ πίνακας A βαθμού $\text{rank}(A) = r$, με παραγοντοποίηση SVD , $A = U \Sigma V^*$. Να αποδείξετε ότι οι $\mu - r$ τελευταίες στήλες του V σχηματίζουν τη βάση του πυρήνα του A , ενώ οι r πρώτες στήλες του U σχηματίζουν βάση της εικόνας του A .
2. Να αποδείξετε ότι η πολική παραγοντοποίηση ενός τετραγωνικού πίνακα οδηγεί στην παραγοντοποίηση SVD .
3. Περιγράψτε (με λεπτομέρειες) την παραγοντοποίηση SVD ενός κανονικού πίνακα και ενός ερμιτιανού πίνακα.

4. Κατασκευάστε μια παραγοντοποίηση SVD του πίνακα $A = \begin{bmatrix} 0 & 1 & 1 \\ \sqrt{2} & 2 & 0 \\ 0 & 1 & 1 \end{bmatrix}$.

5. Κατασκευάστε μια παραγοντοποίηση SVD του πίνακα $A = \begin{bmatrix} -1 & -1 & -1 \\ 0 & 1 & -1 \\ -1 & 0 & 1 \\ 0 & 0 & -1 \end{bmatrix}$.

6. Έστω $A^\dagger$ ο ψευδοαντίστροφος ενός πίνακα $A \in \mathbb{C}^{\nu \times \mu}$. Να αποδείξετε ότι:

(α') οι πίνακες $AA^\dagger$ και $A^\dagger A$ είναι ερμιτιανοί,

(β') $AA^\dagger A = A$,

(γ') $A^\dagger A A^\dagger = A^\dagger$.

7. Για έναν πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, να αποδείξετε με εφαρμογή της παραγοντοποίησης SVD τα ακόλουθα:

(α') Η μέγιστη ιδιάζουσα τιμή του A είναι

$$s_1 = \max\{|y^* A x| : x, y \in \mathbb{C}^\nu, \|x\|_2 = \|y\|_2 = 1\}.$$

(β') Η νόρμα Frobenius του A ικανοποιεί τη σχέση

$$\|A\|_2 \leq \|A\|_F \leq \sqrt{\text{rank}(A)} \|A\|_2.$$

8. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ με ιδιάζουσες τιμές $s_1 \geq s_2 \geq \dots \geq s_\nu \geq 0$. Να αποδείξετε ότι ο A είναι ορθομοναδιαίος αν και μόνο αν $s_1 = s_2 = \dots = s_\nu = 1$.
9. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ με ιδιάζουσες τιμές $s_1 \geq s_2 \geq \dots \geq s_\nu \geq 0$. Να αποδείξετε ότι $\text{trace}(A) \leq s_1 + s_2 + \dots + s_\nu$.

10. Να αποδείξετε ότι για κάθε $\nu \times \mu$ ($\nu \geq \mu$) πίνακα A βαθμού $\text{rank}(A) = \mu$, ισχύει $\|A(A^*A)^{-1}A^*\|_2 = 1$.
11. Να αποδείξετε ότι για κάθε νόρμα πίνακα, ο αντίστοιχος βαθμός κατάστασης $\kappa(A)$ ικανοποιεί τις σχέσεις:
 - (α') $\kappa(AB) \leq \kappa(A)\kappa(B)$, και
 - (β') $\kappa(aA) = \kappa(A)$, για κάθε $a \neq 0$.
12. Ισχύει ή όχι το ότι δύο $\nu \times \nu$ πίνακες A και B είναι ορθομοναδιαία όμοιοι αν και μόνο αν έχουν τις ίδιες ιδιάζουσες τιμές;
13. Έστω δύο διανύσματα $x, y \in \mathbb{C}^\nu$. Κατασκευάστε μια παραγοντοποίηση SVD του πίνακα $A = xy^* \in \mathbb{C}^{\nu \times \nu}$.
14. Με εφαρμογή της παραγοντοποίησης SVD , να αποδείξετε ότι κάθε $\nu \times \nu$ πίνακας είναι το όριο μιας ακολουθίας $\nu \times \nu$ αντιστρέψιμων πινάκων (δηλαδή, ότι το σύνολο των $\nu \times \nu$ αντιστρέψιμων πινάκων είναι πυκνό υποσύνολο του $\mathbb{C}^{\nu \times \nu}$).
15. Με εφαρμογή της παραγοντοποίησης SVD , να λυθεί η Άσκηση 1.5.9.
16. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ με παραγοντοποίηση SVD , $A = U\Sigma V^*$. Υπολογίστε τις ιδιάζουσες τιμές και κατασκευάστε μία διαγωνοποίηση του $(2\nu) \times (2\nu)$ πίνακα $\begin{bmatrix} 0 & A^* \\ A & 0 \end{bmatrix}$. Τί παρατηρείτε;
17. Με τη μέθοδο παραγοντοποίησης QR , κατασκευάστε τον ψευδοαντίστροφο του πίνακα $A = \begin{bmatrix} -1 & -1 & 1 \\ 1 & 3 & 3 \\ -1 & -1 & 5 \\ 1 & 3 & 7 \end{bmatrix}$. Έπειτα, να τον κατασκευάσετε με τη μέθοδο παραγοντοποίησης SVD .

Κεφάλαιο 5

Κανονικοί Πίνακες

Κανονικοί πίνακες (φυσιολογικοί τελεστές) εμφανίζονται σε ένα μεγάλο πλήθος σύγχρονων προβλημάτων, τόσο στα θεωρητικά όσο και στα εφαρμοσμένα μαθηματικά. Η έντονη χρήση τους αποτυπώνεται στην ύπαρξη 90 περίπου ισοδύναμων ορισμών στη βιβλιογραφία. Στις παραγράφους που ακολουθούν θα εξετάσουμε την έννοια του κανονικού πίνακα, αποδεικνύοντας 17 ισοδύναμες συνθήκες κανονικότητας και μελετώντας τις 3 σημαντικότερες (από τις τουλάχιστον 20) συγγενείς αποστάσεις από την κανονικότητα.

5.1 Η Έννοια του Κανονικού Πίνακα

Όπως είναι γνωστό, ένας τετραγωνικός πίνακας $A \in \mathbb{C}^{n \times n}$ καλείται *κανονικός* όταν αντιμετωπίζεται με τον αναστροφосуζυγή του, δηλαδή όταν ικανοποιεί τη σχέση

$$AA^* = A^*A.$$

Πρόταση 5.1.1. (Συνθήκη 1) Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν, για κάθε (βαθμωτό) πολυώνυμο $p(\lambda) = a_m \lambda^m + \dots + a_1 \lambda + a_0$, ο πίνακας $p(A) = a_m A^m + \dots + a_1 A + a_0 I_n$ είναι κανονικός.

Απόδειξη. Αν ο A είναι κανονικός, τότε για κάθε πολυώνυμο $p(\lambda)$, προκύπτει με απλές πράξεις ότι $p(A)p(A)^* = p(A)^*p(A)$, δηλαδή ο πίνακας $p(A)$ είναι κανονικός. Το αντίστροφο είναι προφανές. $\square$

Πρόταση 5.1.2. (Συνθήκη 2) Ένας αντιστρέψιμος πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν ο A^{-1} είναι κανονικός.

Απόδειξη. Αν ο A είναι αντιστρέψιμος κανονικός πίνακας, τότε ισχύει

$$(A^{-1})^* A^{-1} = (A^*)^{-1} A^{-1} = (A A^*)^{-1} = (A^* A)^{-1} = A^{-1} (A^*)^{-1} = A^{-1} (A^{-1})^*,$$

δηλαδή ο A^{-1} είναι κανονικός.

Αντίστροφα, αν ο A^{-1} είναι κανονικός, τότε από το ευθύ θα είναι κανονικός και ο πίνακας $(A^{-1})^{-1} = A$. $\square$

Πρόταση 5.1.3. (Συνθήκη 3) Ένας αντιστρέψιμος πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ είναι κανονικός αν και μόνο αν ο $A^{-1}A^*$ είναι ορθομοναδιαίος.

Απόδειξη. Αν ο πίνακας A είναι αντιστρέψιμος και κανονικός, τότε

$$(A^{-1}A^*)(A^{-1}A^*)^* = A^{-1}A^*A(A^{-1})^* = A^{-1}AA^*(A^{-1})^* = I_\nu$$

και

$$(A^{-1}A^*)^*(A^{-1}A^*) = A(A^{-1})^*A^{-1}A^* = AA^{-1}(A^{-1})^*A^* = I_\nu.$$

Άρα $(A^{-1}A^*)(A^{-1}A^*)^* = (A^{-1}A^*)^*(A^{-1}A^*) = I_\nu$.

Αντίστροφα, αν ο $A^{-1}A^*$ είναι ορθομοναδιαίος, τότε

$$(A^{-1}A^*)^* = (A^{-1}A^*)^{-1} \Leftrightarrow A(A^{-1})^* = (A^{-1})^*A \Leftrightarrow A^{-1}A^* = A^*A^{-1}.$$

Πολυπλασιάζοντας την τελευταία σχέση, από αριστερά και από δεξιά, με A , παίρνουμε $A^*A = AA^*$. Δηλαδή, ο πίνακας A είναι κανονικός. $\square$

Πρόταση 5.1.4. (Συνθήκη 4) Ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ είναι κανονικός αν και μόνο αν, για οποιοδήποτε ορθομοναδιαίο πίνακα $U \in \mathbb{C}^{\nu \times \nu}$, ο πίνακας U^*AU είναι κανονικός.

Απόδειξη. Για τυχαίο ορθομοναδιαίο πίνακα $U \in \mathbb{C}^{\nu \times \nu}$, έχουμε ότι

$$\begin{aligned} (U^*AU)(U^*AU)^* &= (U^*AU)^*(U^*AU) \Leftrightarrow U^*AUU^*A^*U = U^*A^*UU^*AU \\ &\Leftrightarrow U^*AA^*U = U^*A^*AU \\ &\Leftrightarrow AA^* = A^*A \end{aligned}$$

και η απόδειξη ολοκληρώθηκε. $\square$

Πρόταση 5.1.5. (Συνθήκη 5) Ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ είναι κανονικός αν και μόνο αν υπάρχει ένας ορθομοναδιαίος πίνακας $U \in \mathbb{C}^{\nu \times \nu}$ κι ένας διαγώνιος πίνακας $\Lambda \in \mathbb{C}^{\nu \times \nu}$ τέτοιοι ώστε $U^*AU = \Lambda$.

Απόδειξη. Από το Λήμμα 1.4.2 του Schur, υπάρχουν ορθομοναδιαίος πίνακας U και άνω τριγωνικός πίνακας T τέτοιοι ώστε

$$A = UTU^* \Leftrightarrow T = U^*AU.$$

Αν λοιπόν ο A είναι κανονικός, τότε από την Πρόταση 5.1.4 (Συνθήκη 4) και ο τριγωνικός πίνακας T είναι κανονικός. Με απλές πράξεις, εύκολα μπορεί κανείς

να επαληθεύσει ότι ο T πρέπει υποχρεωτικά να έχει μηδενικά όλα τα μη διαγώνια στοιχεία του. Δηλαδή, ο T είναι διαγώνιος.

Αντίστροφα, ας υποθέσουμε ότι $A = U \Lambda U^*$, όπου ο πίνακας Λ , ως διαγώνιος, ικανοποιεί τη σχέση $\Lambda \Lambda^* = \Lambda^* \Lambda$ και είναι κανονικός. Από την Πρόταση 5.1.4 (Συνθήκη 4), ο πίνακας A είναι επίσης κανονικός. $\square$

Συνεχίζοντας τη μελέτη της κανονικότητας ενός πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, θα συνδέσουμε τη συγκεκριμένη έννοια με τις ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα.

Πρόταση 5.1.6. (Συνθήκη 6) Ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ είναι κανονικός αν και μόνο αν υπάρχει ορθοκανονική βάση του $\mathbb{C}^{\nu \times \nu}$ που αποτελείται από ιδιοδιανύσματα του A .

Απόδειξη. Προκύπτει άμεσα από την Πρόταση 5.1.5 (Συνθήκη 5). $\square$

Πρόταση 5.1.7. (Συνθήκη 7) Ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ είναι κανονικός αν και μόνο αν ο A είναι διαγωνοποιήσιμος και οποιαδήποτε ιδιοδιανύσματα του τα οποία προέρχονται από διαφορετικές ιδιοτιμές είναι ορθογώνια.

Απόδειξη. Προκύπτει άμεσα από την Πρόταση 5.1.5 (Συνθήκη 5). $\square$

Πρόταση 5.1.8. (Συνθήκη 8) Ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ είναι κανονικός αν και μόνο αν υπάρχει ένα (βαθμωτό) πολυώνυμο $p(\lambda)$ τέτοιο ώστε $A^* = p(A)$.

Απόδειξη. Από την Πρόταση 5.1.5 (Συνθήκη 5), χωρίς βλάβη της γενικότητας, μπορούμε να θεωρήσουμε ότι $A = \text{diag} \{ \lambda_1, \lambda_2, \dots, \lambda_\nu \}$. Αν επιλέξουμε ένα πολυώνυμο παρεμβολής $p(\lambda)$, βαθμού το πολύ $\nu - 1$, που ικανοποιεί τη σχέση $p(\lambda_i) = \bar{\lambda}_i$ ($i = 1, 2, \dots, \nu$), τότε άμεσα προκύπτει ότι $p(A) = A^*$.

Αντίστροφα, αν $A^* = p(A)$, τότε είναι προφανές ότι $A p(A) = p(A) A$, κι επομένως, $A A^* = A^* A$. $\square$

Πρόταση 5.1.9. (Συνθήκη 9) Ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ είναι κανονικός αν και μόνο αν ο A είναι διαγωνοποιήσιμος και κάθε ιδιοδιάνυσμα $x \in \mathbb{C}^\nu$ του A είναι ιδιοδιάνυσμα και του A^* .

Απόδειξη. Αν ο A είναι κανονικός, τότε από την Πρόταση 5.1.5 (Συνθήκη 5),

$$U^* A U = \Lambda \Leftrightarrow U^* A^* U = \Lambda^*$$

για κάποιο διαγώνιο πίνακα $\Lambda \in \mathbb{C}^{\nu \times \nu}$ και για κάποιο ορθομοναδιαίο $U \in \mathbb{C}^{\nu \times \nu}$. Όμως οι διαγώνιοι πίνακες Λ και Λ^* έχουν τα ίδια ιδιοδιανύσματα, τα διανύσματα της κανονικής βάσης. Επομένως, οι A και A^* έχουν τα ίδια ιδιοδιανύσματα, τα διανύσματα που είναι οι στήλες του ορθομοναδιαίου πίνακα U .

Αντίστροφα, ας υποθέσουμε ότι ο πίνακας A είναι διαγωνοποιήσιμος και ότι τα ιδιοδιανύσματα του ταυτίζονται με τα ιδιοδιανύσματα του A^* . Τότε τα ιδιοδιανύσματα του A ταυτίζονται με τα ιδιοδιανύσματα του ερμιτιανού πίνακα $A + A^*$, και συνεπώς, σχηματίζουν μια ορθοκανονική βάση του $\mathbb{C}^n$. Η απόδειξη ολοκληρώνεται από την Πρόταση 5.1.6 (Συνθήκη 6). $\square$

Κάθε τετραγωνικός πίνακας A μπορεί να γραφεί στη μορφή $A = H(A) + S(A)$, όπου $H(A) = (A + A^*)/2$ είναι το ερμιτιανό μέρος και $S(A) = (A - A^*)/2$ το αντιερμιτιανό μέρος του A . Οι ακόλουθες συνθήκες συνδέουν την έννοια της κανονικότητας με αυτήν την ανάλυση καρτεσιανού τύπου.

Πρόταση 5.1.10. (Συνθήκη 10) Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν $H(A)S(A) = S(A)H(A)$.

Απόδειξη. Αν ο πίνακας A είναι κανονικός, τότε προφανώς ισχύει ότι

$$\begin{aligned} H(A)S(A) &= \frac{A + A^*}{2} \frac{A - A^*}{2} = \frac{1}{4} (A^2 - A A^* + A^* A + (A^*)^2) \\ &= \frac{A - A^*}{2} \frac{A + A^*}{2} = S(A)H(A). \end{aligned}$$

Αντίστροφα, έχουμε

$$\begin{aligned} H(A)S(A) &= S(A)H(A) \Rightarrow \frac{1}{4} (A^2 - A A^* + A^* A + (A^*)^2) \\ &= \frac{1}{4} ((A^*)^2 - A^* A + A A^* + A^2) \\ &\Rightarrow -A A^* + A^* A = A A^* - A^* A \\ &\Rightarrow 2 A^* A = 2 A A^*, \end{aligned}$$

δηλαδή ο πίνακας A είναι κανονικός. $\square$

Πρόταση 5.1.11. (Συνθήκη 11) Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν $H(A)A = A H(A)$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πρόταση 5.1.12. (Συνθήκη 12) Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν $S(A)A = A S(A)$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πρόταση 5.1.13. (Συνθήκη 13) Έστω ότι ο ερμιτιανός πίνακας $H(A)$ έχει διακεκριμένες ιδιοτιμές. Τότε ο πίνακας A είναι κανονικός αν και μόνο αν κάθε ιδιοδιάνυσμα του $H(A)$ είναι επίσης ιδιοδιάνυσμα του $S(A)$.

Απόδειξη. Αν ο A είναι κανονικός, τότε από τη Συνθήκη 5, υπάρχει ορθομοναδιαίος πίνακας $U \in \mathbb{C}^{n \times n}$ (με στήλες τα ιδιοδιανύσματα του A) ώστε ο πίνακας $U^*AU = U^*(H(A) + S(A))U = U^*H(A)U + U^*S(A)U$ να είναι διαγώνιος.

Αντίστροφα, υποθέτουμε ότι ο πίνακας $H(A)$ έχει διακεκριμένες ιδιοτιμές και κάθε ιδιοδιάνυσμα του $H(A)$ είναι επίσης ιδιοδιάνυσμα του $S(A)$. Τότε υπάρχει ορθομοναδιαίος πίνακας $U \in \mathbb{C}^{n \times n}$ (με στήλες τα ιδιοδιανύσματα του $H(A)$) ώστε οι πίνακες $U^*H(A)U$ και $U^*S(A)U$ να είναι διαγώνιοι. Προφανώς, και ο $U^*AU = U^*(H(A) + S(A))U$ θα είναι διαγώνιος. $\square$

Όπως αναφέραμε και στην Παράγραφο 4.2, η πολική παραγοντοποίηση ενός πίνακα $A \in \mathbb{C}^{n \times n}$ είναι η γραφή του στη μορφή $A = PW$, όπου ο $P \in \mathbb{C}^{n \times n}$ είναι (ερμιτιανός) θετικά ημιορισμένος και ο $W \in \mathbb{C}^{n \times n}$ είναι ορθομοναδιαίος.

Πρόταση 5.1.14. (Συνθήκη 14) Ένας αντιστρέψιμος πίνακας $A \in \mathbb{C}^{n \times n}$ με πολική παραγοντοποίηση $A = PW$ είναι κανονικός αν και μόνο αν $WP = PW$.

Απόδειξη. Ο ερμιτιανός πίνακας P διαγωνοποιείται στη μορφή $P = U\Lambda U^*$, όπου ο $U \in \mathbb{C}^{n \times n}$ είναι ορθομοναδιαίος και ο $\Lambda \in \mathbb{C}^{n \times n}$ είναι διαγώνιος με τα διαγώνια στοιχεία του θετικά (λόγω της αντιστρεψιμότητας του A). Αν λοιπόν υποθέσουμε ότι ο A είναι κανονικός, τότε

$$\begin{aligned} AA^* = A^*A &\Rightarrow (U\Lambda U^*W)(U\Lambda U^*W)^* = (U\Lambda U^*W)^*(U\Lambda U^*W) \\ &\Rightarrow U\Lambda^2 U^* = W^*U\Lambda^2 U^*W \\ &\Rightarrow \Lambda^2 = (U^*WU)^*\Lambda^2(U^*WU). \end{aligned}$$

Δηλαδή, ο ορθομοναδιαίος πίνακας U^*WU έχει στήλες τα ιδιοδιανύσματα του Λ^2 , τα οποία είναι και ιδιοδιανύσματα του Λ . Επομένως,

$$\begin{aligned} \Lambda &= (U^*WU)^*\Lambda(U^*WU) \Rightarrow U\Lambda U^* = W^*U\Lambda U^*W \\ &\Rightarrow W(U\Lambda U^*) = (U\Lambda U^*)W \\ &\Rightarrow WP = PW. \end{aligned}$$

Αντίστροφα, αν $WP = PW$, τότε με απλές πράξεις μπορεί κανείς να επαληθεύσει ότι $A^*A = AA^*$. $\square$

Πρόταση 5.1.15. (Συνθήκη 15) Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν $AW = WA$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πρόταση 5.1.16. (Συνθήκη 16) Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν $AP = PA$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πρόταση 5.1.17. (Συνθήκη 17) Ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$, με (όχι απαραίτητα διακεκριμένες) ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_\nu$, είναι κανονικός αν και μόνο αν οι ιδιάζουσες τιμές του είναι $|\lambda_1|, |\lambda_2|, \dots, |\lambda_\nu|$.

Απόδειξη. Χωρίς βλάβη της γενικότητας, μπορούμε να θεωρήσουμε ότι

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_\nu| \quad \text{και} \quad \lambda_i = e^{i\theta_i} |\lambda_i| \quad (i = 1, 2, \dots, \nu).$$

Από την Πρόταση 5.1.5 (Συνθήκη 5), έχουμε ότι ο $A \in \mathbb{C}^{\nu \times \nu}$ είναι κανονικός αν και μόνο αν υπάρχει ένας ορθομοναδιαίος πίνακας $V = [v_1 \ v_2 \ \dots \ v_\nu] \in \mathbb{C}^{\nu \times \nu}$ (με στηλές $v_1, v_2, \dots, v_\nu \in \mathbb{C}^\nu$) κι ένας διαγώνιος πίνακας $\Lambda \in \mathbb{C}^{\nu \times \nu}$ τέτοιοι ώστε

$$\begin{aligned} A &= V \Lambda V^* \\ &= [v_1 \ v_2 \ \dots \ v_\nu] \begin{bmatrix} e^{i\theta_1} |\lambda_1| & 0 & \dots & 0 \\ 0 & e^{i\theta_2} |\lambda_2| & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & e^{i\theta_\nu} |\lambda_\nu| \end{bmatrix} \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_\nu^* \end{bmatrix} \\ &= [e^{i\theta_1} v_1 \ e^{i\theta_2} v_2 \ \dots \ e^{i\theta_\nu} v_\nu] \begin{bmatrix} |\lambda_1| & 0 & \dots & 0 \\ 0 & |\lambda_2| & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & |\lambda_\nu| \end{bmatrix} \begin{bmatrix} v_1^* \\ v_2^* \\ \vdots \\ v_\nu^* \end{bmatrix}. \end{aligned}$$

Η παραπάνω γραφή είναι προφανώς μία παραγοντοποίηση SVD του A . Επομένως, αν ο A είναι κανονικός, τότε οι ιδιάζουσες τιμές του είναι $|\lambda_1|, |\lambda_2|, \dots, |\lambda_\nu|$.

Για το αντίστροφο, από το Λήμμα 1.4.2 του Schur, υποθέτουμε χωρίς βλάβη της γενικότητας ότι ο A είναι άνω τριγωνικός με τις ιδιοτιμές του στη διαγώνιο. Θεωρούμε επίσης τον ερμιτιανό πίνακα

$$\begin{aligned} A^* A &= \begin{bmatrix} \bar{\lambda}_1 & 0 & \dots & 0 \\ * & \bar{\lambda}_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & \bar{\lambda}_\nu \end{bmatrix} \begin{bmatrix} \lambda_1 & * & \dots & * \\ 0 & \lambda_2 & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_\nu \end{bmatrix} \\ &= \begin{bmatrix} |\lambda_1|^2 & * & \dots & * \\ * & |\lambda_2|^2 + \rho_{2,2} & \dots & * \\ \vdots & \vdots & \ddots & \vdots \\ * & * & \dots & |\lambda_\nu|^2 + \rho_{\nu,\nu} \end{bmatrix}, \end{aligned}$$

όπου οι $\nu - 1$ αριθμοί $\rho_{2,2}, \dots, \rho_{\nu,\nu}$ είναι μη αρνητικοί και τα διαγώνια στοιχεία $|\lambda_1|^2, |\lambda_2|^2 + \rho_{2,2}, \dots, |\lambda_\nu|^2 + \rho_{\nu,\nu}$ είναι τα τετράγωνα των (Ευκλείδειων) μέτρων των στηλών του A . Συνεπώς, αν ένα οποιοδήποτε στοιχείο του τριγωνικού πίνακα A πάνω από την κύρια διαγώνιο είναι μη μηδενικό, τότε το ίχνος του $A^* A$ είναι μεγαλύτερο του αθροίσματος $|\lambda_1|^2 + |\lambda_2|^2 + \dots + |\lambda_\nu|^2$, κάτι που οδηγεί σε άτοπο. Άρα, κατά την εφαρμογή του Λήμματος 1.4.2 του Schur, ο A διαγωνοποιείται με ορθομοναδιαίο μετασχηματισμό ομοιότητας και είναι κανονικός πίνακας. $\square$

5.2 Αποστάσεις από την Κανονικότητα

Οι ορισμοί των αποστάσεων από την κανονικότητα που ακολουθούν βασίζονται στην επιλογή της νόρμας πινάκων που χρησιμοποιούμε. Συνήθως, στη βιβλιογραφία, επιλέγονται η φασματική (τελεστική) νόρμα $\| \circ \|_2$ και η νόρμα Frobenius $\| \circ \|_F$ που μελετήσαμε στο πρώτο κεφάλαιο.

Ορισμός 5.2.1. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ και μία νόρμα πινάκων $\| \circ \|$.

- (i) Η τοπολογική απόσταση του A από την κανονικότητα (με βάση τη νόρμα $\| \circ \|$) ορίζεται ως

$$\mathrm{dn}_{\| \circ \|}^{\mathrm{top}}(A) = \min \{ \|A - N\| : N \in \mathbb{C}^{\nu \times \nu}, N^*N = NN^* \}.$$

- (ii) Η απόσταση μέσω αντιμεταθέτη του A από την κανονικότητα (με βάση τη νόρμα $\| \circ \|$) ορίζεται ως

$$\mathrm{dn}_{\| \circ \|}^{\mathrm{com}}(A) = \|A^*A - AA^*\|^{1/2}.$$

- (iii) Υπενθυμίζουμε ότι, από το Λήμμα 1.4.2 του Schur, ο $A \in \mathbb{C}^{\nu \times \nu}$ γράφεται στη μορφή

$$A = UTU^* = U(\Lambda_U + R_U)U^*$$

για κάποιο ορθομοναδιαίο πίνακα U , όπου ο $T = \Lambda_U + R_U$ είναι άνω τριγωνικός, ο Λ_U είναι διαγώνιος και ο R_U είναι γνήσια άνω τριγωνικός. Η απομάκρυνση κατά *Henrici* του A από την κανονικότητα (με βάση τη νόρμα $\| \circ \|$) ορίζεται ως

$$\delta_{\| \circ \|}(A) = \min \{ \|R_U\| : U^*AU = \Lambda_U + R_U, U^*U = UU^* = I_\nu \}.$$

Για κάθε $A \in \mathbb{C}^{\nu \times \nu}$, οι αποστάσεις $\mathrm{dn}_{\| \circ \|}^{\mathrm{top}}(A)$, $\mathrm{dn}_{\| \circ \|}^{\mathrm{com}}(A)$ και $\delta_{\| \circ \|}(A)$ λαμβάνουν μη αρνητικές τιμές, ενώ στην περίπτωση που ο πίνακας A είναι κανονικός, προκύπτει άμεσα ότι $\mathrm{dn}_{\| \circ \|}^{\mathrm{top}}(A) = \mathrm{dn}_{\| \circ \|}^{\mathrm{com}}(A) = \delta_{\| \circ \|}(A) = 0$. Επιπλέον, αξίζει να παρατηρήσουμε ότι:

- Η τοπολογική απόσταση $\mathrm{dn}_{\| \circ \|}^{\mathrm{top}}(A)$ συνδέεται με την έννοια του πλησιέστερου στον A κανονικού πίνακα, αλλά ο ακριβής υπολογισμός της αποτελεί γενικά ένα δύσκολο πρόβλημα.
- Η απόσταση μέσω αντιμεταθέτη $\mathrm{dn}_{\| \circ \|}^{\mathrm{com}}(A)$ υπολογίζεται εύκολα, αλλά δεν συνδέεται με την έννοια του πλησιέστερου στον A κανονικού πίνακα.
- Η απομάκρυνση κατά *Henrici* $\delta_{\| \circ \|}(A)$ είναι άμεσα υπολογίσιμη μόνο για τη νόρμα Frobenius (βλέπε την Πρόταση 5.2.1 παρακάτω), συνδέεται με την έννοια του πλησιέστερου στον A κανονικού πίνακα και μετρά την απόσταση

του A από τους κανονικούς πίνακες που έχουν τις ίδιες ακριβώς ιδιοτιμές με τον A , λαμβάνοντας υπόψη και τις πολλαπλότητες. Λόγω αυτής της τελευταίας παρατήρησης, χρησιμοποιούμε τον όρο “απομάκρυνση” αντί του όρου “απόσταση”.

Παρατήρηση 5.2.1. Έστω ότι η νόρμα $\|\cdot\|$ είναι ορθομοναδιαία αναλλοίωτη. Αν για κάθε κανονικό πίνακα $N \in \mathbb{C}^{\nu \times \nu}$ θεωρήσουμε όλες τις ορθομοναδιαίες διαγωνοποιήσεις της μορφής $N = U\Lambda U^*$ (όπου ο πίνακας U είναι ορθομοναδιαίος και ο Λ είναι διαγώνιος), τότε ο Ορισμός 5.2.1 (i) της τοπολογικής απόστασης από την κανονικότητα γράφεται

$$\mathrm{dn}_{\|\cdot\|}^{\mathrm{top}}(A) = \min \{ \|U^*AU - \Lambda\| : \Lambda \text{ διαγώνιος, } U^*U = UU^* = I_\nu \}. \quad (5.2.1)$$

Ας υποθέσουμε τώρα ότι ο $N \in \mathbb{C}^{\nu \times \nu}$ είναι ένας από τους πλησιέστερους στον A κανονικούς πίνακες, με βάση την τοπολογική απόσταση $\mathrm{dn}_{\|\cdot\|_2}^{\mathrm{top}}(\cdot)$, δηλαδή $\|A - N\|_2 = \mathrm{dn}_{\|\cdot\|_2}^{\mathrm{top}}(A)$. Αν θεωρήσουμε μία διαγωνοποίηση του N μέσω ενός ορθομοναδιαίου μετασχηματισμού ομοιότητας, $N = U_N \Lambda_N U_N^*$, και συμβολίσουμε με $D(U_N^*AU_N)$ τον $\nu \times \nu$ διαγώνιο πίνακα με κύρια διαγώνιο την κύρια διαγώνιο του $U_N^*AU_N$, τότε από την (5.2.1) παρατηρούμε ότι $\|U_N^*AU_N - \Lambda_N\|_2 = \mathrm{dn}_{\|\cdot\|_2}^{\mathrm{top}}(A)$, $\Lambda_N = D(U_N^*AU_N)$ και

$$\|N\|_2 = \|D(U_N^*AU_N)\|_2 \leq \|U_N^*AU_N\|_2 = \|A\|_2. \quad (5.2.2)$$

Ο υπολογισμός της απομάκρυνσης $\delta_{\|\cdot\|}(A)$ δεν είναι γενικά εύκολος αφού στο Λήμμα 1.4.2 του Schur, ο ορθομοναδιαίος πίνακας U δεν είναι μοναδικός. Επειδή όμως η νόρμα Frobenius $\|\cdot\|_F$ είναι ορθομοναδιαία αναλλοίωτη, έχουμε την ακόλουθη πρόταση.

Πρόταση 5.2.1. Για κάθε πίνακα $A \in \mathbb{C}^{\nu \times \nu}$ με (όχι κατ' ανάγκη διακεκριμένες) ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_\nu$, ισχύει

$$\delta_{\|\cdot\|_F}(A) = \sqrt{\|A\|_F^2 - \sum_{i=1}^{\nu} |\lambda_i|^2}.$$

Απόδειξη. Για κάθε ορθομοναδιαίο πίνακα $U \in \mathbb{C}^{\nu \times \nu}$ που δίνει την τριγωνοποίηση κατά Schur του A (Λήμμα 1.4.2), $A = U(\Lambda_U + R_U)U^*$, ισχύει

$$\begin{aligned} \|A\|_F^2 &= \|U^*AU\|_F^2 = \|\Lambda_U + R_U\|_F^2 \\ &= \|\Lambda_U\|_F^2 + \|R_U\|_F^2 = \sum_{i=1}^{\nu} |\lambda_i|^2 + \|R_U\|_F^2. \end{aligned}$$

Δηλαδή, $\|R_U\|_F^2 = \|A\|_F^2 - \sum_{i=1}^{\nu} |\lambda_i|^2$. □

Χρησιμοποιώντας τη γνωστή σχέση $\|\circ\|_2 \leq \|\circ\|_F \leq \sqrt{\nu} \|\circ\|_2$ της Πρότασης 1.3.4, εύκολα μπορεί κανείς να επαληθεύσει το ακόλουθο αποτέλεσμα.

Πρόταση 5.2.2. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Τότε ισχύουν οι ανισότητες

$$\mathrm{dn}_{\|\circ\|_2}^{\mathrm{top}}(A) \leq \mathrm{dn}_{\|\circ\|_F}^{\mathrm{top}}(A) \leq \sqrt{\nu} \mathrm{dn}_{\|\circ\|_2}^{\mathrm{top}}(A),$$

$$\mathrm{dn}_{\|\circ\|_2}^{\mathrm{com}}(A) \leq \mathrm{dn}_{\|\circ\|_F}^{\mathrm{com}}(A) \leq \sqrt{\nu} \mathrm{dn}_{\|\circ\|_2}^{\mathrm{com}}(A)$$

και

$$\delta_{\|\circ\|_2}(A) \leq \delta_{\|\circ\|_F}(A) \leq \sqrt{\nu} \delta_{\|\circ\|_2}(A).$$

Πρόταση 5.2.3. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Τότε ισχύει

$$\mathrm{dn}_{\|\circ\|_F}^{\mathrm{top}}(A) \leq \delta_{\|\circ\|_F}(A).$$

Απόδειξη. Θεωρούμε μία παραγοντοποίηση Schur (Λήμμα 1.4.2) του πίνακα A , $A = U T U^* = U(\Lambda_U + R_U)U^*$, όπου ο R_U είναι γνήσια άνω τριγωνικός με $\|R_U\|_F = \delta_{\|\circ\|_F}(A)$. Τότε έχουμε

$$\begin{aligned} \delta_{\|\circ\|_F}(A) &= \|R_U\|_F = \|U^* A U - \Lambda_U\|_F = \|A - U \Lambda_U U^*\|_F \\ &\geq \min \{ \|A - U \Lambda_U U^*\|_F : U^* U = U U^* = I_\nu \} = \mathrm{dn}_{\|\circ\|_F}^{\mathrm{top}}(A) \end{aligned}$$

και η απόδειξη ολοκληρώθηκε. $\square$

Πρόταση 5.2.4. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Τότε ισχύει

$$\mathrm{dn}_{\|\circ\|_F}^{\mathrm{com}}(A) \leq 2 \sqrt{\|A\|_2 \mathrm{dn}_{\|\circ\|_F}^{\mathrm{top}}(A)}.$$

Απόδειξη. Θεωρούμε έναν από τους πλησιέστερους στον A κανονικούς πίνακες, με βάση την απόσταση $\mathrm{dn}_{\|\circ\|_F}^{\mathrm{top}}(\cdot)$, έστω τον $N \in \mathbb{C}^{\nu \times \nu}$. Τότε με απλές πράξεις και εφαρμογή της τριγωνικής ανισότητας και της Άσκησης 1.5.9, έχουμε

$$\begin{aligned} \|A^* A - A A^*\|_F &= \|A^*(A - N) + (A - N)^* N - (A - N)N^* - A(A^* - N^*)\|_F \\ &\leq \|A^*\|_2 \|A - N\|_F + \|(A - N)^*\|_F \|N\|_2 \\ &\quad + \|A - N\|_F \|N^*\|_2 + \|A\|_2 \|(A - N)^*\|_F \\ &= 2 \|A\|_2 \|A - N\|_F + 2 \|N\|_2 \|A - N\|_F. \end{aligned}$$

Όμως, από τη σχέση (5.2.2), $\|N\|_2 \leq \|A\|_2$. Επομένως,

$$\|A^* A - A A^*\|_F \leq 4 \|A\|_2 \|A - N\|_F$$

και η απόδειξη ολοκληρώθηκε. $\square$

5.3 Παραδείγματα

Ολοκληρώνοντας το κεφάλαιο, δίνουμε τα παραδείγματα ενός κανονικού και ενός μη κανονικού πίνακα προκειμένου να επαληθεύσουμε αριθμητικά τα αποτελέσματα των δύο προηγούμενων παραγράφων.

Ας θεωρήσουμε τον 3×3 πίνακα

$$A = \begin{bmatrix} 1+i & 1 & 0 \\ 1 & 1+i & 0 \\ 0 & 0 & 1-i \end{bmatrix},$$

ο οποίος είναι κανονικός καθώς

$$A^*A = AA^* = \begin{bmatrix} 3 & 2 & 0 \\ 2 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix}.$$

Ο A είναι αντιστρέψιμος και ο αντίστροφος του πίνακας

$$A^{-1} = \begin{bmatrix} 0.2 - i0.6 & 0.2 + i0.4 & 0 \\ 0.2 + i0.4 & 0.2 - i0.6 & 0 \\ 0 & 0 & 0.5 + i0.5 \end{bmatrix}$$

είναι επίσης κανονικός, καθώς

$$(A^{-1})^*A^{-1} = A^{-1}(A^{-1})^* = \begin{bmatrix} 0.6 & -0.4 & 0 \\ -0.4 & 0.6 & 0 \\ 0 & 0 & 0.5 \end{bmatrix},$$

επαληθεύοντας την Πρόταση 5.1.2 (Συνθήκη 2). Η παραγοντοποίηση κατά Schur (Λήμμα 1.4.2) του πίνακα A οδηγεί στην διαγωνοποίηση

$$A = U \Lambda U^* = \begin{bmatrix} -\frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \\ -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} 2+i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & 1-i \end{bmatrix} \begin{bmatrix} -\frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ \frac{\sqrt{2}}{2} & -\frac{\sqrt{2}}{2} & 0 \\ 0 & 0 & 1 \end{bmatrix},$$

όπως αναμενόταν από την Πρόταση 5.1.5 (Συνθήκη 5). Το ερμιτιανό και αντιερμιτιανό μέρος του A ,

$$H(A) = \begin{bmatrix} 1 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{και} \quad S(A) = \begin{bmatrix} i & 0 & 0 \\ 0 & i & 0 \\ 0 & 0 & -i \end{bmatrix},$$

ικανοποιούν τη σχέση

$$H(A)S(A) = S(A)H(A) = \begin{bmatrix} i & i & 0 \\ i & i & 0 \\ 0 & 0 & -i \end{bmatrix},$$

επαληθεύοντας την Πρόταση 5.1.10 (Συνθήκη 10). Τέλος, στην πολική παραγοντοποίηση του A ,

$$\begin{aligned} A &= P W \\ &= \begin{bmatrix} 1.618 & 0.618 & 0 \\ 0.618 & 1.618 & 0 \\ 0 & 0 & \sqrt{2} \end{bmatrix} \begin{bmatrix} 0.4472 + i 0.7236 & 0.4472 - i 0.2764 & 0 \\ 0.4472 - i 0.2764 & 0.4472 + i 0.7236 & 0 \\ 0 & 0 & \frac{\sqrt{2}}{2}(1 - i) \end{bmatrix}, \end{aligned}$$

μπορεί κανείς να δει ότι $P W = W P = A$, επαληθεύοντας την Πρόταση 5.1.14 (Συνθήκη 14).

Για το δεύτερο μας παράδειγμα, θεωρούμε τον 4×4 πίνακα

$$B = \begin{bmatrix} 2 & 10 & 4 & 2 \\ 0 & -1 & -i 3 & 5 \\ 0 & 0 & 2 & i 4 \\ 0 & 0 & 0 & 1 \end{bmatrix}.$$

Από την Πρόταση 5.2.1, έχουμε ότι $\delta_{\|\cdot\|_F}(B) = 13.0384$, ενώ από την άνω τριγωνική μορφή του B προκύπτει ότι ο πλησιέστερος στον B κανονικός πίνακας, ως προς την απομάχρυνση $\delta_{\|\cdot\|_F}(\cdot)$, είναι ο διαγώνιος πίνακας $\text{diag}\{2, -1, 2, 1\}$. Εφαρμόζοντας τη σχέση (5.2.1), προσεγγίζουμε τις αποστάσεις

$$\text{dn}_{\|\cdot\|_2}^{\text{top}}(B) \cong 5.88 \quad \text{και} \quad \text{dn}_{\|\cdot\|_F}^{\text{top}}(B) \cong 8.74.$$

Επιπλέον, για την απόσταση από την κανονικότητα μέσω αντιμεταθέτη, έχουμε

$$\text{dn}_{\|\cdot\|_2}^{\text{com}}(B) = \|B^* B - B B^*\|_2^{1/2} = 11.1215$$

και

$$\text{dn}_{\|\cdot\|_F}^{\text{com}}(B) = \|B^* B - B B^*\|_F^{1/2} = 12.9378.$$

Προφανώς, επαληθεύονται οι ανισότητες της Πρότασης 5.2.2 για τις αποστάσεις $\text{dn}_{\|\cdot\|_2}^{\text{top}}(B)$, $\text{dn}_{\|\cdot\|_F}^{\text{top}}(B)$, $\text{dn}_{\|\cdot\|_2}^{\text{com}}(B)$ και $\text{dn}_{\|\cdot\|_F}^{\text{com}}(B)$. Επίσης, μπορεί να παρατηρήσει κανείς ότι

$$\text{dn}_{\|\cdot\|_F}^{\text{top}}(B) \cong 8.74 \leq 13.0384 = \delta_{\|\cdot\|_F}(B)$$

και

$$\text{dn}_{\|\cdot\|_F}^{\text{com}}(B) = 12.9378 \leq 2 \sqrt{11.2935 \cdot 8.74} \cong 2 \sqrt{\|B\|_2 \text{dn}_{\|\cdot\|_F}^{\text{top}}(B)}.$$

επαληθεύοντας τις Προτάσεις 5.2.3 και 5.2.4, αντίστοιχα.

5.4 Ασκήσεις

1. Να αποδείξετε την Πρόταση 5.1.11.
2. Να αποδείξετε την Πρόταση 5.1.12.
3. Να αποδείξετε την Πρόταση 5.1.15.
4. Να αποδείξετε την Πρόταση 5.1.16.
5. Να αποδείξετε ότι ένας πίνακας είναι ορθομοναδιαίος αν και μόνο αν είναι κανονικός και όλες του οι ιδιοτιμές έχουν μέτρο ίσο με 1.
6. Έστω ένας $\nu \times \nu$ κανονικός πίνακας $A = B + iC$ με $B, C \in \mathbb{R}^{\nu \times \nu}$.
Να αποδείξετε ότι ο $(2\nu) \times (2\nu)$, πραγματικός πίνακας $\begin{bmatrix} B & -C \\ C & B \end{bmatrix}$ είναι επίσης κανονικός.
7. Να αποδείξετε ότι ένας αντιστρέψιμος πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ είναι κανονικός αν και μόνο αν υπάρχει ορθομοναδιαίος πίνακας $U \in \mathbb{C}^{\nu \times \nu}$ τέτοιος ώστε $A^* = AU$.
8. Έστω ένας κανονικός πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Αν για έναν πίνακα $B \in \mathbb{C}^{\nu \times \nu}$, ισχύει $AB = BA$, να αποδείξετε ότι $A^*B = BA^*$.
9. Έστω ένας κανονικός πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ με όλες τις ιδιοτιμές του απλές. Να αποδείξετε ότι κάθε πίνακας $B \in \mathbb{C}^{\nu \times \nu}$ που ικανοποιεί τη σχέση $AB = BA$ είναι επίσης κανονικός.

Κεφάλαιο 6

Αριθμητικό Πεδίο Πίνακα

6.1 Ορισμός του Αριθμητικού Πεδίου

Το αριθμητικό πεδίο (*numerical range, field of values*) ενός τετραγωνικού πίνακα $A \in \mathbb{C}^{\nu \times \nu}$ ορίζεται ως το σύνολο

$$F(A) = \{x^* A x \in \mathbb{C} : x \in \mathbb{C}^\nu, x^* x = 1\}.$$

Τα αριθμητικά πεδία κάποιων απλών πινάκων είναι τα ακόλουθα:

1.

$$\begin{aligned} F(a I_\nu) &= \{x^* (a I_\nu) x : x \in \mathbb{C}^\nu, x^* x = 1\} \\ &= \{a (x^* x) : x \in \mathbb{C}^\nu, x^* x = 1\} \\ &= \{a\}. \end{aligned}$$

2.

$$\begin{aligned} F\left(\begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}\right) &= \left\{ [\bar{x}_1 \ \bar{x}_2] \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : |x_1|^2 + |x_2|^2 = 1 \right\} \\ &= \{|x_1|^2 : 0 \leq |x_1|^2 \leq 1\}, \end{aligned}$$

δηλαδή το αριθμητικό πεδίο είναι το κλειστό (πραγματικό) διάστημα $[0, 1]$.

3.

$$\begin{aligned} F\left(\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}\right) &= \left\{ [\bar{x}_1 \ \bar{x}_2] \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : |x_1|^2 + |x_2|^2 = 1 \right\} \\ &= \{2 x_2 \bar{x}_1 : |x_1|^2 + |x_2|^2 = 1\}, \end{aligned}$$

δηλαδή το αριθμητικό πεδίο ταυτίζεται με τον κλειστό μοναδιαίο (κυκλικό) δίσκο $\Delta(0, 1) = \{z \in \mathbb{C} : |z| \leq 1\}$.

Η αριθμητική ακτίνα (*numerical radius*) ενός τετραγωνικού πίνακα A ορίζεται ως

$$r(A) = \max \{|z| : z \in F(A)\}.$$

Όπως θα δείξουμε αργότερα, η $r(A)$ είναι διανυσματική νόρμα στους πίνακες αλλά δεν αποτελεί νόρμα πινάκων.

6.2 Βασικές Ιδιότητες του Αριθμητικού Πεδίου

Το αριθμητικό πεδίο ενός $\nu \times \nu$ μιγαδικού πίνακα παρουσιάζει πολλές και ενδιαφέρουσες ιδιότητες. Άλλες σχετίζονται με τη γεωμετρία του, ενώ άλλες με το γεγονός ότι το σύνολο αυτό είναι εικόνα μιας συνεχούς απεικόνισης από την επιφάνεια της μοναδιαίας σφαίρας του $\mathbb{C}^\nu$ στο μιγαδικό επίπεδο.

Πρόταση 6.2.1. (Συμπάγεια) Για κάθε πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, το $F(A)$ είναι ένα συμπαγές (δηλαδή, κλειστό και φραγμένο) υποσύνολο του $\mathbb{C}$.

Απόδειξη. Το $F(A)$, όπως ήδη παρατηρήσαμε, είναι η εικόνα της (Ευκλείδειας) μοναδιαίας σφαίρας $\{x \in \mathbb{C}^\nu : x^*x = 1\}$ στο μιγαδικό επίπεδο, μέσω της απεικόνισης $x \rightarrow x^*Ax$. Όμως η σφαίρα είναι συμπαγές (δηλαδή, κλειστό και φραγμένο) σύνολο και η εν λόγω απεικόνιση συνεχής. Επομένως, από το Θεώρημα του Weierstrass, το $F(A)$ είναι συμπαγές υποσύνολο του μιγαδικού επιπέδου (ως συνεχής εικόνα συμπαγούς συνόλου). $\square$

Πρόταση 6.2.2. Για κάθε $A \in \mathbb{C}^{\nu \times \nu}$ και $a \in \mathbb{C}$, ισχύουν

$$F(A + aI_\nu) = F(A) + a \quad \text{και} \quad F(aA) = aF(A).$$

Απόδειξη. Εύκολα βλέπουμε ότι

$$F(A + aI_\nu) = \{x^*Ax + x^*aI_\nu x : x \in \mathbb{C}^\nu, x^*x = 1\} = F(A) + a$$

και

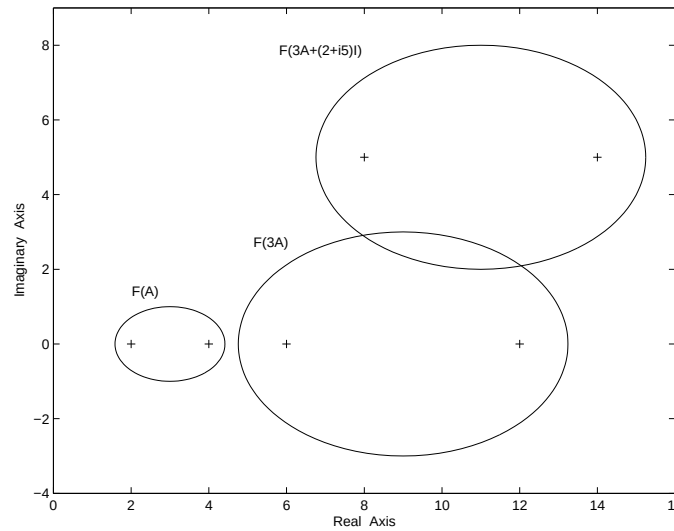
$$F(aA) = \{ax^*Ax : x \in \mathbb{C}^\nu, x^*x = 1\} = aF(A).$$

Η απόδειξη ολοκληρώθηκε. $\square$

Παράδειγμα 6.2.1. Θεωρούμε τον 2×2 πίνακα $A = \begin{bmatrix} 2 & 2 \\ 0 & 4 \end{bmatrix}$. Το αριθμητικό του πεδίο, όπως φαίνεται και στο Σχήμα 6.1, είναι κλειστός ελλειπτικός δίσκος με εστίες τις δύο ιδιοτιμές του, $\lambda_1 = 1$ και $\lambda_2 = 4$. Συγκεκριμένα στην επόμενη παράγραφο, θα δούμε ότι το αριθμητικό πεδίο ενός 2×2 πίνακα είναι πάντα έλλειψη με εστίες τις ιδιοτιμές του. Στο ίδιο σχήμα δίνονται και τα αριθμητικά πεδία των

πινάκων $3A = \begin{bmatrix} 6 & 6 \\ 0 & 12 \end{bmatrix}$ και $3A + (2 + i5)I_2 = \begin{bmatrix} 8 + i5 & 6 \\ 0 & 14 + i5 \end{bmatrix}$. Όπως είναι αναμενόμενο από την Πρόταση 6.2.2,

$$F(3A) = 3F(A) \quad \text{και} \quad F(3A + (2 + i5)I_2) = 3F(A) + 2 + i5.$$



Σχήμα 6.1: Τα αριθμητικά πεδία $F(A)$, $F(3A)$ και $F(3A + (2 + i5)I_2)$.

Πρόταση 6.2.3. Για έναν πίνακα $A = [a_{ij}] \in \mathbb{C}^{\nu \times \nu}$, ισχύει $F(A) = \{a\}$ αν και μόνο αν $A = aI_\nu$.

Απόδειξη. Το αντίστροφο έχειδειχθεί στην εισαγωγή του κεφαλαίου, στο πρώτο παράδειγμα. Για το ευθύ, από την Πρόταση 6.2.2, χωρίς βλάβη της γενικότητας, θεωρούμε ότι $a = 0$ και

$$x^* A x = \sum_{i,j=1}^{\nu} (a_{ij} \bar{x}_i x_j) = 0,$$

για κάθε διάνυσμα $x = [x_i] \in \mathbb{C}^\nu$ με $\sum_{i=1}^{\nu} |x_i| = 1$. Επιλέγοντας για x τα διανύσματα της κανονικής βάσης, προκύπτει άμεσα ότι τα διαγώνια στοιχεία του A είναι μηδενικά. Επιλέγοντας τώρα μοναδιαία διανύσματα x με $x_i = \frac{\sqrt{2}}{2}$ και $x_j = \frac{\sqrt{2}}{2}$, και με $x_i = \frac{\sqrt{2}}{2}$ και $x_j = i\frac{\sqrt{2}}{2}$, για κάθε $i \neq j$, βλέπουμε αντίστοιχα ότι $a_{ij} = -a_{ji}$ και $a_{ij} = a_{ji}$ ($i \neq j$). Επομένως, $a_{ij} = 0$ ($i \neq j$) και καταλήγουμε στο ότι $A = 0$. $\square$

Έστω ένας τετραγωνικός πίνακας A . Όπως και στο προηγούμενο κεφάλαιο, συμβολίζουμε με

$$H(A) = \frac{A + A^*}{2} \quad \text{και} \quad S(A) = \frac{A - A^*}{2}$$

το ερμιτιανό μέρος και αντιερμιτιανό μέρος του A , αντίστοιχα. Υπενθυμίζουμε ότι $A = H(A) + S(A)$, καθώς και ότι οι πίνακες $H(A)$ και $-iS(A)$ είναι ερμιτιανοί. Για ένα μιγαδικό αριθμό, το φανταστικό και το πραγματικό του μέρος είναι οι προβολές του στον φανταστικό και στον πραγματικό άξονα, αντίστοιχα. Κατά ανάλογο τρόπο, το αριθμητικό πεδίο του ερμιτιανού και αντιερμιτιανού μέρους του A είναι οι προβολές του $F(A)$ στον πραγματικό και το φανταστικό άξονα, αντίστοιχα.

Πρόταση 6.2.4. Για κάθε πίνακα $A \in \mathbb{C}^{n \times n}$, ισχύουν

$$F(H(A)) = \operatorname{Re}(F(A)) \quad \text{και} \quad F(S(A)) = i \operatorname{Im}(F(A)).$$

Απόδειξη. Για τυχαίο μοναδιαίο διάνυσμα $x \in \mathbb{C}^n$, έχουμε

$$x^* H(A) x = x^* \frac{A + A^*}{2} x = \frac{x^* A x + x^* A^* x}{2} = \frac{x^* A x + \overline{x^* A x}}{2} = \operatorname{Re}(x^* A x)$$

και

$$x^* S(A) x = x^* \frac{A - A^*}{2} x = \frac{x^* A x - x^* A^* x}{2} = \frac{x^* A x - \overline{x^* A x}}{2} = i \operatorname{Im}(x^* A x).$$

Η απόδειξη ολοκληρώθηκε. $\square$

Πόρισμα 6.2.5. Για έναν πίνακα $A \in \mathbb{C}^{n \times n}$, ισχύει ότι $F(A) \subset \mathbb{R}$ αν και μόνο αν ο A είναι ερμιτιανός.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πρόταση 6.2.6. Για κάθε πίνακα $A \in \mathbb{C}^{n \times n}$, το φάσμα του, $\sigma(A)$, περιέχεται στο αριθμητικό του πεδίο $F(A)$.

Απόδειξη. Έστω $\lambda \in \sigma(A)$ και $x \in \mathbb{C}^n$ ένα μοναδιαίο (δηλαδή, $x^* x = 1$) ιδιοδιάνυσμα του A που αντιστοιχεί στην ιδιοτιμή λ . Τότε ισχύει $\lambda = \lambda(x^* x) = x^*(\lambda x) = x^* A x \in F(A)$. $\square$

Η παραπάνω ιδιότητα μας δείχνει άμεσα ότι οι ιδιοτιμές ενός θετικά ορισμένου πίνακα είναι θετικοί αριθμοί. Πράγματι, έστω $A \in \mathbb{C}^{n \times n}$ θετικά ορισμένος (δηλαδή $x^* A x > 0$ για κάθε διάνυσμα $x \neq 0$) και (λ, x) ένα ιδιοζεύγος του. Τότε, αν $y = x/\|x\|$, έπεται ότι $\|y\|_2 = 1$ και $\lambda y^* y = y^*(\lambda y) = y^* A y \geq 0$.

Πρόταση 6.2.7. (Υπο-προσθετικότητα) Για κάθε $A, B \in \mathbb{C}^{\nu \times \nu}$, ισχύει

$$F(A + B) \subseteq F(A) + F(B).$$

Απόδειξη. Από τον ορισμό του αριθμητικού πεδίου, έχουμε

$$\begin{aligned} F(A + B) &= \{x^*(A + B)x : x \in \mathbb{C}^\nu, x^*x = 1\} \\ &= \{x^*Ax + x^*Bx : x \in \mathbb{C}^\nu, x^*x = 1\} \\ &\subseteq \{x^*Ax : x \in \mathbb{C}^\nu, x^*x = 1\} + \{y^*By : y \in \mathbb{C}^\nu, y^*y = 1\} \\ &= F(A) + F(B). \end{aligned}$$

Η απόδειξη ολοκληρώθηκε. □

Είναι γνωστό ότι ένας ορθομοναδιαίος μετασχηματισμός (δηλαδή, ένας μετασχηματισμός ισομετρίας) αφήνει αναλλοίωτη την επιφάνεια της μοναδιαίας σφαίρας του $\mathbb{C}^\nu$. Επομένως, θα αφήνει αναλλοίωτο και το αριθμητικό πεδίο. Πράγματι, έχουμε την παρακάτω ιδιότητα.

Πρόταση 6.2.8. Για κάθε $A, U \in \mathbb{C}^{\nu \times \nu}$, με U ορθομοναδιαίο, ισχύει

$$F(U^*AU) = F(A).$$

Απόδειξη. Εύκολα μπορεί κανείς να δει ότι

$$\begin{aligned} F(U^*AU) &= \{x^*(U^*AU)x : x \in \mathbb{C}^\nu, x^*x = 1\} \\ &= \{(Ux)^*A(Ux) : Ux \in \mathbb{C}^\nu, (Ux)^*(Ux) = 1\}. \end{aligned}$$

Η απόδειξη ολοκληρώθηκε. □

Βάσει της παραπάνω ιδιότητας, έχουμε την ακόλουθη σημαντική πρόταση.

Πρόταση 6.2.9. (Κανονικότητα) Για κάθε κανονικό πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, ισχύει $F(A) = \text{Co}(\sigma(A))$, δηλαδή το αριθμητικό πεδίο κανονικού πίνακα είναι η κυρτή θήκη των ιδιοτιμών του.

Απόδειξη. Ο κανονικός πίνακας A είναι ορθομοναδιαία όμοιος με τον διαγώνιο πίνακα $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_\nu\}$, όπου $\lambda_1, \lambda_2, \dots, \lambda_\nu$ είναι οι ιδιοτιμές του A (όχι κατ' ανάγκη διακεκριμένες). Δηλαδή, υπάρχει ορθομοναδιαίος πίνακας U

τέτοιος ώστε $A = U^* \Lambda U$. Έτσι, από την Πρόταση 6.2.8, έχουμε

$$\begin{aligned}
 F(A) &= F(U^* \Lambda U) = F(\Lambda) \\
 &= \{x^* \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_\nu\} x : x \in \mathbb{C}^\nu, x^* x = 1\} \\
 &= \left\{ \sum_{i=1}^{\nu} \bar{x}_i x_i \lambda_i : \sum_{i=1}^{\nu} \bar{x}_i x_i = 1 \right\} \\
 &= \left\{ \sum_{i=1}^{\nu} |x_i|^2 \lambda_i : \sum_{i=1}^{\nu} |x_i|^2 = 1 \right\} \\
 &= \text{Co}(\sigma(A))
 \end{aligned}$$

και η απόδειξη ολοκληρώθηκε. $\square$

Ως πόρισμα της προηγούμενης πρότασης, θα εξετάσουμε την περίπτωση όπου ο πίνακας A είναι ερμιτιανός, οπότε και έχει αποκλειστικά πραγματικές ιδιοτιμές.

Πόρισμα 6.2.10. Αν ένας $A \in \mathbb{C}^{\nu \times \nu}$ είναι ερμιτιανός, τότε το αριθμητικό πεδίο του είναι το κλειστό ευθύγραμμο τμήμα με άκρα την ελάχιστη και μέγιστη ιδιοτιμή του A .

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Παράδειγμα 6.2.2. Ας θεωρήσουμε τον άνω τριγωνικό πίνακα

$$A = \begin{bmatrix} 1 + i3 & -1 & 1 - i2 \\ 0 & 4 + i3 & 2 - i \\ 0 & 0 & 5 + i3 \end{bmatrix}$$

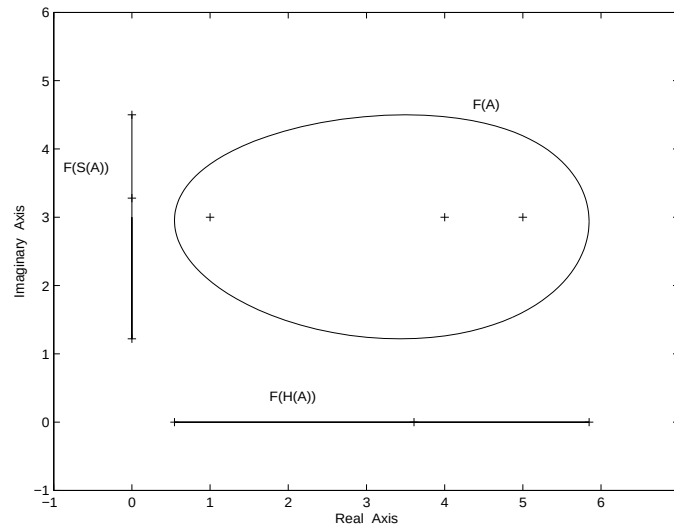
(γενικά μπορούμε, χωρίς βλάβη της γενικότητας, να χρησιμοποιούμε στα παραδείγματα μας τριγωνικούς πίνακες λόγω της τριγωνποίησης κατά Schur που δίνεται στο Λήμμα 1.4.2 και του αναλλοίωτου του αριθμητικού πεδίου στους ορθομοναδιαίους μετασχηματισμούς ομοιότητας). Στο Σχήμα 6.2, ο ωοειδής δίσκος είναι το $F(A)$. Το ερμιτιανό μέρος του A είναι

$$H(A) = \begin{bmatrix} 1 & -0.5 & 0.5 - i \\ -0.5 & 4 & 1 - i0.5 \\ 0.5 + i & 1 - i0.5 & 5 \end{bmatrix}$$

με ιδιοτιμές $\lambda_1 = 0.545$, $\lambda_2 = 3.608$ και $\lambda_3 = 5.864$. Το αντιερμιτιανό μέρος του πίνακα A ,

$$S(A) = \begin{bmatrix} i3 & -0.5 & 0.5 - i \\ 0.5 & i3 & 1 - i0.5 \\ -0.5 - i & -1 - i0.5 & i3 \end{bmatrix},$$

έχει ιδιοτιμές $\rho_1 = i1.219$, $\rho_2 = i3.28$ και $\rho_3 = i4.5$. Το $F(H(A))$ είναι η προβολή του $F(A)$ στον πραγματικό άξονα, ενώ το πεδίο $F(S(A))$ είναι η προβολή



Σχήμα 6.2: Τα αριθμητικά πεδία $F(A)$, $F(H(A))$ και $F(S(A))$.

του $F(A)$ στον φανταστικό άξονα, όπως άλλωστε υπαγορεύει και η Πρόταση 6.2.4. Ο $H(A)$ είναι ερμιτιανός πίνακας και το $F(H(A))$ είναι το κλειστό πραγματικό διάστημα με άκρα την ελάχιστη και τη μέγιστη ιδιοτιμή του, όπως γνωρίζουμε και από το Πόρισμα 6.2.10. Ο πίνακας $S(A)$ από την άλλη, είναι αντιερμιτιανός (άρα και κανονικός) με φανταστικές ιδιοτιμές, και από την Πρόταση 6.2.9, το αριθμητικό πεδίο $F(S(A))$ ταυτίζεται με την κυρτή θήκη αυτών των ιδιοτιμών.

Πόρισμα 6.2.11. Ένας ερμιτιανός πίνακας είναι θετικά ορισμένος αν και μόνο αν οι ιδιοτιμές του είναι θετικές.

Πόρισμα 6.2.12. Ένας ερμιτιανός πίνακας H είναι θετικά ορισμένος ή θετικά ημιορισμένος αν και μόνο αν $F(H) \subset (0, +\infty)$ ή $[0, +\infty)$, αντίστοιχα.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. □

Πόρισμα 6.2.13. Για τυχαίο πίνακα $A \in \mathbb{C}^{n \times n}$, ισχύουν τα ακόλουθα:

- (i) Το πεδίο $F(A)$ ανήκει στο δεξιό ανοιχτό ημιεπίπεδο του $\mathbb{C}$ αν και μόνο αν ο ερμιτιανός πίνακας $A + A^*$ είναι θετικά ορισμένος.
- (ii) Το πεδίο $F(A)$ ανήκει στο δεξιό κλειστό ημιεπίπεδο του $\mathbb{C}$ αν και μόνο αν ο ερμιτιανός πίνακας $A + A^*$ είναι θετικά ημιορισμένος.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. □

Για να παρουσιάσουμε την επόμενη ιδιότητα του αριθμητικού πεδίου, πρέπει να υπενθυμίσουμε ότι το ευθύ άθροισμα δύο πινάκων $A \in \mathbb{C}^{\nu_1 \times \nu_1}$ και $B \in \mathbb{C}^{\nu_2 \times \nu_2}$ ορίζεται ως ο $(\nu_1 + \nu_2) \times (\nu_1 + \nu_2)$ πίνακας

$$A \oplus B = \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}.$$

Πρόταση 6.2.14. Για κάθε $A \in \mathbb{C}^{\nu_1 \times \nu_1}$ και $B \in \mathbb{C}^{\nu_2 \times \nu_2}$, ισχύει

$$F(A \oplus B) = \text{Co}(F(A) \cup F(B)).$$

Απόδειξη. Αν $A \oplus B \in \mathbb{C}^{(\nu_1 + \nu_2) \times (\nu_1 + \nu_2)}$, τότε για κάθε μοναδιαίο $z \in \mathbb{C}^{\nu_1 + \nu_2}$ με $z = \begin{bmatrix} x \\ y \end{bmatrix}$ ($x \in \mathbb{C}^{\nu_1}, y \in \mathbb{C}^{\nu_2}$), προφανώς ισχύει $z^*(A \oplus B)z = x^*Ax + y^*By$. Εδώ $y^*y = 1$ αν και μόνο αν $x = 0$. Στην περίπτωση αυτή $z^*(A \oplus B)z = y^*By$, συνεπώς $F(B) \subseteq F(A \oplus B)$. Όμοια μπορεί κανείς να δει ότι $F(A) \subseteq F(A \oplus B)$. Από τις σχέσεις αυτές έπεται άμεσα ότι $F(A) \cup F(B) \subseteq F(A \oplus B)$.

Όπως θα δείξουμε παρακάτω σε ξεχωριστή παράγραφο, το αριθμητικό πεδίο ενός τετραγωνικού πίνακα είναι κυρτό υποσύνολο του μιγαδικού επιπέδου. Έτσι προκύπτει ότι $\text{Co}(F(A) \cup F(B)) \subseteq F(A \oplus B)$. Αρκεί λοιπόν να δείξουμε τον αντίθετο εγκλεισμό για να ολοκληρωθεί η απόδειξη. Έστω ένα μοναδιαίο διάνυσμα $z = \begin{bmatrix} x \\ y \end{bmatrix}$ ($x \in \mathbb{C}^{\nu_1}, y \in \mathbb{C}^{\nu_2}$). Τότε ισχύει

$$z^*(A \oplus B)z = x^*Ax + y^*By = x^*x \frac{x^*Ax}{x^*x} + y^*y \frac{y^*By}{y^*y} \in \text{Co}(F(A) \cup F(B)),$$

όπου $x^*x + y^*y = z^*z = 1$. Δηλαδή, ο αριθμός $x^*Ax + y^*By$ είναι κυρτός συνδυασμός στοιχείων των $F(A)$ και $F(B)$. Έτσι καταλήγουμε στη ζητούμενη ισότητα. $\square$

Από το πρώτο μέρος της παραπάνω απόδειξης προκύπτει άμεσα το ακόλουθο αποτέλεσμα.

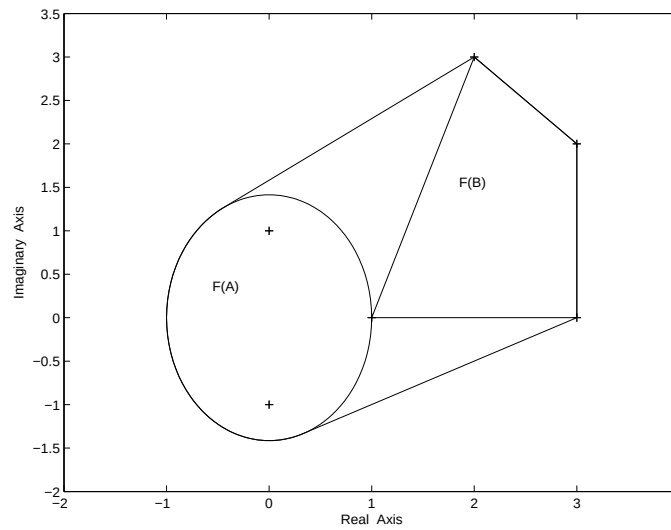
Πρόταση 6.2.15. Για κάθε κύριο υποπίνακα $\hat{A}$ ενός πίνακα $A \in \mathbb{C}^{\nu \times \nu}$ ισχύει $F(\hat{A}) \subseteq F(A)$.

Πρόταση 6.2.16. Έστω ένας τετραγωνικός πίνακας A .

- (i) Αν ο A είναι πραγματικός, τότε το $F(A)$ είναι συμμετρικό ως προς τον άξονα των πραγματικών αριθμών.
- (ii) Γενικά ισχύει ότι $F(A^*) = \overline{F(A)}$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Παράδειγμα 6.2.3. Έστω οι πίνακες $A = \begin{bmatrix} i & 2 \\ 0 & -i \end{bmatrix}$ και $B = \text{diag}\{1, 3, 3 + i2, 2 + i3\}$. Τα αριθμητικά τους πεδία φαίνονται στο Σχήμα 6.3. Το πεδίο $F(A)$ είναι ελλειπτικός δίσκος με εστίες τις ιδιοτιμές του A , $\lambda_1 = i$ και $\lambda_2 = -i$. Ο πίνακας B είναι κανονικός, αφού είναι διαγώνιος, κι επομένως το $F(B)$ είναι η κυρτή θήκη των ιδιοτιμών του σύμφωνα με την Πρόταση 6.2.9. Έτσι το $F(B)$ είναι το τετράπλευρο με κορυφές τις τέσσερις ιδιοτιμές του πίνακα B , $\rho_1 = 1$, $\rho_2 = 3$, $\rho_3 = 3 + i2$ και $\rho_4 = 2 + i3$. Το αριθμητικό πεδίο του πίνακα $F(A \oplus B)$



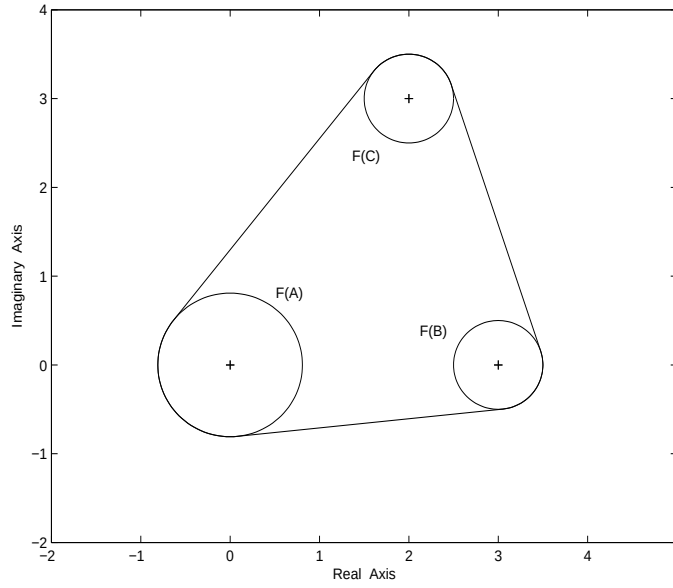
Σχήμα 6.3: Τα αριθμητικά πεδία $F(A)$, $F(B)$ και $F(A \oplus B)$.

είναι η κυρτή θήκη των $F(A)$ και $F(B)$, όπως υπαγορεύεται και από την Πρόταση 6.2.14.

Παράδειγμα 6.2.4. Έστω οι πίνακες (Jordan blocks)

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 3 & 1 \\ 0 & 3 \end{bmatrix} \quad \text{και} \quad C = \begin{bmatrix} 2 + i3 & 1 \\ 0 & 2 + i3 \end{bmatrix}.$$

Τα σύνολα $F(A)$, $F(B)$ και $F(C)$ φαίνονται στο Σχήμα 6.4. Το $F(A)$ είναι κυκλικός δίσκος με κέντρο το 0 και ακτίνα $a = \cos\left(\frac{\pi}{5}\right)$. Όπως θα δούμε στη συνέχεια, το αριθμητικό πεδίο ενός $\nu \times \nu$ Jordan block είναι κυκλικός δίσκος με κέντρο την ιδιοτιμή του και ακτίνα $a = \cos\left(\frac{\pi}{\nu+1}\right)$. Έτσι τα $F(B)$ και $F(C)$ είναι επίσης κυκλικοί δίσκοι με κέντρα τις ιδιοτιμές τους 3 και $2 + i3$, αντίστοιχα. Τα αριθμητικά πεδία $F(A)$ και $F(B)$ είναι συμμετρικά ως προς τον άξονα των πραγματικών αριθμών, επαληθεύοντας την Πρόταση 6.2.16. Το αριθμητικό πεδίο $F(A \oplus B \oplus C)$ είναι η κυρτή θήκη των $F(A)$, $F(B)$ και $F(C)$, όπως μπορούμε να δούμε και στο σχήμα, γεγονός αναμενόμενο από την Πρόταση 6.2.14.



Σχήμα 6.4: Τα αριθμητικά πεδία $F(A)$, $F(B)$, $F(C)$ και $F(A \oplus B \oplus C)$.

6.3 Κυρτότητα

Στην παράγραφο αυτή θα μελετήσουμε μία θεμελιώδη ιδιότητα του αριθμητικού πεδίου τετραγωνικών πινάκων, την κυρτότητα. Ας ξεκινήσουμε όμως βλέποντας τι συμβαίνει στη περίπτωση των 2×2 πινάκων.

Θεώρημα 6.3.1. (Ελλειπτικό Θεώρημα) Έστω $A \in \mathbb{C}^{2 \times 2}$ με ιδιοτιμές λ_1, λ_2 . Το αριθμητικό πεδίο του A είναι ένας κλειστός ελλειπτικός δίσκος με εστίες λ_1, λ_2 και μήκος μικρού άξονα ίσο με $\sqrt{\text{trace}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2}$.

Απόδειξη. Αν ο πίνακας A είναι κανονικός, τότε είναι ορθομοναδιαία όμοιος με τον διαγώνιο πίνακα $\text{diag}\{\lambda_1, \lambda_2\}$. Έτσι, από τις Προτάσεις 6.2.8 και 6.2.9,

$$F(A) = F(\text{diag}\{\lambda_1, \lambda_2\}) = \{\lambda_1|x_1|^2 + \lambda_2|x_2|^2 : x_1, x_2 \in \mathbb{C}, |x_1|^2 + |x_2|^2 = 1\}$$

είναι το ευθύγραμμο τμήμα με άκρα λ_1 και λ_2 , το οποίο θεωρείται ότι είναι ο εκφυλισμένος κλειστός ελλειπτικός δίσκος με εστίες λ_1, λ_2 και μήκος μικρού άξονα $0 = \sqrt{\text{trace}(\text{diag}\{\lambda_1, \lambda_2\}^* \text{diag}\{\lambda_1, \lambda_2\}) - |\lambda_1|^2 - |\lambda_2|^2}$.

Ας υποθέσουμε τώρα ότι ο πίνακας A δεν είναι κανονικός. Αν θεωρήσουμε τον πίνακα $A_0 = A - \frac{\text{trace}(A)}{2} I_2$, τότε $F(A_0) = F(A) - \frac{\text{trace}(A)}{2}$. Επιπλέον, από το Λήμμα 1.4.2 του Schur, ο πίνακας A_0 είναι ορθομοναδιαία όμοιος με έναν πίνακα της μορφής

$$B = \begin{bmatrix} \lambda_1 - \frac{\text{trace}(A)}{2} & b \\ 0 & \lambda_2 - \frac{\text{trace}(A)}{2} \end{bmatrix}$$

για κάποιο $b \in \mathbb{C}$. Επομένως,

$$F(B) = F(A_0) = F(A) - \frac{\text{trace}(A)}{2}.$$

Αν $\lambda_1 = \lambda_2$ (δηλαδή, ο A έχει μια διπλή ιδιοτιμή), τότε το αριθμητικό πεδίο του πίνακα $B = \begin{bmatrix} 0 & b \\ 0 & 0 \end{bmatrix}$ είναι ο κλειστός κυκλικός δίσκος

$$F(B) = \{b x_2 \bar{x}_1 : x_1, x_2 \in \mathbb{C}, |x_1|^2 + |x_2|^2 = 1\} = \Delta(0, |b|),$$

ο οποίος θεωρείται ότι είναι ο κλειστός ελλειπτικός δίσκος με εστίες στο 0 και μήκος μικρού άξονα ίσο με τη διάμετρο $|b| = \sqrt{\text{trace}(B^*B) - 0^2 - 0^2}$.

Αν $\lambda_1 \neq \lambda_2$ (δηλαδή, ο A έχει δύο απλές ιδιοτιμές), τότε ο B έχει δύο αντίθετες ιδιοτιμές, έστω a και $-a$. Διαιρώντας τον πίνακα B με a και εφαρμόζοντας κατάλληλο ορθομοναδιαίο μετασχηματισμό ομοιότητας (με διαγώνιο ορθομοναδιαίο πίνακα), χωρίς βλάβη της γενικότητας, μπορούμε να θεωρήσουμε ότι ο B είναι της μορφής

$$B = \begin{bmatrix} 1 & 2c \\ 0 & -1 \end{bmatrix}.$$

με $c > 0$. Επιπλέον, αν $H(B)$ και $S(B)$ είναι το ερμιτιανό και αντιερμιτιανό μέρος του B αντίστοιχα, παρατηρούμε ότι ο πίνακας

$$C = H(B) + \frac{\sqrt{1+c^2}}{c} S(B) = \begin{bmatrix} 1 & c + \sqrt{1+c^2} \\ c - \sqrt{1+c^2} & -1 \end{bmatrix}$$

(ο οποίος προκύπτει από μια “παραμόρφωση” του αντιερμιτιανού μέρους του B) έχει διπλή ιδιοτιμή το 0 και είναι ορθομοναδιαία όμοιος με τον

$$C_0 = \begin{bmatrix} 0 & 2\sqrt{1+c^2} \\ 0 & 0 \end{bmatrix}$$

Το αριθμητικό πεδίο $F(C_0)$ είναι κλειστός κυκλικός δίσκος με κέντρο στο 0 και ακτίνα $\sqrt{1+c^2}$, και ένας μιγαδικός αριθμός $t + is$ ($t, s \in \mathbb{R}$) ανήκει στο $F(B)$ αν και μόνο αν ο $t + i \frac{\sqrt{1+c^2}}{c} s$ ανήκει στο $F(C_0)$ ($= F(C)$). Αφού λοιπόν το σύνολο του $F(C_0)$ είναι

$$\partial F(C_0) = \{e^{i\theta} \sqrt{1+c^2} : \theta \in \mathbb{R}\},$$

έπεται ότι το σύνολο του $F(B)$ είναι η έλλειψη

$$\partial F(B) = \{\sqrt{1+c^2} \cos \theta + ic \sin \theta : \theta \in \mathbb{R}\}.$$

Επομένως, το πεδίο $F(B)$ είναι κλειστός ελλειπτικός δίσκος με εστίες 1 και -1 , μήκος μεγάλου άξονα $2\sqrt{1+c^2}$ και μήκος μικρού άξονα $2c$. Η απόδειξη ολοκληρώθηκε. $\square$

Η κυρτότητα λοιπόν του αριθμητικού πεδίου ενός 2×2 πίνακα προκύπτει από το γεγονός ότι αυτό είναι πάντα κλειστός ελλειπτικός δίσκος. Στη συνέχεια, προχωράμε στην απόδειξη της κυρτότητας του αριθμητικού πεδίου τετραγωνικών πινάκων οποιασδήποτε (πεπερασμένης) διάστασης.

Θεώρημα 6.3.2. (Toeplitz-Hausdorff) Για κάθε πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, το αριθμητικό πεδίο $F(A)$ είναι κυρτό.

Απόδειξη. Για να απόδειξουμε την κυρτότητα ενός συνόλου S , αρκεί να δείξουμε ότι $as + (1-a)t \in S$ για κάθε $s, t \in S$ και $a \in [0, 1]$. Επομένως, για έναν πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, το αριθμητικό πεδίο $F(A)$ είναι κυρτό αν και μόνο αν $ax^*Ax + (1-a)y^*Ay \in F(A)$ για κάθε $a \in [0, 1]$ και $x, y \in \mathbb{C}^\nu$ μοναδιαία διανύσματα. Όμως για κάθε ζεύγος γραμμικά ανεξαρτήτων μοναδιαίων διανυσμάτων $x, y \in \mathbb{C}^\nu$, υπάρχει ένας ορθομοναδιαίος πίνακας $U \in \mathbb{C}^{\nu \times \nu}$ και δύο μοναδιαία διανύσματα $\hat{x}, \hat{y} \in \mathbb{C}^\nu$ με όλα τα στοιχεία τους εκτός των δύο πρώτων να είναι μηδενικά, τέτοια ώστε $x = U\hat{x}$ και $y = U\hat{y}$. Αυτό αποδεικνύεται ως εξής: Έστω $x, y \in \mathbb{C}^\nu$ γραμμικά ανεξάρτητα διανύσματα και ο $\nu \times 2$ πίνακας $B = [x \ y]$. Από την παραγοντοποίηση *SVD* του B , έχουμε

$$B = U \begin{bmatrix} s_1 & 1 \\ 0 & s_2 \\ 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \end{bmatrix} V^*,$$

όπου ο πίνακας U είναι ορθομοναδιαίος $\nu \times \nu$, ο V είναι ορθομοναδιαίος 2×2 και $s_1 \geq s_2 > 0$ (αφού $\text{rank}(B) = 2$). Έτσι ο πίνακας B γράφεται στη μορφή

$$B = U \begin{bmatrix} \begin{bmatrix} s_1 & 0 \\ 0 & s_2 \end{bmatrix} V^* \\ 0_{(\nu-2) \times 2} \end{bmatrix}.$$

Δηλαδή, $B = [x \ y] = U\hat{B}$, όπου ο $\hat{B} = [\hat{x} \ \hat{y}]$ είναι ένας $\nu \times 2$ πίνακας με μηδενικά όλα τα στοιχεία του κάτω από τη δεύτερη γραμμή. Συνεπώς, έχουμε $x = U\hat{x}$ και $y = U\hat{y}$.

Επομένως, για κάθε γραμμικά ανεξάρτητα μοναδιαία διανύσματα $x, y \in \mathbb{C}^\nu$, ισχύει ότι

$$\begin{aligned} ax^*Ax + (1-a)y^*Ay &= a\hat{x}^*U^*AU\hat{x} + (1-a)\hat{y}^*U^*AU\hat{y} \\ &= a\hat{\xi}^*C\hat{\xi} + (1-a)\hat{\eta}^*C\hat{\eta}, \end{aligned}$$

όπου C είναι ο πρώτος 2×2 κύριος υποπίνακας του U^*AU και $\xi, \eta \in \mathbb{C}^2$ είναι τα διανύσματα με στοιχεία, τα δύο πρώτα στοιχεία των $\hat{x}, \hat{y}$. Από το Ελλειπτικό Θεώρημα 6.3.1 και την Πρόταση 6.2.8, προκύπτει η ζητούμενη κυρτότητα του αριθμητικού πεδίου. $\square$

Από την προηγούμενη παράγραφο ακόμα, είδαμε πως η ιδιότητα της κυρτότητας πέρα του γεγονότος ότι είναι η ίδια πολύ ενδιαφέρουσα, γίνεται ιδιαίτερα χρήσιμη και στην απόδειξη άλλων ιδιοτήτων.

Θεώρημα 6.3.3. Έστω ένας πίνακας $A \in \mathbb{C}^{n \times n}$. Τότε το 0 δεν ανήκει στο $F(A)$ αν και μόνο αν υπάρχει $\theta \in [0, 2\pi]$ τέτοιος ώστε ο ερμιτιανός πίνακας $H(e^{i\theta}A) = \frac{1}{2}(e^{i\theta}A + e^{-i\theta}A^*)$ να είναι θετικά ορισμένος.

Απόδειξη. Από το Πρόταση 6.2.4, το ευθύ προκύπτει άμεσα αφού αν ο $H(e^{i\theta}A)$ είναι θετικά ορισμένος για κάποιον $\theta \in [0, 2\pi]$, τότε το $F(e^{i\theta}A)$ βρίσκεται στο δεξιό ανοιχτό ημιεπίπεδο του $\mathbb{C}$. Επομένως, το 0 δεν ανήκει στο $F(e^{i\theta}A)$, άρα δεν ανήκει ούτε στο $F(A)$.

Για το αντίστροφο, υπενθυμίζουμε ότι από τα διαχωριστικά θεωρήματα, υπάρχει ευθεία L στο επίπεδο τέτοια ώστε καθένα από τα ξένα συμπαγή και κυρτά σύνολα $\{0\}$ και $F(A)$ να βρίσκονται εκατέρωθεν της L . Στρέφοντας τους άξονες κατάλληλα ώστε η L να συμπίπτει με μία κατακόρυφη ευθεία στο δεξιό ανοιχτό ημιεπίπεδο, προκύπτει ότι για τη συγκεκριμένη αυτή στροφή $\theta \in [0, 2\pi]$, το αριθμητικό πεδίο $F(e^{i\theta}A) = e^{i\theta}F(A)$ ανήκει στο δεξιό ανοιχτό ημιεπίπεδο. Άρα ο ερμιτιανός πίνακας $H(e^{i\theta}A) = \frac{1}{2}(e^{i\theta}A + e^{-i\theta}A^*)$ είναι θετικά ορισμένος. $\square$

Πόρισμα 6.3.4. Έστω ένας πίνακας $A \in \mathbb{C}^{n \times n}$. Τότε ισχύουν τα ακόλουθα:

- (i) $0 \in \text{Int}(F(A))$ αν και μόνο αν για κάθε $\theta \in [0, 2\pi]$, ο $H(e^{i\theta}A)$ είναι αόριστος.
- (ii) $0 \in \partial F(A)$ αν και μόνο αν υπάρχει $\theta_0 \in [0, 2\pi]$ τέτοιο ώστε ο πίνακας $H(e^{i\theta_0}A)$ να είναι μη αντιστρέψιμος και θετικά ημιορισμένος.
- (iii) $0 \notin F(A)$ αν και μόνο αν υπάρχει $\theta_0 \in [0, 2\pi]$ τέτοιο ώστε ο πίνακας $H(e^{i\theta_0}A)$ να είναι θετικά ορισμένος.

6.4 Αριθμητική Ακτίνα

Όπως αναφέραμε και στην αρχή του κεφαλαίου, η αριθμητική ακτίνα ενός πίνακα $A \in \mathbb{C}^{n \times n}$ ορίζεται ως $r(A) = \max\{|z| : z \in F(A)\}$. Στην παράγραφο αυτή παρουσιάζουμε κάποιες βασικές ιδιότητες της.

Πρόταση 6.4.1. Για κάθε $A \in \mathbb{C}^{n \times n}$, ισχύει $\frac{1}{2}\|A\|_2 \leq r(A) \leq \|A\|_2$.

Απόδειξη. Έχουμε

$$r(A) = \max_{\|x\|_2=1} |x^*Ax| \leq \max_{\|x\|_2=1} (\|Ax\|_2\|x\|_2) = \max_{\|x\|_2=1} (\|Ax\|_2) = \|A\|_2$$

και

$$\begin{aligned}\|A\|_2 &= \left\| \frac{A + A^*}{2} + \frac{A - A^*}{2} \right\|_2 \\ &\leq \left\| \frac{A + A^*}{2} \right\|_2 + \left\| \frac{A - A^*}{2} \right\|_2 \\ &= r\left(\frac{A + A^*}{2}\right) + r\left(\frac{A - A^*}{2}\right),\end{aligned}$$

όπου η τελευταία ισότητα ισχύει λόγω της κανονικότητας του ερμιτιανού και αντιερμιτιανού μέρους. Επιπλέον, έχουμε

$$\begin{aligned}r\left(\frac{A + A^*}{2}\right) + r\left(\frac{A - A^*}{2}\right) &\leq \frac{r(A) + r(A^*)}{2} + \frac{r(A) + r(A^*)}{2} \\ &= r(A) + r(A^*) = 2r(A),\end{aligned}$$

και η απόδειξη ολοκληρώθηκε. $\square$

Η αριθμητική ακτίνα είναι διανυσματική νόρμα (δηλαδή, χωρίς να ικανοποιεί την υπο-πολλαπλασιαστική ιδιότητα), αλλά δεν είναι νόρμα πινάκων.

Πρόταση 6.4.2. Η αριθμητική ακτίνα στο διανυσματικό χώρο των τετραγωνικών πινάκων είναι διανυσματική νόρμα.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πρόταση 6.4.3. Η νόρμα $4r(\cdot)$ είναι νόρμα πινάκων στο $\mathbb{C}^{n \times n}$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πρόταση 6.4.4. Έστω ένας πίνακας $A \in \mathbb{C}^{n \times n}$ με φασματική ακτίνα

$$\rho(A) = \max\{|\lambda| : \lambda \in \sigma(A)\}.$$

Τότε ισχύει $\rho(A) \leq r(A)$.

Απόδειξη. Προκύπτει άμεσα από το γεγονός ότι το φάσμα ενός πίνακα περιέχεται στο αριθμητικό του πεδίο (βλέπε την Πρόταση 6.2.6). $\square$

Ολοκληρώνοντας την παράγραφο, θα προχωρήσουμε στη παρουσίαση μιας ιδιότητας της αριθμητικής ακτίνας η οποία φαίνεται πως έχει εφαρμογές στις επαναληπτικές αριθμητικές μεθόδους των πινάκων. Η απόδειξη είναι τεχνική και παραλείπεται.

Πρόταση 6.4.5. (Ανισότητα δυνάμεων) Έστω ένας πίνακας $A \in \mathbb{C}^{n \times n}$. Τότε για κάθε $n \in \mathbb{N}$, ισχύει $r(A^n) \leq r(A)^n$.

6.5 Προσέγγιση του Αριθμητικού Πεδίου

Γενικά, για ένα σημείο a του αριθμητικού πεδίου $F(A)$, δεν μπορούμε να υπολογίσουμε (όλα) τα μοναδιαία διανύσματα $x \in \mathbb{C}^\nu$ που ικανοποιούν τη σχέση $x^* A x = a$. Κάτι τέτοιο είναι εφικτό μόνο στην περίπτωση “ακραίων” συνοριακών σημείων.

Λήμμα 6.5.1. *Αν $A \in \mathbb{C}^{\nu \times \nu}$ και $x_0 \in \mathbb{C}^\nu$ με $x_0^* x_0 = 1$, τότε τα ακόλουθα είναι ισοδύναμα:*

$$(i) \quad \operatorname{Re}(x_0^* A x_0) = \max\{\operatorname{Re}(a) : a \in F(A)\},$$

$$(ii) \quad x_0^* H(A) x_0 = \max\{b : b \in F(H(A))\},$$

$$(iii) \quad H(A) x_0 = \lambda_{\max}(H(A)) x_0,$$

όπου $\lambda_{\max}(H(A))$ είναι η μεγαλύτερη ιδιοτιμή του ερμιτιανού πίνακα $H(A)$.

Απόδειξη. Η ισοδυναμία των (i) και (ii) έπεται από το γεγονός ότι

$$\operatorname{Re}(x^* A x) = \frac{1}{2} (x^* A x + \overline{x^* A x}) = \frac{1}{2} (x^* A x + x^* A^* x) = x^* H(A) x.$$

Έστω τώρα $\{y_1, y_2, \dots, y_n\}$ μία βάση μοναδιαίων ιδιοδιανυσμάτων του $H(A)$ που αντιστοιχούν στις ιδιοτιμές $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_\nu$ του $H(A)$, δηλαδή $H(A) y_i = \lambda_i y_i$ ($i = 1, 2, \dots, \nu$). Τότε το διάνυσμα x_0 μπορεί να γραφεί στη μορφή

$$x_0 = \sum_{i=1}^{\nu} c_i y_i, \quad \text{με} \quad \sum_{i=1}^{\nu} \bar{c}_i c_i = 1,$$

ώστε $x_0^* x_0 = 1$. Επομένως, η ποσότητα

$$x_0^* H(A) x_0 = \sum_{i=1}^{\nu} \bar{c}_i c_i \lambda_i = \sum_{i=1}^{\nu} |c_i|^2 \lambda_i$$

λαμβάνει τη μέγιστη δυνατή τιμή της όταν $c_1 = 1$ και $c_2 = \dots = c_\nu = 0$, δίνοντας έτσι την ισοδυναμία μεταξύ των (i) και (ii). $\square$

Από το προηγούμενο λήμμα, επαληθεύεται ότι

$$\max\{\operatorname{Re}(a) : a \in F(A)\} = \max\{h : h \in F(H(A))\} = \lambda_{\max}(H(A)).$$

Αυτό σημαίνει ότι το “δεξιότερο” σημείο του $F(A)$ έχει πραγματικό μέρος τη μέγιστη ιδιοτιμή του ερμιτιανού μέρους του πίνακα A , $\lambda_{\max}(H(A))$. Ουσιαστικά, το λήμμα λέει ότι αν υπολογισθεί η $\lambda_{\max}(H(A))$ και ένα μοναδιαίο ιδιοδιάνυσμα

της x , τότε το “δεξιότερο” συνοριακό σημείο του $F(A)$ είναι το x^*Ax . Επιπλέον, η ευθεία $\{\lambda_{\max}(H(A)) + it : t \in \mathbb{R}\}$ είναι εφαπτομένη του κυρτού συνόλου $F(A)$ σε αυτό το συνοριακό σημείο. Χρησιμοποιώντας λοιπόν τη σχέση $e^{-i\theta} F(e^{i\theta} A) = F(A)$, είναι δυνατόν να πάρουμε όσα συνοριακά σημεία και όσες εφαπτόμενες ευθείες θέλουμε, στρέφοντας το $F(A)$ και πραγματοποιώντας τον απαιτούμενο υπολογισμό των ιδιοζευγών.

Για μία γωνία $\theta \in [0, 2\pi]$, ορίζουμε $\lambda_\theta = \lambda_{\max}(H(e^{i\theta} A))$ και $x_\theta \in \mathbb{C}^n$ ένα αντίστοιχο μοναδιαίο ιδιοδιάνυσμα. Δηλαδή, $H(e^{i\theta} A)x_\theta = \lambda_\theta x_\theta$ με $x_\theta^* x_\theta = 1$. Συμβολίζουμε επίσης με l_θ την εφαπτόμενη ευθεία $\{e^{-i\theta}(\lambda_\theta + it) : t \in \mathbb{R}\}$ και συμβολίζουμε το ημιεπίπεδο που ορίζεται από την l_θ και περιέχει το πεδίο $F(A)$, με $H_\theta = e^{-i\theta}\{z : \operatorname{Re}(z) \leq \lambda_\theta\}$. Βάσει των παραπάνω έχουμε το ακόλουθο αποτέλεσμα.

Θεώρημα 6.5.2. Για κάθε $A \in \mathbb{C}^{n \times n}$ και κάθε $\theta \in [0, 2\pi]$, ο μιγαδικός αριθμός $p_\theta = x_\theta^* A x_\theta$ είναι ένα συνοριακό σημείο του $F(A)$. Η ευθεία l_θ είναι εφαπτομένη του $F(A)$ με $p_\theta \in l_\theta \cap F(A)$ και $F(A) \subset H_\theta$ για κάθε $\theta \in [0, 2\pi]$.

Επειδή το $F(A)$ είναι κυρτό, γίνεται εποπτικά φανερό ότι κάθε συνοριακό σημείο του εμφανίζεται ως ένα p_θ και για κάθε $a \notin F(A)$ υπάρχει μία ευθεία l_θ που αφήνει το $F(A)$ και το A σε διαφορετικά ημιεπίπεδα, δηλαδή $a \notin H_\theta$ ενώ $F(A) \in H_\theta$. Έτσι, το αριθμητικό πεδίο μπορεί να γραφεί ως άπειρη τομή κλειστών ημιεπιπέδων, όπως φαίνεται στο παρακάτω θεώρημα.

Θεώρημα 6.5.3. Για κάθε $A \in \mathbb{C}^{n \times n}$,

$$F(A) = \operatorname{Co}(\{p_\theta : 0 \leq \theta \leq 2\pi\}) = \bigcap_{0 \leq \theta \leq 2\pi} H_\theta.$$

Αφού δε γίνεται να υπολογίσουμε άπειρα σημεία p_θ και άπειρες ευθείες l_θ , απλά αρκούμαστε σε ένα διακριτό ανάλογο του τελευταίου θεωρήματος, επιλέγοντας μία διαμέριση του $[0, 2\pi]$, έστω

$$\Theta = \{\theta_1, \theta_2, \dots, \theta_k\}, \quad \text{όπου } 0 = \theta_1 \leq \theta_2 \leq \dots \leq \theta_k = 2\pi.$$

Αν p_{θ_i} ($i = 1, 2, \dots, k$) είναι τα διαδοχικά συνοριακά σημεία του $F(A)$ και H_{θ_i} ($i = 1, 2, \dots, k$) είναι τα αντίστοιχα ημιεπίπεδα που ορίζονται από τις εφαπτόμενες ευθείες l_{θ_i} ($i = 1, \dots, k$) του $F(A)$ στα σημεία p_{θ_i} , τότε ορίζουμε το εγγεγραμμένο στο $F(A)$ πολύγωνο $F_{in}(A, \Theta) = \operatorname{Co}(\{p_{\theta_1}, \dots, p_{\theta_k}\})$ και το περιγεγραμμένο στο $F(A)$ πολύγωνο $F_{out}(A, \Theta) = H_{\theta_1} \cap H_{\theta_2} \cap \dots \cap H_{\theta_k}$. Τα δύο αυτά σύνολα αποτελούν μία εσωτερική και μία εξωτερική προσέγγιση για το $F(A)$, αντίστοιχα, αφού είναι προφανές ότι για κάθε διαμέριση $\Theta = \{\theta_1, \theta_2, \dots, \theta_k\}$, ισχύει

$$F_{in}(A, \Theta) \subseteq F(A) \subseteq F_{out}(A, \Theta).$$

6.6 Σύνορο του Αριθμητικού Πεδίου

Το αριθμητικό πεδίο ενός 2×2 πίνακα $A = \begin{bmatrix} \lambda_1 & b \\ 0 & \lambda_2 \end{bmatrix}$, είναι

- ένα σημείο του μιγαδικού επιπέδου αν και μόνο αν $\lambda_1 = \lambda_2$ και $b = 0$,
- το κλειστό ευθύγραμμο τμήμα με άκρα τα λ_1, λ_2 αν και μόνο αν $\lambda_1 \neq \lambda_2$ και $b = 0$,
- ένας κλειστός κυκλικός δίσκος με ακτίνα $|b|/2$ αν και μόνο αν $\lambda_1 = \lambda_2$,
- ένας (μη εκφυλισμένος) κλειστός ελλειπτικός δίσκος με εστίες τις λ_1, λ_2 αν και μόνο αν $\lambda_1 \neq \lambda_2$ και $b \neq 0$.

Τι γίνεται όμως για πίνακες μεγαλύτερης διάστασης; Σε μεγαλύτερες διαστάσεις συναντάμε μία πολύ πιο πλούσια ποικιλία σχημάτων και μορφών για το αριθμητικό πεδίο τετραγωνικών πινάκων. Από την Πρόταση 6.2.9, κάθε κυρτό πολύγωνο είναι το αριθμητικό πεδίο ενός κανονικού πίνακα του οποίου οι ιδιοτιμές είναι οι κορυφές του πολυγώνου αυτού. Έτσι, κάθε φραγμένο και κυρτό σύνολο μπορεί να προσεγγιστεί από το αριθμητικό πεδίο ενός τετραγωνικού πίνακα του οποίου η διάσταση μπορεί να είναι αυθαίρετα μεγάλη.

Στην παράγραφο αυτή, βασικός στόχος μας είναι να παρουσιάσουμε κάποια στοιχεία για τη μορφή του αριθμητικού πεδίου ενός τετραγωνικού πίνακα, καθώς και τη σχέση του συνόρου του $F(A)$, $\partial F(A)$, με το φάσμα του A . Ένα $a \in \partial F(A)$ καλείται *γωνιακό σημείο* (*angular point*, *sharp point*) του $F(A)$ αν υπάρχουν γωνίες θ_1 και θ_2 με $0 \leq \theta_1 \leq \theta_2 < 2\pi$ και $\theta_1 - \theta_2 < \pi$, για τις οποίες $\operatorname{Re}(e^{i\theta}a) = \max\{\operatorname{Re}(b) : b \in F(e^{i\theta}A)\}$ για κάθε $\theta \in (\theta_1, \theta_2)$.

Θεώρημα 6.6.1. Έστω ένας πίνακας $A \in \mathbb{C}^{n \times n}$. Αν $a \in \mathbb{C}$ είναι ένα γωνιακό σημείο του $F(A)$, τότε το a είναι ιδιοτιμή του A .

Απόδειξη. Αν $a \in \mathbb{C}$ είναι ένα γωνιακό σημείο του $F(A)$, τότε γνωρίζουμε από το Λήμμα 6.5.1 ότι υπάρχει ένα μοναδιαίο διάνυσμα $x_a \in \mathbb{C}^n$, ανεξάρτητο του θ , τέτοιο ώστε

$$x_a^* H(e^{i\theta}A) x_a = x_a^* \lambda_{\max}(H(e^{i\theta}A)) x_a = \lambda_{\max}(H(e^{i\theta}A))$$

για κάθε $\theta \in (\theta_1, \theta_2)$. Αφού ο $H(e^{i\theta}A)$ είναι ερμιτιανός, συμπεραίνουμε ότι για κάθε $\theta \in (\theta_1, \theta_2)$, ισχύει $H(e^{i\theta}A) x_a = \lambda_{\theta} x_a$, όπου $\lambda_{\theta} = \lambda_{\max}(H(e^{i\theta}A))$. Παραγωγίζοντας την τελευταία σχέση ως προς θ , έχουμε

$$H(i e^{i\theta}A) x_a = \frac{d \lambda_{\theta}}{d \theta} x_a,$$

ή ισοδύναμα,

$$S(e^{i\theta} A) x_a = -i \frac{d\lambda_\theta}{d\theta} x_a.$$

Προσθέτοντας τώρα κατά μέλη τη τελευταία σχέση με την $H(e^{i\theta} A) x_a = \lambda_\theta x_a$, βλέπουμε ότι

$$H(e^{i\theta} A) x_a + S(e^{i\theta} A) x_a = \left(\lambda_\theta - i \frac{d\lambda_\theta}{d\theta} \right) x_a.$$

Συνεπώς,

$$e^{i\theta} A x_a = \left(\lambda_\theta - i \frac{d\lambda_\theta}{d\theta} \right) x_a,$$

ή ισοδύναμα,

$$A x_a = e^{-i\theta} \left(\lambda_\theta - i \frac{d\lambda_\theta}{d\theta} \right) x_a.$$

Αφού

$$a = x_a^* A x_a = x_a^* e^{-i\theta} \left(\lambda_\theta - i \frac{d\lambda_\theta}{d\theta} \right) x_a = e^{-i\theta} \left(\lambda_\theta - i \frac{d\lambda_\theta}{d\theta} \right),$$

καταλήγουμε στο ότι το γωνιακό σημείο a είναι ιδιοτιμή του πίνακα A . $\square$

Πόρισμα 6.6.2. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Το πεδίο $F(A)$ έχει το πολύ ν γωνιακά σημεία και είναι κυρτό πολύγωνο αν και μόνο αν $F(A) = \text{Co}(\sigma(A))$.

Πρέπει να αναφέρουμε βέβαια ότι αν και κάθε γωνιακό σημείο του αριθμητικού πεδίου είναι ιδιοτιμή στο σύνορο του αριθμητικού πεδίου, το αντίστροφο δεν ισχύει.

Μια ιδιοτιμή $\lambda \in \sigma(A)$ καλείται *κανονική* αν η γεωμετρική της πολλαπλότητα ισούται με την αλγεβρική της πολλαπλότητα και κάθε ιδιοδιάνυσμα του A που αντιστοιχεί στη λ είναι κάθετο σε κάθε ιδιοδιάνυσμα άλλης ιδιοτιμής. Ακολουθεί ένα θεώρημα που μας εξασφαλίζει τη κανονικότητα κάθε ιδιοτιμής που βρίσκεται στο σύνορο του αριθμητικού πεδίου ενός τετραγωνικού πίνακα.

Θεώρημα 6.6.3. Αν $A \in \mathbb{C}^{\nu \times \nu}$ και $a \in \partial F(A) \cap \sigma(A)$, τότε η ιδιοτιμή a είναι κανονική ιδιοτιμή του πίνακα A . Επιπλέον, αν m είναι η πολλαπλότητα της a , τότε ο A είναι ορθομοναδιαία όμοιος με έναν πίνακα της μορφής $a I_m \oplus B$, όπου $B \in \mathbb{C}^{(\nu-m) \times (\nu-m)}$ και $a \notin \sigma(B)$.

Απόδειξη. Αν η αλγεβρική πολλαπλότητα της ιδιοτιμής a είναι m , τότε σύμφωνα με το Λήμμα 1.4.2 του Schur, ο A είναι ορθομοναδιαία όμοιος με έναν άνω τριγωνικό πίνακα T του οποίου τα πρώτα m διαγώνια στοιχεία είναι ίσα με a και τα υπόλοιπα είναι οι υπόλοιπες ιδιοτιμές του A . Υποθέτουμε ότι ο T έχει μη μηδενικό στοιχείο εκτός της κυρίας διαγωνίου σε μία από τις m πρώτες γραμμές του.

Τότε υπάρχει ένας κύριος 2×2 υποπίνακας $T_0 = \begin{bmatrix} a & c \\ 0 & b \end{bmatrix}$, με $c \neq 0$. Όμως, το πεδίο $F(T_0)$ είναι ή κλειστός κυκλικός δίσκος με ακτίνα $|c|/2$ (όταν $a = b$) ή μη εκφυλισμένος κλειστός έλλειπτικός δίσκος με εστίες a, b . Έτσι, έχουμε ότι το a είναι εσωτερικό σημείο του $F(T_0)$. Όμως, $F(T_0) \subseteq F(T) = F(A)$, από τις Προτάσεις 6.2.8 και 6.2.15. Δηλαδή, το σημείο a είναι εσωτερικό του $F(A)$, άτοπο αφού το a είναι συνοριακό σημείο του $F(A)$. Από το άτοπο αυτό καταλήγουμε στο συμπέρασμα ότι δεν υπάρχουν μη μηδενικά στοιχεία εκτός της κυρίας διαγωνίου στις m πρώτες γραμμές του T . Έτσι, $T = a I_m \oplus B$, $B \in \mathbb{C}^{(\nu-m) \times (\nu-m)}$. Τα υπόλοιπα συμπεράσματα του θεωρήματος προκύπτουν άμεσα. $\square$

Πόρισμα 6.6.4. Αν $\nu - 1$ ιδιοτιμές του πίνακα $A \in \mathbb{C}^{\nu \times \nu}$ (συνυπολογίζοντας και τις αλγεβρικές πολλαπλότητες) ανήκουν στο σύνορο του $F(A)$, τότε ο πίνακας A είναι κανονικός.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πόρισμα 6.6.5. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Τότε $F(A) = \text{Co}(\sigma(A))$ αν και μόνο αν ο A είναι κανονικός ή είναι ορθομοναδιαία όμοιος με έναν πίνακα της μορφής $\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix}$, όπου ο A_1 είναι κανονικός και $F(A_2) \subset F(A_1)$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πόρισμα 6.6.6. Αν $A \in \mathbb{C}^{\nu \times \nu}$ και $\nu \leq 4$, τότε ο A είναι κανονικός αν και μόνο αν $F(A) = \text{Co}(\sigma(A))$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

6.7 Πίνακες Jordan και Ομοιότητα

Στην παράγραφο αυτή, θα αναφερθούμε στα αριθμητικά πεδία των πινάκων Jordan, καθώς και στις σχέσεις που συνδέουν το αριθμητικό πεδίο ενός τετραγωνικού πίνακα με τα αριθμητικά πεδία των ομοίων του πινάκων. Αρχικά όμως, πρέπει (ως σύντομη υπενθύμιση) να αναφερθούμε στον τρόπο κατασκευής της κανονικής μορφής Jordan ενός τετραγωνικού πίνακα.

Έστω $A \in \mathbb{C}^{\nu \times \nu}$ και λ_0 μία ιδιοτιμή του. Ένα μη μηδενικό διάνυσμα $x \in \mathbb{C}^\nu$ λέγεται γενικευμένο ιδιοδιάνυσμα τάξης k του πίνακα A που αντιστοιχεί στην ιδιοτιμή λ_0 όταν ο k είναι ο ελάχιστος φυσικός αριθμός για τον οποίο ισχύει $(A - \lambda_0 I_\nu)^k = 0$, ενώ $(A - \lambda_0 I_\nu)^{k-1} \neq 0$. Από ένα γενικευμένο ιδιοδιάνυσμα x τάξης k , δημιουργείται μία αλυσίδα διανυσμάτων

$$x_k = x, x_{k-1} = (A - \lambda_0 I_\nu) x_k, x_{k-2} = (A - \lambda_0 I_\nu)^2 x_k, \dots, x_1 = (A - \lambda_0 I_\nu)^{k-1} x_k.$$

Έτσι, κατασκευάζουμε μία αλυσίδα *Jordan* $x_1, x_2, \dots, x_{k-1}, x_k$ μήκους k , της οποίας τα διανύσματα είναι γραμμικά ανεξάρτητα. Το x_1 είναι ιδιοδιάνυσμα του A που αντιστοιχεί στην ιδιοτιμή $\lambda_0 \in \sigma(A)$. Τα υπόλοιπα διανύσματα, $x_2, x_3, \dots, x_k$, της αλυσίδας λέγονται *γενικευμένα ιδιοδιανύσματα* και ικανοποιούν τις σχέσεις

$$A x_r = x_{r-1} + \lambda_0 x_r, \quad r = 1, 2, \dots, k.$$

Επομένως,

$$\begin{aligned} A [x_1 \ x_2 \ \cdots \ x_k] &= [A x_1 \ A x_2 \ \cdots \ A x_k] \\ &= [\lambda_0 x_1 \ \lambda_0 x_2 + x_1 \ \cdots \ \lambda_0 x_k + x_{k-1}] \\ &= [x_1 \ x_2 \ \cdots \ x_k] \begin{bmatrix} \lambda_0 & 1 & \cdots & 0 \\ 0 & \lambda_0 & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda_0 \end{bmatrix}, \end{aligned}$$

όπου ο $k \times k$ πίνακας

$$J_k(\lambda_0) = \begin{bmatrix} \lambda_0 & 1 & \cdots & 0 \\ 0 & \lambda_0 & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & \lambda_0 \end{bmatrix}$$

είναι ένα *Jordan block* της ιδιοτιμής λ_0 .

Στη συνέχεια, προχωράμε στη διατύπωση του γνωστού θεωρήματος που αφορά την παραγοντοποίηση (κανονική μορφή) *Jordan* των τετραγωνικών μιγαδικών πινάκων. Σημειώνουμε ότι *πίνακας Jordan* ονομάζεται κάθε block-διαγώνιος πίνακας $J \in \mathbb{C}^{n \times n}$ που έχει υποπίνακες *Jordan* (*Jordan blocks*) στη διαγώνιο του.

Θεώρημα 6.7.1. Για κάθε $A \in \mathbb{C}^{n \times n}$, υπάρχει αντιστρέψιμος $S \in \mathbb{C}^{n \times n}$ τέτοιος ώστε $S^{-1} A S = J$, όπου $J \in \mathbb{C}^{n \times n}$ είναι ένας πίνακας *Jordan* που αντιστοιχεί στον A .

Θεώρημα 6.7.2. Αν $J_k(\lambda_0)$ είναι το $k \times k$ *Jordan block* της ιδιοτιμής λ_0 , τότε το αριθμητικό του πεδίο είναι $F(J_k(\lambda_0)) = \Delta(\lambda_0, \cos(\frac{\pi}{k+1}))$, δηλαδή ο κυκλικός δίσκος με κέντρο το λ_0 και ακτίνα $\cos(\frac{\pi}{k+1})$.

Το θεώρημα αυτό, του οποίου η απόδειξη είναι τεχνική και παραλείπεται, μας δείχνει τη μορφή που έχει το αριθμητικό πεδίο ενός πίνακα *Jordan* (βλέπε το Παράδειγμα 6.2.4). Ας δούμε όμως τι ακριβώς συμβαίνει και τι σχέση υπάρχει μεταξύ των αριθμητικών πεδίων όμοιων πινάκων. Για δύο όμοιους πίνακες $A, B \in \mathbb{C}^{n \times n}$ με αριθμητικά πεδία $F(A)$ και $F(B)$, είναι φανερό ότι ισχύει

$$\text{Co}(\sigma(A)) = \text{Co}(\sigma(B)) \subseteq F(A) \cap F(B)$$

λόγω της κυρτότητας του αριθμητικού πεδίου και του γεγονότος ότι το φάσμα περιέχεται σε αυτό. Γενικά, τα αριθμητικά πεδία όμοιων πινάκων δεν παρουσιάζουν κάποια άλλη προφανή μεταξύ τους σχέση. Ωστόσο, από τον Givens (1952) έχει αποδειχθεί το παρακάτω θεώρημα για το οποίο παρουσιάζουμε μία νέα απλή απόδειξη που βασίζεται στην κανονική μορφή Jordan.

Θεώρημα 6.7.3. Για κάθε πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, ισχύει

$$\bigcap_{\det(S) \neq 0} F(S A S^{-1}) = \text{Co}(\sigma(A)),$$

όπου η τομή περιλαμβάνει όλους τους αντιστρέψιμους πίνακες $S \in \mathbb{C}^{\nu \times \nu}$.

Απόδειξη. Η σχέση

$$\text{Co}(\sigma(A)) \subseteq \bigcap_{\det(S) \neq 0} F(S A S^{-1})$$

είναι προφανής. Έστω τώρα ότι ο A γράφεται στη μορφή Jordan, $A = N J_A N^{-1}$. Παρατηρούμε ότι ο πίνακας Jordan J_A είναι όμοιος με κάθε πίνακα $J_{A,a}$ που προκύπτει από τον J_A με αντικατάσταση των άσων της υπερδιαγωνίου με έναν $a > 0$. Για παράδειγμα, για ένα $k \times k$ Jordan block $J_k(\mu)$, έχουμε

$$\begin{bmatrix} a^{-1} & 0 & \dots & 0 \\ 0 & a^{-2} & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a^{-k} \end{bmatrix} \begin{bmatrix} \mu & 1 & \dots & 0 \\ 0 & \mu & \ddots & \vdots \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & \mu \end{bmatrix} \begin{bmatrix} a & 0 & \dots & 0 \\ 0 & a^2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & 0 \\ 0 & \dots & 0 & a^k \end{bmatrix} = \begin{bmatrix} \mu & a & \dots & 0 \\ 0 & \mu & \ddots & \vdots \\ \vdots & \ddots & \ddots & a \\ 0 & \dots & 0 & \mu \end{bmatrix},$$

όπου το αριθμητικό πεδίο του τελευταίου πίνακα είναι ο κυκλικός δίσκος

$$\Delta\left(\mu, a \cos\left(\frac{\pi}{k+1}\right)\right) \subseteq \Delta(\mu, a).$$

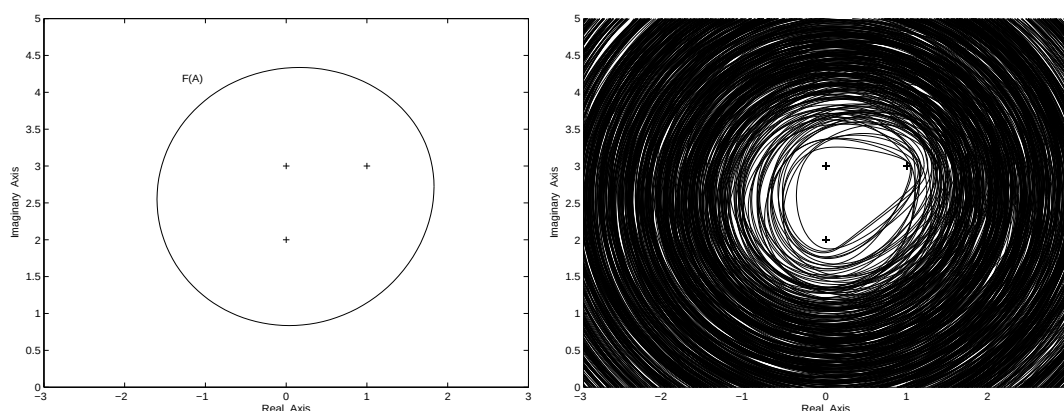
Άρα $J_A = M J_{A,a} M^{-1}$ για κάποιον αντιστρέψιμο $M \in \mathbb{C}^{\nu \times \nu}$ και προφανώς, $A = (NM) J_{A,a} (NM)^{-1}$. Δηλαδή, για κάθε $a > 0$, ο A είναι όμοιος με τον $J_{A,a}$ που έχει αριθμητικό πεδίο

$$F(J_{A,a}) \subseteq \bigcup_{\mu \in \sigma(A)} \Delta(\mu, a).$$

Αφού λοιπόν η τελευταία σχέση ισχύει για οσοδήποτε μικρό $a > 0$ θέλουμε, έπεται ότι για κάθε $z \notin \text{Co}(\sigma(A))$, υπάρχει $a > 0$ με $z \notin F(J_{A,a})$, δηλαδή

$$z \notin \bigcap_{\det(S) \neq 0} F(S A S^{-1}).$$

Άρα ισχύει η ζητούμενη ισότητα. $\square$



Σχήμα 6.5: Το αριθμητικό πεδίο $F(A)$ και τα αριθμητικά πεδία ομοίων του A πινάκων.

Παράδειγμα 6.7.1. Θεωρούμε τον πίνακα

$$A = \begin{bmatrix} 1 + i3 & -1 & 1 - i2 \\ 0 & i2 & 2 - i \\ 0 & 0 & i3 \end{bmatrix}$$

με φάσμα $\sigma(A) = \{1 + i3, i2, i3\}$. Στο αριστερό μέρος του Σχήματος 6.5, δίνεται το $F(A)$, ενώ με “+” σημειώνονται οι ιδιοτιμές του A (που ασφαλώς περιέχονται στο $F(A)$). Στο δεξιό μέρος του σχήματος, έχουμε τα αριθμητικά πεδία 1000 (τυχαίων) όμοιων πινάκων του A . Το Θεώρημα 6.7.3 επαληθεύεται καθώς η τομή των αριθμητικών πεδίων των όμοιων του A πινάκων τείνει να ταυτιστεί με την κυρτή θήκη του φάσματος $\sigma(A) = \{1 + i3, i2, i3\}$.

6.8 Ασκήσεις

1. Να αποδείξετε το Πόρισμα 6.2.5.
2. Να αποδείξετε τα Πορίσματα 6.2.10 και 6.2.12.
3. Να αποδείξετε το Πόρισμα 6.2.13.
4. Να αποδείξετε την Πρόταση 6.2.16.
5. Να αποδείξετε τις Προτάσεις 6.4.2 και 6.4.3.
6. Να αποδείξετε το Πόρισμα 6.6.4.
7. Να αποδείξετε το Πόρισμα 6.6.5.
8. Να αποδείξετε το Πόρισμα 6.6.6.

9. Να αποδείξετε ότι αν το αριθμητικό πεδίο $F(A)$ είναι ένα ευθύγραμμο τμήμα του μιγαδικού επιπέδου, τότε ο πίνακας A είναι κανονικός και συνδέεται άμεσα με κάποιον ερμιτιανό πίνακα.
10. Αν ο πίνακας A είναι αντιστρέψιμος, τότε να αποδείξετε ότι $0 \in F(A)$ αν και μόνο αν $0 \in F(A^{-1})$.
11. Κατασκευάστε έναν 5×5 μη κανονικό πίνακα που έχει αριθμητικό πεδίο ένα τρίγωνο στο μιγαδικό επίπεδο.
12. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ του οποίου το αριθμητικό πεδίο έχει ένα γωνιακό σημείο $a \in \mathbb{C}$. Να αποδείξετε ότι κάθε μοναδιαίο διάνυσμα $y \in \mathbb{C}^{\nu}$ για το οποίο ισχύει $y^* A y = a$ είναι κοινό ιδιοδιάνυσμα των πινάκων A , $H(A)$ και $S(A)$.
13. Βάσει της μεθοδολογίας που περιγράφεται στην Παράγραφο 4.5, σχεδιάστε έναν αλγόριθμο που να προσεγγίζει το αριθμητικό πεδίο ενός τυχαίου πίνακα $A \in \mathbb{C}^{\nu \times \nu}$ μέσω ενός κυρτού πολυγώνου εγγεγραμμένου στο σύνολο του $F(A)$.
14. Βρείτε τη σχέση του πεδίου $F(A)$ με το σύνολο

$$G(A) = \{x^* A x \in \mathbb{C} : x \in \mathbb{C}^{\nu}, x^* x \leq 1\}.$$

Κεφάλαιο 7

Θεωρία Perron-Frobenius

7.1 Εισαγωγικό Παράδειγμα

Ας υποθέσουμε ότι σε μία χώρα υπάρχουν $\nu \geq 2$ πόλεις, $\Pi_1, \Pi_2, \dots, \Pi_\nu$, μεταξύ των οποίων η μεταφορά πληθυσμών γίνεται ως εξής: Καθημερινά, ένα σταθερό ποσοστό a_{ij} του πληθυσμού της πόλης Π_j μετακινείται στην πόλη Π_i , για όλα τα $i \neq j$. Το ποσοστό a_{ii} του πληθυσμού, παραμένει στην πόλη Π_i . Έτσι αν συμβολίσουμε τον πληθυσμό της πόλης Π_i κατά την ημέρα k με $p_i^{(k)}$, τότε έχουμε την ακόλουθη αναδρομική σχέση μεταξύ των κατανομών πληθυσμού κατά τις ημέρες k και $k+1$ ($k = 0, 1, 2, \dots$):

$$p_i^{(k+1)} = a_{i1}p_1^{(k)} + a_{i2}p_2^{(k)} + \dots + a_{i\nu}p_\nu^{(k)}, \quad i = 1, 2, \dots, \nu.$$

Αν θεωρήσουμε τον $\nu \times \nu$ πίνακα συντελεστών μεταφοράς $A = [a_{ij}]$ και το διάνυσμα κατανομής πληθυσμού κατά την ημέρα k , $p^{(k)} = [p_i^{(k)}] \in \mathbb{R}^\nu$, τότε

$$p^{(k+1)} = Ap^{(k)} = A^2p^{(k-1)} = A^{k+1}p^{(0)}, \quad k = 0, 1, 2, \dots,$$

όπου $p^{(0)}$ είναι η αρχική κατανομή πληθυσμού. Αφού οι συντελεστές a_{ij} αναπαριστούν ποσοστά πληθυσμού, ισχύει $0 \leq a_{ij} \leq 1$ ($i, j = 1, 2, \dots, \nu$) και $a_{i1} + a_{i2} + \dots + a_{i\nu} = 1$ για κάθε $i = 1, 2, \dots, \nu$.

Στα πλαίσια του προγραμματισμού για τις επενδύσεις και τις αναβαθμίσεις στις διάφορες πόλεις, η κυβέρνηση της συγκεκριμένης χώρας θέλει να γνωρίζει ποια θα είναι η κατανομή του συνολικού πληθυσμού στο μακρινό μέλλον, $p = \lim p^{(k)}$ (δηλαδή, θέλει να γνωρίζει την ασυμπτωτική συμπεριφορά του $p^{(k)}$). Εφόσον όμως $p^{(k)} = A^k p^{(0)}$, είναι προφανές ότι μπορούμε να περιοριστούμε στην ασυμπτωτική συμπεριφορά του A^k .

Για παράδειγμα, ας θεωρήσουμε την απλή περίπτωση $\nu = 2$, όπου έχουμε $a_{11} + a_{12} = a_{21} + a_{22} = 1$. Αν θέσουμε $a_{12} = \alpha$ και $a_{21} = \beta$, τότε ο πίνακας

συντελεστών μεταφοράς γράφεται

$$A = \begin{bmatrix} 1-\alpha & \beta \\ \alpha & 1-\beta \end{bmatrix}$$

και έχει ιδιοτιμές $\lambda_1 = 1 - \alpha - \beta$ και $\lambda_2 = 1$. Αφού $0 \leq \alpha, \beta \leq 1$, έχουμε ότι $\lambda_2 = 1 \geq |\lambda_1| = |1 - \alpha - \beta|$, άρα η φασματική ακτίνα του A , $\rho(A) = 1 = \lambda_2$, είναι ιδιοτιμή του A . Επίσης, για τις ιδιοτιμές $\lambda_1 = 1 - \alpha - \beta$ και $\lambda_2 = 1$, υπολογίζουμε τα αντίστοιχα ιδιοδιανύσματα $x_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$ και $x_2 = \begin{bmatrix} \beta \\ \alpha \end{bmatrix}$.

Εξετάζουμε τώρα τρεις περιπτώσεις:

1. Αν $\alpha = \beta = 0$, τότε $A = I_2$ και $p^{(k)} = p^{(0)}$ για κάθε $k = 1, 2, \dots$
2. Αν $\alpha = \beta = 1$, τότε $A = A^3 = A^5 = \dots = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ και $A^2 = A^4 = A^6 = \dots = I_2$. Στην περίπτωση αυτή το όριο $\lim p^{(k)} = (\lim A^k) p^{(0)}$ δεν υπάρχει, εκτός αν $p_1^{(0)} = p_2^{(0)}$ (δηλαδή, εκτός αν οι αρχικοί πληθυσμοί είναι ίσοι).
3. Σε κάθε άλλη περίπτωση ο πίνακας συντελεστών μεταφοράς είναι διαγωνοποιήσιμος και

$$A = \begin{bmatrix} \beta & 1 \\ \alpha & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 - \alpha - \beta \end{bmatrix} \left(\frac{1}{\alpha + \beta} \begin{bmatrix} 1 & 1 \\ \alpha & -\beta \end{bmatrix} \right).$$

Επομένως,

$$A^k = \begin{bmatrix} \beta & 1 \\ \alpha & -1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & (1 - \alpha - \beta)^k \end{bmatrix} \left(\frac{1}{\alpha + \beta} \begin{bmatrix} 1 & 1 \\ \alpha & -\beta \end{bmatrix} \right)$$

και το διάνυσμα κατανομής πληθυσμού τείνει στο όριο

$$\lim p^{(k)} = (\lim A^k) p^{(0)} = \frac{1}{\alpha + \beta} \begin{bmatrix} \beta & \beta \\ \alpha & \alpha \end{bmatrix} p^{(0)} = \begin{bmatrix} \frac{\beta}{\alpha + \beta} (p_1^{(0)} + p_2^{(0)}) \\ \frac{\alpha}{\alpha + \beta} (p_1^{(0)} + p_2^{(0)}) \end{bmatrix}.$$

Ειδικότερα, αν θεωρήσουμε ότι τα $p_1^{(0)}$ και $p_2^{(0)}$ εκφράζουν τα ποσοστά του συνολικού πληθυσμού της χώρας που αντιστοιχούν σε κάθε πόλη (κι όχι τους ακριβείς πληθυσμούς των δύο πόλεων), τότε

$$p_1^{(0)} + p_2^{(0)} = 1 \quad \text{και} \quad \lim p^{(k)} = \frac{1}{\alpha + \beta} \begin{bmatrix} \beta \\ \alpha \end{bmatrix}.$$

Παρατηρούμε ότι το διάνυσμα κατανομής πληθυσμού είναι ανεξάρτητο της αρχικής κατανομής πληθυσμού. Το όριο $\lim A^k$ είναι ένας πίνακας, του οποίου οι

στήλες είναι συγγραμμικές με το ιδιοδιάνυσμα $x_2 = \begin{bmatrix} \beta \\ \alpha \end{bmatrix}$ της ιδιοτιμής $\lambda_2 = 1$, δηλαδή της φασματικής ακτίνας του A . Επίσης, συγγραμμικό με το x_2 είναι και το $\lim p^{(k)}$.

Ανακεφαλαιώνοντας, παρατηρούμε τα ακόλουθα:

- Η φασματική ακτίνα $\rho(A)$ είναι απλή ιδιοτιμή του A .
- Τα στοιχεία του ιδιοδιανύσματος $x_2 = \begin{bmatrix} \beta \\ \alpha \end{bmatrix}$ που αντιστοιχεί στην $\rho(A)$ μπορούν να επιλεγούν έτσι ώστε να είναι μη αρνητικά.
- Η φασματική ακτίνα $\rho(A)$ είναι η μοναδική ιδιοτιμή μεγίστου μέτρου.
- Το όριο $\lim (A/\rho(A))^k$ υπάρχει και είναι ένας πίνακας βαθμού 1, του οποίου οι στήλες είναι συγγραμμικά διανύσματα με το ιδιοδιάνυσμα που αντιστοιχεί στην ιδιοτιμή $\rho(A)$.

Τα ενδιαφέροντα αυτά συμπεράσματα ισχύουν για όλους τους τετραγωνικούς πίνακες με θετικά στοιχεία, όπως θα δούμε στη συνέχεια του κεφαλαίου.

7.2 Εισαγωγή στους Μη Αρνητικούς Πίνακες

Έστω $A = [a_{ij}]$, $B = [b_{ij}] \in \mathbb{R}^{\nu \times \nu}$. Γράφουμε:

- $A > 0$ αν $a_{ij} > 0$ για κάθε $i, j = 1, 2, \dots, \nu$,
- $A \geq 0$ αν $a_{ij} \geq 0$ για κάθε $i, j = 1, 2, \dots, \nu$,
- $A \geq B$ αν $A - B \geq 0$,
- $A > B$ αν $A - B > 0$.

Οι αντίστροφες σχέσεις $<$ και $\leq$ ορίζονται ανάλογα. Επίσης, υπενθυμίζουμε το συμβολισμό $|A| = [|a_{ij}|] \in \mathbb{R}^{\nu \times \nu}$. Αν $A \geq 0$, τότε λέμε ότι ο A είναι *μη αρνητικός πίνακας*, και αν $A > 0$, τότε λέμε ότι ο A είναι *θετικός πίνακας*. Οι ιδιότητες που αναφέρονται στο επόμενο πρόταξη προκύπτουν άμεσα από τους ορισμούς. Οι αποδείξεις είναι προφανείς και παραλείπονται.

Πρόταση 7.2.1. Έστω $A, B, C, D \in \mathbb{C}^{\nu \times \nu}$. Τότε ισχύουν τα παρακάτω:

- (i) $|A| \geq 0$, ενώ $|A| = 0$ αν και μόνο αν $A = 0$.
- (ii) $|aA| = |a| |A|$, για κάθε $a \in \mathbb{C}$.
- (iii) $|A + B| \leq |A| + |B|$.
- (iv) Αν $A, B \geq 0$ και $a, b \geq 0$, τότε $aA + bB \geq 0$.
- (v) Αν $A \geq B$ και $C \geq D$, τότε $A + C \geq B + D$.
- (vi) Αν $A \geq B$ και $B \geq C$, τότε $A \geq C$.
- (vii) Αν $x \in \mathbb{C}^n$, τότε $|Ax| \leq |A| |x|$.
- (viii) $|AB| \leq |A| |B|$.
- (ix) $|A^k| \leq |A|^k$, για κάθε $k = 1, 2, \dots$.
- (x) Αν $0 \leq A \leq B$ και $0 \leq C \leq D$, τότε $0 \leq AC \leq BD$.
- (xi) Αν $0 \leq A \leq B$, τότε $0 \leq A^k \leq B^k$ για κάθε $k = 1, 2, \dots$.
- (xii) Αν $A \geq 0$, τότε $A^k \geq 0$, και αν $A > 0$, τότε $A^k > 0$ για κάθε $k = 1, 2, \dots$.
- (xiii) Αν $A > 0$, $x \geq 0$ και $x \neq 0$, τότε $Ax > 0$.
- (xiv) Αν $A \geq 0$, $x > 0$ και $Ax = 0$, τότε $A = 0$.
- (xv) Αν $|A| \geq |B|$, τότε $\|A\|_F \geq \|B\|_F$.
- (xvi) $\|A\|_F = \||A|\|_F$.

Ένα πρώτο μη τετριμμένο αποτέλεσμα είναι η ανισότητα που ακολουθεί και αφορά τη φασματική ακτίνα. Για την απόδειξη της, είναι απαραίτητη η σχέση

$$\lim \|A^k\|_F^{1/k} = \rho(A) \quad (7.2.1)$$

από το Θεώρημα 1.4.5.

Θεώρημα 7.2.2. Έστω πίνακες $A \in \mathbb{C}^{\nu \times \nu}$ και $B \in \mathbb{R}^{\nu \times \nu}$ με $|A| \leq B$. Τότε

$$\rho(A) \leq \rho(|A|) \leq \rho(B).$$

Απόδειξη. Για κάθε $k = 1, 2, \dots$, ισχύει ότι $|A^k| \leq |A|^k \leq B^k$ (από τις ιδιότητες (ix) και (xi) της Πρότασης 7.2.1). Έτσι από τις ιδιότητες (xv) και (xvi) έχουμε ότι $\|A^k\|_F \leq \||A|^k\|_F \leq \|B^k\|_F$ και $\|A^k\|_F^{1/k} \leq \||A|^k\|_F^{1/k} \leq \|B^k\|_F^{1/k}$, για όλα τα $k = 1, 2, \dots$. Αν τώρα θεωρήσουμε ότι $k \rightarrow +\infty$ και εφαρμόσουμε την (7.2.1), συμπεραίνουμε ότι $\rho(A) \leq \rho(|A|) \leq \rho(B)$. $\square$

Πόρισμα 7.2.3. Αν $A, B \in \mathbb{R}^{\nu \times \nu}$ με $0 \leq A \leq B$, τότε $\rho(A) \leq \rho(B)$.

Πόρισμα 7.2.4. Αν $A \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας και $\hat{A}$ ένας οποιοσδήποτε κύριος υποπίνακας του A , τότε $\rho(\hat{A}) \leq \rho(A)$. Ειδικότερα, για τα διαγώνια στοιχεία του A , $a_{11}, a_{22}, \dots, a_{\nu\nu}$, ισχύει $\max_{i=1,2,\dots,\nu} a_{ii} \leq \rho(A)$.

Απόδειξη. Έστω $1 \leq r \leq \nu$ και έστω $\hat{A}$ ένας $r \times r$ κύριος υποπίνακας του A . Συμβολίζουμε με $\tilde{A}$ τον $\nu \times \nu$ πίνακα που σχηματίζεται από τα στοιχεία του $\hat{A}$ (στις ίδιες θέσεις που είχαν στο πίνακα A) και στις υπόλοιπες θέσεις συμπληρώνουμε με μηδενικά. Τότε $\rho(\hat{A}) = \rho(\tilde{A})$ (διότι οι $\hat{A}$ και $\tilde{A}$ έχουν τις ίδιες ακριβώς μη μηδενικές ιδιοτιμές) και $0 \leq \hat{A} \leq A$. Επομένως, $\rho(\hat{A}) = \rho(\tilde{A}) \leq \rho(A)$ από το Πόρισμα 7.2.3. $\square$

Το φράγμα $a_{ii} \leq \rho(A)$ στο προηγούμενο πόρισμα είναι το πρώτο (εύχρηστο) κάτω φράγμα που βρίσκουμε για τη φασματική ακτίνα $\rho(A)$. Για την κατασκευή καλύτερων φραγμάτων, απαραίτητη προϋπόθεση συνεχίζει να είναι η μη αρνητικότητα των στοιχείων του πίνακα A .

Λήμμα 7.2.5. Αν οι γραμμές ενός τετραγωνικού πίνακα $A \geq 0$ έχουν σταθερό άθροισμα ξ , τότε $\rho(A) = \|A\|_{\infty} = \xi$. Αν οι στήλες του πίνακα A έχουν σταθερό άθροισμα ξ , τότε $\rho(A) = \|A\|_1 = \xi$.

Απόδειξη. Από το Θεώρημα 1.4.1, γνωρίζουμε ότι $\rho(A) \leq \|A\|$ για κάθε νόρμα πίνακα $\|\cdot\|$. Αν όμως το άθροισμα των γραμμών είναι σταθερό, τότε αυτό είναι ίσο με $\|A\|_{\infty}$. Έτσι, για το διάνυσμα $x = [1 \ 1 \ \dots \ 1]^T$, ισχύει $Ax = \|A\|_{\infty}x$. Δηλαδή, η νόρμα $\|A\|_{\infty}$ είναι ιδιοτιμή του πίνακα A με αντίστοιχο ιδιοδιάνυσμα το x , και $\rho(A) = \|A\|_{\infty}$. Όμοια εργαζόμαστε και για τις στήλες, εφαρμόζοντας τα παραπάνω για τον πίνακα A^T . $\square$

Θεώρημα 7.2.6. Έστω $A = [a_{ij}] \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας. Τότε

$$\min_{i=1,2,\dots,\nu} \sum_{j=1}^{\nu} a_{ij} \leq \rho(A) \leq \max_{i=1,2,\dots,\nu} \sum_{j=1}^{\nu} a_{ij}$$

και

$$\min_{j=1,2,\dots,\nu} \sum_{i=1}^{\nu} a_{ij} \leq \rho(A) \leq \max_{j=1,2,\dots,\nu} \sum_{i=1}^{\nu} a_{ij}.$$

Απόδειξη. Θέτουμε

$$\alpha = \min_{i=1,2,\dots,\nu} \sum_{j=1}^{\nu} a_{ij}$$

και κατασκευάζουμε ένα νέο πίνακα $B = [b_{ij}]$ με $A \geq B \geq 0$ και $\sum_{j=1}^{\nu} b_{ij} = \alpha$ για όλα τα $i = 1, 2, \dots, \nu$. Για παράδειγμα, αν $\alpha = 0$, τότε θέτουμε $B = 0$, και αν $\alpha > 0$, τότε μπορούμε να θέσουμε $b_{ij} = \alpha a_{ij} \left(\sum_{j=1}^{\nu} a_{ij} \right)^{-1}$. Από το προηγούμενο λήμμα, έχουμε $\rho(B) = \alpha$, και από το Πόρισμα 7.2.3, $\rho(B) \leq \rho(A)$. Το άνω φράγμα εξασφαλίζεται παρόμοια.

Για το δεύτερο μέρος του θεωρήματος, εργαζόμαστε ομοίως επί του αναστρόφου πίνακα A^T . $\square$

Πόρισμα 7.2.7. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας. Αν $\sum_{j=1}^{\nu} a_{ij} > 0$, για κάθε $i = 1, 2, \dots, \nu$, τότε $\rho(A) > 0$.

Γνωρίζουμε ότι οι μετασχηματισμοί ομοιότητας διατηρούν αναλλοίωτο το φάσμα, άρα και τη φασματική ακτίνα. Έτσι, μπορούμε να γενικεύσουμε το Θεώρημα 7.2.6 εισάγοντας κάποιες επιπλέον ελεύθερες παραμέτρους. Συγκεκριμένα, αν $S = \text{diag}\{x_1, x_2, \dots, x_{\nu}\}$, με $x_1, x_2, \dots, x_{\nu} > 0$, τότε $S^{-1}AS \geq 0$ αν $A \geq 0$. Εφαρμόζοντας το Θεώρημα 7.2.6 για τον πίνακα $S^{-1}AS = [a_{ij}x_jx_i^{-1}]$, έχουμε το ακόλουθο αποτέλεσμα.

Θεώρημα 7.2.8. Έστω $A = [a_{ij}] \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας. Τότε για κάθε θετικό διάνυσμα $x = [x_i] \in \mathbb{R}^{\nu}$, έχουμε

$$\min_{i=1,2,\dots,\nu} \left\{ \frac{1}{x_i} \sum_{j=1}^{\nu} a_{ij}x_j \right\} \leq \rho(A) \leq \max_{i=1,2,\dots,\nu} \left\{ \frac{1}{x_i} \sum_{j=1}^{\nu} a_{ij}x_j \right\}$$

και

$$\min_{j=1,2,\dots,\nu} \left\{ x_j \sum_{i=1}^{\nu} a_{ij} \frac{1}{x_i} \right\} \leq \rho(A) \leq \max_{j=1,2,\dots,\nu} \left\{ x_j \sum_{i=1}^{\nu} a_{ij} \frac{1}{x_i} \right\}.$$

Απόδειξη. Όμοια με την απόδειξη του Θεωρήματος 7.2.6. $\square$

Πόρισμα 7.2.9. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας και $x \in \mathbb{R}^{\nu}$ ένα θετικό διάνυσμα. Αν υπάρχουν πραγματικοί αριθμοί $a, b > 0$ τέτοιοι ώστε $ax \leq Ax \leq bx$, τότε $a \leq \rho(A) \leq b$. Αν $ax < Ax$ τότε $a < \rho(A)$, και αν $Ax < bx$ τότε $\rho(A) < b$.

Απόδειξη. Αν $ax \leq Ax$, τότε

$$a \leq \min_{i=1,2,\dots,\nu} \left\{ \frac{1}{x_i} \sum_{j=1}^{\nu} a_{ij}x_j \right\}.$$

Από το Θεώρημα 7.2.8, παρατηρούμε ότι $a \leq \rho(A)$. Αν τώρα $ax < Ax$, τότε υπάρχει ένας $\hat{a} > a$ τέτοιος ώστε $\hat{a}x \leq Ax$. Έτσι $\rho(A) \geq \hat{a} > a$, δηλαδή $\rho(A) > a$. Το άνω φράγμα επαληθεύεται ομοίως. $\square$

Πόρισμα 7.2.10. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας. Αν ο A έχει ένα θετικό ιδιοδιάνυσμα, τότε η αντίστοιχη ιδιοτιμή είναι η $\rho(A)$. Δηλαδή, αν $A \geq 0$, $Ax = \lambda x$ και $x > 0$, τότε $\lambda = \rho(A)$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πόρισμα 7.2.11. Έστω $A = [a_{ij}] \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας. Αν ο A έχει ένα θετικό ιδιοδιάνυσμα $x = [x_i] \in \mathbb{R}^\nu$, τότε

$$\rho(A) = \max_{x>0} \min_{i=1,2,\dots,\nu} \frac{1}{x_i} \sum_{j=1}^{\nu} a_{ij}x_j = \min_{x>0} \max_{i=1,2,\dots,\nu} \frac{1}{x_i} \sum_{j=1}^{\nu} a_{ij}x_j.$$

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πόρισμα 7.2.12. Έστω $A = [a_{ij}] \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας. Αν ο A έχει ένα θετικό ιδιοδιάνυσμα $x = [x_i] \in \mathbb{R}^\nu$, τότε για κάθε $s = 1, 2, \dots$ και $i = 1, 2, \dots, \nu$, ισχύει

$$\frac{\min_{k=1,2,\dots,\nu} x_k}{\max_{k=1,2,\dots,\nu} x_k} \sum_{j=1}^{\nu} a_{ij}^{(s)} \leq \rho(A)^s \leq \frac{\max_{k=1,2,\dots,\nu} x_k}{\min_{k=1,2,\dots,\nu} x_k} \sum_{j=1}^{\nu} a_{ij}^{(s)},$$

όπου $A^s = [a_{ij}^{(s)}]$.

Απόδειξη. Αν $Ax = \rho(A)x$, τότε προφανώς $A^s x = \rho(A)^s x$, για κάθε $s = 1, 2, \dots$. Επιπλέον, αφού $A \geq 0$, προκύπτει ότι $A^s \geq 0$ και

$$\rho(A)^s \max_{k=1,2,\dots,\nu} x_k \geq \rho(A^s) x_i = (A^s x)_i = \sum_{j=1}^{\nu} a_{ij}^{(s)} x_j \geq \left(\min_{k=1,2,\dots,\nu} x_k \right) \sum_{j=1}^{\nu} a_{ij}^{(s)}$$

για κάθε $i = 1, 2, \dots, \nu$. Αφού $x > 0$, διαιρώντας κατάλληλα λαμβάνουμε το κάτω φράγμα. Η απόδειξη του άνω φράγματος είναι όμοια και αφήνεται στον αναγνώστη ως άσκηση. $\square$

7.3 Θετικοί Πίνακες

Η θεωρία των μη αρνητικών πινάκων αποκτά την απλούστερη και κομψότερη μορφή της όταν περιοριστεί στους θετικούς πίνακες. Η συγκεκριμένη θεωρία θεμελιώθηκε και μελετήθηκε για πρώτη φορά από τον Ο. Perron στις αρχές του 20ου αιώνα.

Λήμμα 7.3.1. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας θετικός πίνακας. Αν $Ax = \lambda x$ για κάποιο μη μηδενικό διάνυσμα $x \in \mathbb{C}^\nu$ και $|\lambda| = \rho(A)$, τότε $A|x| = \rho(A)|x|$ και $|x| > 0$.

Απόδειξη. Υπολογίζουμε $\rho(A)|x| = |\lambda||x| = |\lambda x| = |Ax| \leq A|x|$ και ορίζουμε το διάνυσμα $y = A|x| - \rho(A)|x| \geq 0$. Από το Πόρισμα 7.2.7 έχουμε ότι $\rho(A) > 0$. Έτσι, αν $y = 0$, τότε $A|x| = \rho(A)|x|$ και $|x| = \rho(A)^{-1}A|x| > 0$. Αν $y \neq 0$, θέτουμε $z = A|x| > 0$ και εφαρμόζουμε την ιδιότητα (xiii) της Πρότασης 7.2.1. Τότε θα έχουμε $0 < Ay = Az - \rho(A)z$, και συνεπώς, $Az > \rho(A)z$. Όμως από το Πόρισμα 7.2.9, προκύπτει ότι $\rho(A) > \rho(A)$, το οποίο είναι άτοπο. Άρα $y = 0$. $\square$

Χρησιμοποιώντας το προηγούμενο λήμμα, μπορούμε εύκολα να καταλήξουμε σε ένα πρώτο θεμελιώδες συμπέρασμα για τους θετικούς πίνακες.

Θεώρημα 7.3.2. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας θετικός πίνακας. Τότε $\rho(A) > 0$, η φασματική ακτίνα $\rho(A)$ είναι ιδιοτιμή του A και υπάρχει θετικό διάνυσμα $x \in \mathbb{R}^\nu$ τέτοιο ώστε $Ax = \rho(A)x$.

Απόδειξη. Από τον ορισμό της φασματικής ακτίνας, υπάρχει ιδιοτιμή $\lambda \in \sigma(A)$ τέτοια ώστε $|\lambda| = \rho(A)$ και αντίστοιχο (μη μηδενικό) ιδιοδιάνυσμα x . Από το Λήμμα 7.3.1, το ζητούμενο θετικό ιδιοδιάνυσμα είναι το $|x|$. $\square$

Από το Πόρισμα 7.2.11 προκύπτει ότι για κάθε θετικό πίνακα $A \in \mathbb{R}^{\nu \times \nu}$, υπάρχει θετικό ιδιοδιάνυσμα $x = [x_i] \in \mathbb{R}^\nu$ τέτοιο ώστε

$$\rho(A) = \max_{x>0} \min_{i=1,2,\dots,\nu} \frac{1}{x_i} \sum_{j=1}^{\nu} a_{ij}x_j = \min_{x>0} \max_{i=1,2,\dots,\nu} \frac{1}{x_i} \sum_{j=1}^{\nu} a_{ij}x_j.$$

Επεκτείνοντας το Λήμμα 7.3.1, μπορούμε να εξάγουμε περισσότερες πληροφορίες για το που μπορεί να βρίσκονται οι ιδιοτιμές ενός θετικού πίνακα.

Λήμμα 7.3.3. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας θετικός πίνακας, $Ax = \lambda x$, $x \neq 0$ και $|\lambda| = \rho(A)$. Τότε για κάποιο $\theta \in [0, 2\pi]$, ισχύει $e^{-i\theta}x = |x| > 0$.

Απόδειξη. Από την υπόθεση του λήμματος, έχουμε ότι $|Ax| = |\lambda x| = \rho(A)|x|$, ενώ από το Λήμμα 7.3.1, γνωρίζουμε ότι $A|x| = \rho(A)|x|$ και $|x| > 0$. Συνδυάζοντας τις σχέσεις αυτές και την τριγωνική ανισότητα, έχουμε ότι για κάθε $i = 1, 2, \dots, \nu$,

$$\begin{aligned} \rho(A)|x_i| &= |\lambda x_i| = \left| \sum_{j=1}^{\nu} a_{ij}x_j \right| \\ &\leq \sum_{j=1}^{\nu} |a_{ij}| |x_j| = \sum_{j=1}^{\nu} a_{ij}|x_j| \\ &= \rho(A)|x_i|, \end{aligned}$$

άρα θα πρέπει $\left| \sum_{j=1}^{\nu} |a_{ij}| |x_j| \right| = \sum_{j=1}^{\nu} |a_{ij}| |x_j|$. Για να ισχύει όμως κάτι τέτοιο, θα πρέπει οι μη μηδενικοί μιγαδικοί αριθμοί $a_{ij}x_j$ ($j = 1, 2, \dots, \nu$) να βρίσκονται πάνω στην ίδια ημιευθεία του μιγαδικού επιπέδου με αρχή το 0. Αν συμβολίσουμε με θ το κοινό όρισμα των σημείων αυτών, έχουμε ότι $e^{-i\theta} a_{ij}x_j > 0$, για κάθε $j = 1, 2, \dots, \nu$. Αφού όμως $A > 0$, έπεται ότι $e^{-i\theta}x > 0$. $\square$

Θεώρημα 7.3.4. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας θετικός πίνακας. Τότε $|\lambda| < \rho(A)$, για κάθε ιδιοτιμή $\lambda \neq \rho(A)$ του A .

Απόδειξη. Από τον ορισμό της φασματικής ακτίνας, έχουμε ότι $|\lambda| \leq \rho(A)$ για κάθε ιδιοτιμή $\lambda \in \sigma(A)$. Υποθέτουμε ότι $\lambda \in \sigma(A)$ με $|\lambda| = \rho(A)$, $Ax = \lambda x$ και $x \neq 0$. Από το Λήμμα 7.3.3, υπάρχει διάνυσμα $w = e^{-i\theta}x > 0$ για κάποιο $\theta \in [0, 2\pi]$, άρα $Aw = \lambda w$. Αλλά τότε $\lambda = \rho(A)$ από το Πόρισμα 7.2.10. $\square$

Γνωρίζουμε ότι αν $A > 0$, τότε η φασματική ακτίνα $\rho(A)$ ταυτίζεται με τη μέγιστου μέτρου ιδιοτιμή του A , η οποία είναι μοναδική. Το επόμενο αποτέλεσμα μας πληροφορεί ότι η $\rho(A)$ είναι ιδιοτιμή γεωμετρικής πολλαπλότητας 1 (δηλαδή, ο ιδιοχώρος που της αντιστοιχεί έχει διάσταση 1). Επίσης, σύντομα θα δούμε ότι η $\rho(A)$ έχει και αλγεβρική πολλαπλότητα 1.

Θεώρημα 7.3.5. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας θετικός πίνακας και $w, z \in \mathbb{R}^{\nu}$ δύο μη μηδενικά διανύσματα τέτοια ώστε $Aw = \rho(A)w$ και $Az = \rho(A)z$. Τότε τα w, z είναι γραμμικά εξαρτημένα.

Απόδειξη. Από το Λήμμα 7.3.3, υπάρχουν πραγματικοί αριθμοί θ_1 και θ_2 τέτοιοι ώστε τα διανύσματα $p = e^{-i\theta_1}z$ και $q = e^{-i\theta_2}w$ να είναι θετικά. Θέτουμε $\beta = \min_{i=1,2,\dots,\nu} (q_i/p_i)$ και ορίζουμε $r = q - \beta p$. Παρατηρούμε ότι $r \geq 0$ και τουλάχιστον μία συντεταγμένη του r είναι ίση με 0, δηλαδή το διάνυσμα r δεν είναι θετικό. Ταυτόχρονα ισχύει

$$Ar = Aq - \beta Ap = \rho(A)q - \beta \rho(A)p = \rho(A)r.$$

Έτσι, αν $r \neq 0$, τότε από την ιδιότητα (xiii) της Πρότασης 7.2.1 έχουμε ότι $r = \rho(A)^{-1}Ar > 0$. Αυτό όμως δεν ισχύει, άρα $r = 0$. Επομένως, $q = \beta p$ και $w = \beta e^{-i(\theta_1+\theta_2)}z$. $\square$

Πόρισμα 7.3.6. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας θετικός πίνακας. Τότε υπάρχει μοναδικό διάνυσμα $x \in \mathbb{R}^{\nu}$ τέτοιο ώστε $Ax = \rho(A)x$, $x > 0$ και $\|x\|_1 = 1$.

Απόδειξη. Γνωρίζουμε ότι για κάθε θετικό πίνακα A , υπάρχει θετικό ιδιοδιάνυσμα $\hat{x}$ της ιδιοτιμής $\rho(A)$ με $A\hat{x} = \rho(A)\hat{x}$. Για το διάνυσμα $x = \hat{x}/\|\hat{x}\|_1$, έχουμε $Ax = \rho(A)x$ και $\|x\|_1 = 1$. Η μοναδικότητα του x προκύπτει από το Θεώρημα 7.3.5 και το γεγονός ότι $\|x\|_1 = 1$. $\square$

Το μοναδικό κανονικοποιημένο ιδιοδιάνυσμα του προηγούμενου πορίσματος συχνά καλείται *διάνυσμα Perron* του πίνακα A , ενώ η ιδιοτιμή $\rho(A)$ καλείται *ρίζα (ιδιοτιμή) Perron* του A . Αν τώρα έχουμε ότι $A > 0$, είναι προφανές ότι $A^T > 0$ και άρα όλα τα προηγούμενα ισχύουν και για τον ανάστροφο πίνακα A^T . Το διάνυσμα Perron του A^T καλείται *αριστερό διάνυσμα Perron* του A .

Το Θεώρημα 7.3.5 μας εξασφαλίζει ότι ο ιδιοχώρος του πίνακα A που αντιστοιχεί στην ιδιοτιμή $\rho(A)$ είναι μονοδιάστατος, συνεπώς ο χώρος αυτός παράγεται από το διάνυσμα Perron του πίνακα A . Άρα κάθε μη αρνητικό διάνυσμα $0 \neq y \in \mathbb{R}^n$ με $Ay = \rho(A)y$, θα είναι πολλαπλάσιο του διανύσματος Perron του πίνακα A .

Μας ενδιαφέρει η συμπεριφορά των δυνάμεων A^k καθώς $k \rightarrow +\infty$. Το επόμενο λήμμα δίνει τις απαραίτητες προϋποθέσεις για διάφορα θεωρήματα που συμβάλλουν στη μελέτη τέτοιων οριακών καταστάσεων για μη αρνητικούς πίνακες.

Λήμμα 7.3.7. Έστω $A \in \mathbb{C}^{n \times n}$, $\lambda \in \mathbb{C}$ και $x, y \in \mathbb{C}^n$ τέτοια ώστε

$$(1) \quad Ax = \lambda x,$$

$$(2) \quad A^T y = \lambda y,$$

$$(3) \quad x^T y = 1.$$

Ορίζουμε τον πίνακα $L = xy^T$. Τότε ισχύουν τα παρακάτω:

$$(a) \quad Lx = x \text{ και } y^T L = y^T.$$

$$(b) \quad L^k = L, \text{ για κάθε } k = 1, 2, \dots$$

$$(c) \quad A^k L = L A^k = \lambda^k L, \text{ για κάθε } k = 1, 2, \dots$$

$$(d) \quad L(A - \lambda L) = 0.$$

$$(e) \quad (A - \lambda L)^k = A^k - \lambda^k L, \text{ για κάθε } k = 1, 2, \dots$$

$$(f) \quad \text{Κάθε μη μηδενική ιδιοτιμή του } A - \lambda L \text{ είναι επίσης ιδιοτιμή του } A.$$

Αν επιπλέον υποθέσουμε ότι

$$(4) \quad \lambda \neq 0,$$

$$(5) \quad \text{ο } \lambda \text{ είναι ιδιοτιμή του } A \text{ με γεωμετρική πολλαπλότητα } 1,$$

τότε

$$(g) \quad \text{ο } \lambda \text{ δεν είναι ιδιοτιμή του } A - \lambda L.$$

Τέλος, αν επιπλέον υποθέσουμε ότι

$$(6) \quad |\lambda| = \rho(A) > 0,$$

(7) ο λ είναι η μοναδική ιδιοτιμή του A με μέτρο $\rho(A)$ και έχουμε διατάξει τις ιδιοτιμές του A έτσι ώστε $|\lambda_1| \leq |\lambda_2| \leq \dots \leq |\lambda_n| = |\lambda| = \rho(A)$,

τότε

$$(h) \quad \rho(A - \lambda L) \leq |\lambda_{n-1}| < \rho(A),$$

$$(i) \quad (\lambda^{-1}A)^k = L + (\lambda^{-1}A - L)^k \rightarrow L \text{ για } k \rightarrow +\infty.$$

Απόδειξη. (a) Έχουμε

$$Lx = xy^Tx \Leftrightarrow (Lx)^T = x^Tyx^T = x^T \Leftrightarrow Lx = x$$

και

$$y^TL = y^Tx y^T \Leftrightarrow (y^TL)^T = yx^Ty = y^T \Leftrightarrow y^TL = y^T.$$

(b) Ισχύει

$$L^k = L^{k-2}LL = L^{k-2}xy^TL = L^{k-2}xy^T = L^{k-1} = \dots = L.$$

(c) Ισχύει

$$A^kL = (AL)^k = (Axy^T)^k = (\lambda xy^T)^k = (\lambda L)^k = \lambda^k L$$

και

$$\lambda^k L = \lambda^k xy^T = ((\lambda yx^T)^T)^k = ((A^T yx^T)^T)^k = (xy^TA)^k = LA^k.$$

(d) Έχουμε ότι $L(A - \lambda L) = LA - \lambda L^2 = LA - LA = 0$.

(e) Η απόδειξη είναι επαγωγική και αφήνεται στον αναγνώστη, ως άσκηση.

(f) Έστω $\mu \neq 0$ μία ιδιοτιμή του $A - \lambda L$. Τότε υπάρχει διάνυσμα $w \neq 0$ τέτοιο ώστε $(A - \lambda L)w = \mu w$. Από το (d) έχουμε $0 = L(A - \lambda L)w = L\mu w$, και επειδή $\mu \neq 0$, έπεται ότι $Lw = 0$. Συνεπώς,

$$(A - \lambda L)w = Aw - \lambda ALw = Aw = \mu w.$$

(g) Έστω ότι $\lambda \neq 0$ είναι μία ιδιοτιμή του πίνακα A γεωμετρικής πολλαπλότητας 1. Τότε για κάθε διάνυσμα $w \neq 0$ με $Aw = \lambda w$, υπάρχει $0 \neq c \in \mathbb{C}$ τέτοιο ώστε $w = cx$. Αν τώρα υποθέσουμε ότι ο λ είναι ιδιοτιμή του $A - \lambda L$

τότε θα έχουμε $\lambda w = (A - \lambda L)w = (A - \lambda L)cx = c\lambda x - c\lambda x = 0$. Αυτό όμως είναι άτοπο, διότι $\lambda \neq 0$ και $w \neq 0$. Συνεπώς, ο λ δεν είναι ιδιοτιμή του $A - \lambda L$.

(h) Από το (f) έχουμε ότι $\sigma(A - \lambda L) \setminus \{0\} \subseteq \sigma(A)$ και από το (g) γνωρίζουμε ότι $\rho(A - \lambda L) \neq \rho(A)$. Επομένως, $\rho(A - \lambda L) \leq |\lambda_{\nu-1}| < |\lambda| = \rho(A)$.

(i) Συνδυάζοντας τα (h) και (e), έχουμε ότι $(\lambda^{-1}A - L)^k = (\lambda^{-1}A)^k - L$. Επίσης, $\rho(\lambda^{-1}A - L) = \rho(A - \lambda L)/\rho(A) = |\lambda_{\nu-1}|/\rho(A) < 1$. Έτσι συμπεραίνουμε ότι $\lim(\lambda^{-1}A - L)^k = 0$. $\square$

Παρατηρούμε ότι οι υποθέσεις του Λήμματος 7.3.7 ικανοποιούνται για κάθε θετικό πίνακα A όταν $\lambda = \rho(A)$.

Θεώρημα 7.3.8. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας θετικός πίνακας. Τότε

$$\lim \left(\frac{A}{\rho(A)} \right)^k = L,$$

όπου $L = xy^T$, $Ax = \rho(A)x$, $A^T y = \rho(A)y$, $x, y > 0$ και $x^T y = 1$.

Απόδειξη. Παρατηρούμε ότι ικανοποιούνται οι προϋποθέσεις (1)–(7) του προηγούμενου λήμματος, με $\lambda = \rho(A)$ και $y = (x^T z)^{-1}z$, όπου x το δεξιά και z το αριστερό διάνυσμα Perron του πίνακα A . Το ζητούμενο προκύπτει από το (i) του λήμματος. $\square$

Πόρισμα 7.3.9. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας θετικός πίνακας. Τότε το όριο $\lim(A/\rho(A))^k$ είναι ένας θετικός πίνακας βαθμού 1.

Θεώρημα 7.3.10. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας θετικός πίνακας. Τότε η φασματική ακτίνα $\rho(A)$ είναι ιδιοτιμή αλγεβρικής πολλαπλότητας 1 (δηλαδή, είναι απλή ρίζα του χαρακτηριστικού πολυωνύμου του A).

Απόδειξη. Υποθέτουμε ότι η $\rho(A)$ είναι ιδιοτιμή αλγεβρικής πολλαπλότητας κ . Από το Λήμμα 1.4.2 του Schur, έχουμε ότι $A = U T U^*$ για κάποιον ορθομοναδιαίο πίνακα $U \in \mathbb{C}^{\nu \times \nu}$ και κάποιον άνω τριγωνικό πίνακα $T \in \mathbb{C}^{\nu \times \nu}$ με στοιχεία στην κύρια διαγώνιο τα $\rho(A), \dots, \rho(A), \lambda_{\kappa+1}, \dots, \lambda_{\nu}$. Επιπλέον, τα μέτρα των ιδιοτιμών $\lambda_{\kappa+1}, \lambda_{\kappa+2}, \dots, \lambda_{\nu}$ είναι μικρότερα της $\rho(A)$. Έτσι έχουμε

$$L = \lim \left(\frac{A}{\rho(A)} \right)^k = \lim U \begin{bmatrix} 1 & \cdots & * & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & * & \cdots & * \\ 0 & \cdots & 0 & \lambda_{\kappa+1}/\rho(A) & \cdots & * \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & \lambda_{\nu}/\rho(A) \end{bmatrix}^k U^*$$

$$= U \begin{bmatrix} 1 & \cdots & * & * & \cdots & * \\ \vdots & \ddots & \vdots & \vdots & & \vdots \\ 0 & \cdots & 1 & * & \cdots & * \\ 0 & \cdots & 0 & 0 & \cdots & * \\ \vdots & & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} U^*,$$

όπου το στοιχείο 1 εμφανίζεται στην κύρια διαγώνιο κ φορές και το 0 εμφανίζεται $\nu - \kappa$ φορές. Ο άνω τριγωνικός πίνακας στην τελευταία ισότητα έχει βαθμό τουλάχιστον κ . Αφού όμως ο πίνακας L είναι βαθμού 1, συμπεραίνουμε ότι δεν είναι δυνατό να ισχύει $\kappa > 1$. $\square$

Ολοκληρώνοντας την παράγραφο, συνοψίζουμε τα βασικά αποτελέσματα της.

Θεώρημα 7.3.11. (Perron) Έστω ένας πίνακας $A \in \mathbb{R}^{\nu \times \nu}$ με $A > 0$. Τότε ισχύουν τα ακόλουθα:

- (i) $\rho(A) > 0$.
- (ii) Η φασματική ακτίνα $\rho(A)$ είναι ιδιοτιμή του πίνακα A .
- (iii) Υπάρχει θετικό διάνυσμα x με $Ax = \rho(A)x$.
- (iv) Η $\rho(A)$ είναι απλή ιδιοτιμή του πίνακα A .
- (v) Για κάθε ιδιοτιμή $\lambda \neq \rho(A)$ του A , ισχύει $|\lambda| < \rho(A)$, δηλαδή η $\rho(A)$ είναι η μοναδική ιδιοτιμή μέγιστου μέτρου.
- (vi) $\lim (\rho(A)^{-1}A)^k = L$, όπου $L = xy^T$, $Ax = \rho(A)x$, $A^T y = \rho(A)y$, $x, y > 0$ και $x^T y = 1$.

Το Θεώρημα Perron είναι από τα πιο σημαντικά αποτελέσματα στη θεωρία μη αρνητικών πινάκων. Οι εφαρμογές του είναι πολλές και θα αποτελέσει σημείο αναφοράς για πολλά από τα αποτελέσματα που θα ακολουθήσουν.

7.4 Μη Αρνητικοί Πίνακες

Στην πράξη, πιο συχνά έχουμε να κάνουμε με μη αρνητικούς πίνακες παρά με θετικούς. Έτσι παρουσιάζεται η ανάγκη να επεκτείνουμε τη θεωρία που έχουμε αναπτύξει, στη γενικότερη περίπτωση όπου τα στοιχεία ενός μη αρνητικού πίνακα $A \in \mathbb{R}^{\nu \times \nu}$ μπορούν να είναι μηδενικά. Εύστοχα θα μπορούσε κάποιος να υποθέσει ότι η χρήση κατάλληλων ορίων είναι ικανή να μας οδηγήσει στα συμπεράσματα που έχουμε ήδη διατυπώσει για τους θετικούς πίνακες. Δυστυχώς, ορισμένα χαρακτηριστικά (όπως ο βαθμός ενός πίνακα ή η διάσταση ενός ιδιόχωρου) δεν επιτρέπουν

την καθολική εφαρμογή αυτής της μεθοδολογίας. Ωστόσο υπάρχουν αποτελέσματα του Θεωρήματος Perron που μπορούν να γενικευτούν παίρνοντας όρια.

Θεώρημα 7.4.1. Αν $A = [a_{ij}] \in \mathbb{R}^{\nu \times \nu}$ είναι ένας μη αρνητικός πίνακας, τότε η φασματική ακτίνα $\rho(A)$ είναι ιδιοτιμή του A και υπάρχει αντίστοιχο μη αρνητικό ιδιοδιάνυσμα $0 \neq x \in \mathbb{R}^{\nu}$ τέτοιο ώστε $Ax = \rho(A)x$.

Απόδειξη. Για κάθε $\varepsilon > 0$, ορίζουμε τον πίνακα $A(\varepsilon) = [a_{ij} + \varepsilon]$. Συμβολίζουμε με $x(\varepsilon)$ το διάνυσμα Perron του $A(\varepsilon)$, άρα $x(\varepsilon) > 0$ και $\|x(\varepsilon)\|_1 = 1$. Αν θεωρήσουμε το σύνολο $\{x(\varepsilon) : \varepsilon > 0\}$, τότε αυτό περιέχεται στο συμπαγές σύνολο $\{x \in \mathbb{C}^{\nu} : \|x\|_1 \leq 1\}$ και άρα θα υπάρχει μία μονότονα φθίνουσα ακολουθία $\varepsilon_k > 0$ ($k = 1, 2, \dots$) με $\lim \varepsilon_k = 0$ τέτοια ώστε να υπάρχει το όριο $\lim x(\varepsilon_k) = x$. Επίσης, αφού $x(\varepsilon_k) > 0$ ($k = 1, 2, \dots$), έχουμε $x = \lim x(\varepsilon_k) \geq 0$ και $x \neq 0$ επειδή $\|x\|_1 = \lim \|x(\varepsilon_k)\|_1 = 1$.

Από το Θεώρημα 7.2.2, έχουμε ότι $\rho(A(\varepsilon_1)) \geq \rho(A(\varepsilon_2)) \geq \dots \geq \rho(A)$, για κάθε $k = 1, 2, \dots$. Άρα η ακολουθία $\rho(A(\varepsilon_k))$ ($k = 1, 2, \dots$) είναι μονότονα φθίνουσα. Έτσι, το όριο $\rho = \lim \rho(A(\varepsilon_k))$ υπάρχει και είναι μεγαλύτερο ή ίσο της $\rho(A)$. Όμως από το γεγονός ότι

$$\begin{aligned} Ax &= \lim(A(\varepsilon_k)x(\varepsilon_k)) = \lim(\rho(A(\varepsilon_k))x(\varepsilon_k)) \\ &= \lim \rho(A(\varepsilon_k)) \lim x(\varepsilon_k) = \rho x, \end{aligned}$$

και επειδή $x \neq 0$, συμπεραίνουμε ότι ο ρ είναι ιδιοτιμή του πίνακα A . Τότε όμως $\rho \leq \rho(A)$ και άρα $\rho = \rho(A)$. $\square$

Στη συνέχεια, χαρακτηρίζουμε τη φασματική ακτίνα χρησιμοποιώντας μη αρνητικά ιδιοδιανύσματα.

Θεώρημα 7.4.2. Έστω $A = [a_{ij}] \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας και $0 \neq x \in \mathbb{R}^{\nu}$ ένα μη αρνητικό διάνυσμα. Αν $Ax \geq ax$ για κάποιο $a \in \mathbb{R}$, τότε $\rho(A) \geq a$.

Απόδειξη. Για $\varepsilon > 0$, ορίζουμε τον πίνακα $A(\varepsilon) = [a_{ij} + \varepsilon] > 0$, ο οποίος έχει θετικό αριστερό ιδιοδιάνυσμα Perron $y(\varepsilon)$, δηλαδή $y(\varepsilon)^T A(\varepsilon) = \rho(A(\varepsilon)) y(\varepsilon)^T$. Επίσης γνωρίζουμε ότι $Ax - ax \geq 0$, άρα $A(\varepsilon)x - ax > Ax - ax \geq 0$, και έτσι $y(\varepsilon)^T [A(\varepsilon)x - ax] = [\rho(A(\varepsilon)) - a] y(\varepsilon)^T x \geq 0$. Αφού $y(\varepsilon)^T x > 0$, έχουμε ότι για κάθε $\varepsilon > 0$, $\rho(A(\varepsilon)) - a \geq 0$. $\square$

Πόρισμα 7.4.3. Έστω $A = [a_{ij}] \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας. Τότε ισχύει ότι

$$\rho(A) = \max_{\substack{x \geq 0 \\ x \neq 0}} \min_{\substack{i=1,2,\dots,\nu \\ x_i \neq 0}} \frac{1}{x_i} \sum_{j=1}^{\nu} a_{ij} x_j.$$

Απόδειξη. Έστω $A \geq 0$ και $x \geq 0$ με $x \neq 0$. Αν θεωρήσουμε την ποσότητα

$$a = \min_{x_i \neq 0} \sum_{j=1}^{\nu} \frac{a_{ij} x_j}{x_i},$$

τότε $Ax \geq ax$, και από το Θεώρημα 7.4.2, $a \leq \rho(A)$. Αλλά από το Θεώρημα 7.4.1, υπάρχει ιδιοδιάνυσμα $x \geq 0$. Χρησιμοποιώντας το διάνυσμα αυτό εξασφαλίζουμε στο άνω φράγμα, την ισότητα $a = \rho(A)$. $\square$

Αν θεωρήσουμε τον πίνακα $A = \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix}$, ο οποίος έχει ιδιοδιάνυσμα το $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$, τότε παρατηρούμε ότι το άνω φράγμα του Πορίσματος 7.2.9 δεν ισχύει στην περίπτωση που το διάνυσμα x είναι μη αρνητικό αλλά όχι θετικό. Πράγματι, για $\beta = 3/2$, έχουμε $Ax < \beta x$ ενώ $2 = \rho(A) > \beta$.

Προσθέτοντας μία επιπλέον προϋπόθεση, το Θεώρημα 7.4.2 μπορεί να γίνει ισχυρότερο και να μας δώσει πληροφορίες για το διάνυσμα x .

Θεώρημα 7.4.4. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας που έχει ένα θετικό αριστερό ιδιοδιάνυσμα. Αν $0 \neq x \in \mathbb{R}^{\nu}$ είναι ένα μη αρνητικό διάνυσμα με $Ax \geq \rho(A)x$, τότε $Ax = \rho(A)x$.

Απόδειξη. Απόδειξη. Έστω ένα θετικό διάνυσμα y τέτοιο ώστε $A^T y = \rho(A)y$. Τότε

$$y^T(Ax - \rho(A)x) = \rho(A)y^T x - \rho(A)y^T x = 0$$

και επομένως, $Ax = \rho(A)x$. $\square$

Χωρίς επιπλέον προϋποθέσεις, δεν μπορούμε να προχωρήσουμε και να γενικεύσουμε το Θεώρημα Perron για μη αρνητικούς πίνακες. Αν $A \in \mathbb{R}^{\nu \times \nu}$ είναι ένας μη αρνητικός πίνακας, τότε η μη αρνητική ιδιοτιμή $\rho(A)$ καλείται επίσης ρίζα Perron του A . Επειδή το αντίστοιχο ιδιοδιάνυσμα δεν είναι απαραίτητα μοναδικό, είναι δύσκολο να καθορίσουμε το διάνυσμα Perron του πίνακα A . Αν για παράδειγμα, θεωρήσουμε το μη αρνητικό πίνακα $A = I_{\nu}$, τότε η ιδιοτιμή $\rho(A) = 1$ έχει αντίστοιχα ιδιοδιανύσματα όλα τα μη αρνητικά διανύσματα.

7.5 Μη Υποβιβάσιμοι, μη Αρνητικοί Πίνακες

Στην παράγραφο αυτή θα μελετήσουμε μία κατηγορία πινάκων οι οποίοι παρουσιάζονται σε πολλές εφαρμογές. Οι μη υποβιβάσιμοι (irreducible), μη αρνητικοί πίνακες παρουσιάζουν πολλές ιδιαιτερότητες τις οποίες μπορούμε να εκμεταλλευτούμε αν γνωρίζουμε ορισμένες βασικές ιδιότητες τους.

Ένας τετραγωνικός πίνακας λέγεται *πίνακας μετάθεσης* αν έχει ακριβώς ένα στοιχείο σε κάθε στήλη και κάθε γραμμή ίσο 1 και όλα τα υπόλοιπα στοιχεία του μηδενικά. Για παράδειγμα, ο πίνακας $P = \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$ είναι ένας 3×3 πίνακας μετάθεσης. Προφανώς, ένας πίνακας μετάθεσης προκύπτει από το μοναδιαίο πίνακα με μεταθέσεις γραμμών και στηλών.

Ορισμός 7.5.1. Ένας πίνακας $A \in \mathbb{R}^{\nu \times \nu}$ ($\nu \geq 2$) καλείται *υποβιβάσιμος* (ανάγωγος, *reducible*) αν υπάρχει πίνακας μετάθεσης P και φυσικός αριθμός $r \leq \nu$ τέτοιοι έτσι ώστε

$$P^T A P = \begin{bmatrix} B & C \\ 0 & D \end{bmatrix},$$

όπου $B \in \mathbb{R}^{r \times r}$, $C \in \mathbb{R}^{r \times (\nu-r)}$, $D \in \mathbb{R}^{(\nu-r) \times (\nu-r)}$ και $0 \in \mathbb{R}^{(\nu-r) \times r}$. Ένας πίνακας $A \in \mathbb{R}^{\nu \times \nu}$ καλείται *μη υποβιβάσιμος* (μη ανάγωγος, *irreducible*) αν δεν είναι υποβιβάσιμος.

Σημειώνουμε ότι οι υποπίνακες B, C, D δεν είναι απαραίτητο να έχουν όλα τα στοιχεία τους μη μηδενικά. Παρατηρούμε επίσης ότι αν $|A| > 0$ τότε ο A είναι μη υποβιβασμός, ενώ αν ο A είναι υποβιβάσιμος τότε θα πρέπει να έχει τουλάχιστον $\nu - 1$ μηδενικά στοιχεία. Είναι επίσης προφανές ότι ένας πίνακας είναι μη υποβιβάσιμος αν και μόνο αν ο ανάστροφος του είναι μη υποβιβάσιμος.

Θεώρημα 7.5.1. Ένας πίνακας $A \in \mathbb{R}^{\nu \times \nu}$ είναι μη υποβιβασμός αν και μόνο αν ο πίνακας $(I_\nu + |A|)^{\nu-1}$ είναι θετικός.

Απόδειξη. Θα δείξουμε ότι ο πίνακας A είναι υποβιβάσιμος αν και μόνο αν ο $(I_\nu + |A|)^{\nu-1}$ έχει τουλάχιστον ένα μηδενικό στοιχείο. Υποθέτουμε ότι ο A είναι υποβιβάσιμος και άρα για κάποιο πίνακα μετάθεσης $P \in \mathbb{R}^{\nu \times \nu}$ θα έχουμε ότι

$$A = P \begin{bmatrix} B & C \\ 0 & D \end{bmatrix} P^T = P \hat{A} P^T,$$

όπου οι B, C, D και 0 είναι υποπίνακες όπως στον Ορισμό 7.5.1. Παρατηρούμε ότι $|A| = P|\hat{A}|P^T$ αφού ο πίνακας P απλά μεταθέτει τις γραμμές και τις στήλες. Επίσης, παρατηρούμε ότι οι πίνακες $|\hat{A}|^2, |\hat{A}|^3, \dots, |\hat{A}|^{\nu-1}$ έχουν όλοι τον ίδιο υποπίνακα $0 \in \mathbb{R}^{(\nu-r) \times r}$ όπως ο $\hat{A}$. Έτσι

$$\begin{aligned} (I_\nu + |A|)^{\nu-1} &= (I_\nu + P|\hat{A}|P^T)^{\nu-1} = [P(I_\nu + |\hat{A}|)P^T]^{\nu-1} = P(I_\nu + |\hat{A}|)^{\nu-1}P^T \\ &= P \left(I_\nu + \binom{\nu-1}{1} |\hat{A}| + \binom{\nu-1}{2} |\hat{A}|^2 + \dots + \binom{\nu-1}{\nu-1} |\hat{A}|^{\nu-1} \right) P^T, \end{aligned}$$

όπου όλοι οι όροι στη τελευταία παράσταση έχουν το ίδιο υποπίνακα $0 \in \mathbb{R}^{(\nu-r) \times r}$. Συνεπώς, ο πίνακας $(I_\nu + |A|)^{\nu-1}$ είναι υποβιβάσιμος και άρα δεν μπορεί να έχει όλα του τα στοιχεία μη μηδενικά.

Το αντίστροφο αποδεικνύεται με τη βοήθεια της θεωρίας γραφημάτων και για το λόγο αυτό παραλείπεται. $\square$

Λήμμα 7.5.2. Έστω $\lambda_1, \lambda_2, \dots, \lambda_\nu$ οι ιδιοτιμές ενός πίνακα $A \in \mathbb{R}^{\nu \times \nu}$ (όχι κατ' ανάγκη διακεκριμένες). Τότε οι $\lambda_1 + 1, \lambda_2 + 1, \dots, \lambda_\nu + 1$ είναι ιδιοτιμές του πίνακα $I_\nu + A$ και $\rho(I_\nu + A) \leq \rho(A) + 1$. Αν $A \geq 0$, τότε $\rho(I_\nu + A) = \rho(A) + 1$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Λήμμα 7.5.3. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας, με $A^k > 0$ για κάποιο $k \geq 1$. Τότε η φασματική ακτίνα $\rho(A)$ είναι απλή ιδιοτιμή του A .

Απόδειξη. Αν οι $\lambda_1, \lambda_2, \dots, \lambda_\nu$ είναι οι ιδιοτιμές του A , τότε οι $\lambda_1^k, \lambda_2^k, \dots, \lambda_\nu^k$ είναι οι ιδιοτιμές του A^k . Γνωρίζουμε επίσης ότι η φασματική ακτίνα $\rho(A)$ είναι ιδιοτιμή του πίνακα A . Έτσι, αν η $\rho(A)$ είναι πολλαπλή ιδιοτιμή του A , τότε και η $\rho(A^k) = \rho(A)^k$ θα είναι πολλαπλή ιδιοτιμή του A^k . Αυτό όμως είναι άτοπο από το Θεώρημα 7.3.10. $\square$

Θα δούμε στη συνέχεια σε ποιο βαθμό μπορεί να γενικευθεί το Θεώρημα Perron για μη αρνητικούς, μη υποβιβάζσιμους πίνακες. Οι γενικεύσεις των αποτελεσμάτων του Perron για μη αρνητικούς πίνακες είναι άμεσα συνδεδεμένες με το όνομα του F.G. Frobenius. Για το λόγο αυτό άλλωστε η συγκεκριμένη θεωρία είναι γνωστή ως θεωρία Perron-Frobenius.

Θεώρημα 7.5.4. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός, μη υποβιβάζσιμος πίνακας. Τότε ισχύουν τα ακόλουθα:

- (i) $\rho(A) > 0$.
- (ii) Η $\rho(A)$ είναι ιδιοτιμή του A .
- (iii) Υπάρχει θετικό διάνυσμα $x \in \mathbb{R}^\nu$ τέτοιο ώστε $Ax = \rho(A)x$.
- (iv) Η $\rho(A)$ είναι απλή ιδιοτιμή του A .

Απόδειξη. Το Πόρισμα 7.2.7 δίνει άμεσα το (i). Από το Θεώρημα 7.4.1, το (ii) ισχύει για κάθε μη αρνητικό πίνακα $A \in \mathbb{R}^{\nu \times \nu}$. Επίσης, από το Θεώρημα 7.4.1, υπάρχει ένα μη αρνητικό διάνυσμα $x \neq 0$ τέτοιο ώστε $Ax = \rho(A)x$. Αλλά τότε $(I_\nu + A)^{\nu-1}x = (1 + \rho(A))^{\nu-1}x$ και έτσι ο πίνακας $(I_\nu + A)^{\nu-1}$ είναι θετικός σύμφωνα με το Λήμμα 7.5.1, άρα το διάνυσμα $(1 + \rho(A))^{\nu-1}x$ είναι θετικό. Επομένως, το διάνυσμα $x = (1 + \rho(A))^{1-\nu}(I_\nu + A)^{\nu-1}x$ είναι επίσης θετικό.

Για το (iv), εφαρμόζουμε το Λήμμα 7.5.2 και δείχνουμε ότι αν η φασματική ακτίνα $\rho(A)$ είναι πολλαπλή ιδιοτιμή του A , τότε η $1 + \rho(A) = \rho(I_\nu + A)$ είναι πολλαπλή ιδιοτιμή του $I_\nu + A$. Αλλά $I_\nu + A \geq 0$ και $(I_\nu + A)^{\nu-1} > 0$ από το Λήμμα 7.5.1. Συνεπώς, η $1 + \rho(A)$ θα πρέπει, σύμφωνα με το Λήμμα 7.5.2, να είναι απλή ιδιοτιμή του $I_\nu + A$. $\square$

Το προηγούμενο θεώρημα μας εξασφαλίζει ότι ο ιδιοχώρος ενός μη αρνητικού, μη υποβιβάσιμου πίνακα που αντιστοιχεί στη ρίζα Perron, είναι μονοδιάστατος και περιέχει ένα μοναδικό θετικό ιδιοδιάνυσμα που τα στοιχεία του έχουν άθροισμα 1, το οποίο καλείται διάνυσμα Perron του πίνακα. Επιπλέον, γνωρίζουμε ότι ένας πίνακας A είναι μη υποβιβάσιμος αν και μόνο αν ο A^T είναι μη υποβιβαστός. Επομένως, κάθε μη αρνητικός, μη υποβιβάσιμος πίνακας έχει ένα θετικό αριστερό ιδιοδιάνυσμα. Έτσι, το Θεώρημα 7.4.4 ισχύει για κάθε μη αρνητικό, μη υποβιβάσιμο πίνακα. Το γεγονός αυτό παίζει σημαντικό ρόλο στο ακόλουθο αποτέλεσμα (επέκταση του Θεωρήματος 7.2.2).

Θεώρημα 7.5.5. Έστω $A, B \in \mathbb{R}^{\nu \times \nu}$. Αν ο A είναι μη αρνητικός, μη υποβιβάσιμος πίνακας και $A \geq |B|$, τότε $\rho(A) \geq \rho(B)$. Αν επιπλέον $\rho(A) = \rho(B)$ και ο $\lambda = e^{i\phi} \rho(B)$ είναι ιδιοτιμή του πίνακα B , τότε υπάρχουν $\theta_1, \theta_2, \dots, \theta_\nu \in \mathbb{R}$ τέτοια ώστε $B = e^{i\phi} D A D^{-1}$, όπου $D = \text{diag}\{e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_\nu}\}$.

Απόδειξη. Από το Θεώρημα 7.2.2 γνωρίζουμε ότι αν $A \geq |B|$, τότε $\rho(A) \geq \rho(B)$. Αν $\rho(A) = \rho(B)$, τότε υπάρχει διάνυσμα $x \neq 0$ τέτοιο ώστε $Bx = \lambda x$ με $|\lambda| = \rho(B) = \rho(A)$, και άρα $\rho(A)|x| = |\lambda x| = |Bx| \leq |B||x| \leq A|x|$. Αφού ο A είναι μη υποβιβάσιμος, από το Θεώρημα 7.4.4 συμπεραίνουμε ότι $A|x| = \rho(A)|x|$ και έτσι $|Bx| = |B||x| = A|x|$. Επίσης, από το Θεώρημα 7.5.4, έχουμε ότι $|x| > 0$, και αφού $|B| \leq A$ και $|B||x| = A|x|$, συμπεραίνουμε ότι $|B| = A$. Αν τώρα ορίσουμε $\theta_k \in \mathbb{R}$ με $e^{i\theta_k} = x_k/|x_k|$ ($k = 1, 2, \dots, \nu$) και $\lambda = e^{i\phi} \rho(A)$, και αν θέσουμε $D = \text{diag}\{e^{i\theta_1}, e^{i\theta_2}, \dots, e^{i\theta_\nu}\}$, τότε

$$x = D|x| \text{ και } \lambda x = e^{i\phi} \rho(A) D|x| = B D|x| = Bx.$$

Έτσι προκύπτει ότι $e^{-i\phi} D^{-1} B D|x| = \rho(A)|x| = A|x|$. Από την ταυτότητα αυτή και από το γεγονός ότι $|x| > 0$ και $|e^{-i\phi} D^{-1} B D| = A$, συμπεραίνουμε ότι $e^{-i\phi} D^{-1} B D = A$. $\square$

Από το Θεώρημα Perron γνωρίζουμε ότι για ένα θετικό πίνακα $A \in \mathbb{R}^{\nu \times \nu}$, η φασματική ακτίνα $\rho(A)$ είναι η μοναδική ιδιοτιμή μεγίστου μέτρου. Όμως για ένα μη αρνητικό πίνακα A , μπορεί να υπάρχουν περισσότερες από μία ιδιοτιμές μεγίστου μέτρου. Στην περίπτωση αυτή, ο πίνακας A θα πρέπει να έχει ειδική μορφή και οι ιδιοτιμές του θα πρέπει να βρίσκονται στις κορυφές κανονικού πολυγώνου.

Θεώρημα 7.5.6. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός, μη υποβιβάσιμος πίνακας. Υποθέτουμε ότι το σύνολο $S = \{\lambda_\nu, \lambda_{\nu-1}, \dots, \lambda_{\nu-s+1}\}$ των ιδιοτιμών μεγίστου μέτρου $\rho(A)$ έχει ακριβώς s διακεκριμένα στοιχεία. Τότε

$$S = \{e^{i2j\pi/s} \rho(A) : j = 0, 1, \dots, s-1\},$$

δηλαδή αυτές οι ιδιοτιμές μεγίστου μέτρου είναι ακριβώς οι s -οστές ρίζες της μονάδας πολλαπλασιασμένες με τη φασματική ακτίνα $\rho(A)$. Επιπλέον, αν λ είναι μία τυχαία ιδιοτιμή του A , τότε ο αριθμός $e^{i2j\pi/s} \lambda$ είναι επίσης ιδιοτιμή του A για κάθε $j = 0, 1, \dots, s-1$.

Απόδειξη. Η απόδειξη είναι τεχνική και παραλείπεται. $\square$

Αν ένας μη αρνητικός πίνακας A είναι μη υποβιβάσιμος και έχει $s > 1$ ιδιοτιμές μεγίστου μέτρου, τότε κάθε μη μηδενική ιδιοτιμή του ανήκει σε κύκλο κέντρου 0 (στο μιγαδικό επίπεδο) πάνω στον οποίο εμφανίζονται ακριβώς s ιδιοτιμές του A που ορίζουν s ίσα τόξα. Ειδικότερα, ο φυσικός αριθμός s πρέπει να διαιρεί ακριβώς το πλήθος των μη αρνητικών ιδιοτιμών. Έτσι αν ο A είναι $\nu \times \nu$ αντιστρέψιμος, μη αρνητικός και μη υποβιβάσιμος πίνακας και ο ν είναι πρώτος αριθμός, τότε ο A θα έχει είτε μία είτε ν ιδιοτιμές μεγίστου μέτρου.

Πόρισμα 7.5.7. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός, μη υποβιβάσιμος πίνακας. Για κάθε $k = 1, 2, \dots$, συμβολίζουμε $A^k = [a_{ij}^{(k)}]$. Αν υπάρχουν ακριβώς $s > 1$ ιδιοτιμές του A μεγίστου μέτρου, τότε $a_{ii}^k = 0$ για κάθε $i = 1, 2, \dots, \nu$, όταν ο k δεν είναι ακέραιο πολλαπλάσιο του s . Ειδικότερα, για $k = 1$, όλα τα διαγώνια στοιχεία του A είναι μηδενικά.

Απόδειξη. Λόγω του Θεωρήματος 7.5.6, μπορούμε να επιλέξουμε μία ιδιοτιμή $\lambda = e^{i\phi}\rho(A)$ του πίνακα A μεγίστου μέτρου, με όρισμα $\phi = 2\pi/k$. Έτσι, ο $e^{i\phi}\rho(A)$ δεν είναι πραγματικός και θετικός αριθμός, όποτε ο k δεν είναι ακέραιο πολλαπλάσιο του s . Εφαρμόζοντας το Θεώρημα 7.5.5 με $B = A$ και $\lambda = e^{i\phi}\rho(A)$, βρίσκουμε ότι $A = e^{i\phi} D A D^{-1}$ και άρα $A^k = e^{ik\phi} D A^k D^{-1}$. Επομένως, ισχύει ότι $a_{ii}^{(k)} = e^{ik\phi} a_{ii}^{(k)}$ για κάθε $i = 1, 2, \dots, \nu$ και $k = 1, 2, \dots$. Αν ο $e^{ik\phi}$ δεν είναι πραγματικός και θετικός αριθμός, θα πρέπει να έχουμε $a_{11}^{(k)} = a_{22}^{(k)} = \dots = a_{\nu\nu}^{(k)} = 0$, όποτε ο φυσικός αριθμός k δεν είναι ακέραιο πολλαπλάσιο του s . $\square$

Ένας μη αρνητικός πίνακας με όλα τα διαγώνια στοιχεία του μηδενικά μπορεί να έχει μοναδική ιδιοτιμή μεγίστου μέτρου. Για παράδειγμα, ο πίνακας

$$A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$$

έχει μία απλή ιδιοτιμή $\lambda = 2$ και μία διπλή ιδιοτιμή $\lambda = -1$. Δηλαδή, έχει μοναδική ιδιοτιμή μεγίστου μέτρου, η οποία ισούται με τη φασματική ακτίνα $\rho(A)$.

Ισχύει επίσης ότι αν ο $A \geq 0$ είναι μη υποβιβάσιμος και έχει $k > 1$ ιδιοτιμές μεγίστου μέτρου, τότε υπάρχει πίνακας μετάθεσης P τέτοιος ώστε

$$P A P^T = \begin{bmatrix} 0 & A_{1,2} & & 0 \\ \vdots & 0 & \ddots & \\ 0 & & \ddots & A_{k-1,k} \\ A_{k,1} & 0 & \dots & 0 \end{bmatrix},$$

όπου οι k υποπίνακες στην κύρια διαγώνιο είναι τετραγωνικοί και οι A_{ij} δεν είναι απαραίτητα μη μηδενικοί.

Πόρισμα 7.5.8. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας, $\rho(A) > 0$ και λ μία ιδιοτιμή του A τέτοια ώστε $|\lambda| = \rho(A)$. Τότε ο $\lambda/\rho(A) = e^{i\phi}$ είναι s -οστή ρίζα της μονάδας για κάποιο $s \leq \nu$ και

$$\{e^{i\phi}\rho(A), e^{i2\phi}\rho(A), \dots, e^{i(s-1)\phi}\rho(A)\} \subseteq \sigma(A).$$

Απόδειξη. Αν ο πίνακας A είναι μη υποβιβάζσιμος, τότε η απόδειξη έπεται άμεσα από το Θεώρημα 7.5.6. Αν ο A είναι υποβιβάζσιμος, τότε υπάρχει ένας πίνακας μετάθεσης P τέτοιος ώστε

$$P A P^T = \begin{bmatrix} A_1 & * & \cdots & * \\ 0 & A_2 & \cdots & * \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & A_r \end{bmatrix},$$

όπου κάθε τετραγωνικός πίνακας A_j είναι είτε μη υποβιβάζσιμος, είτε μηδενικός. Οι ιδιοτιμές του πίνακα A είναι η ένωση των ιδιοτιμών των πινάκων $A_1, A_2, \dots, A_r$, και η δομή των ιδιοτιμών μεγίστου μέτρου για κάθε A_j δίνεται από το Πόρισμα 7.5.7. $\square$

Παράδειγμα 7.5.1. Έστω οι μη αρνητικοί πίνακες

$$A_1 = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \text{και} \quad A = \begin{bmatrix} A_1 & * \\ 0 & A_2 \end{bmatrix}.$$

Το φάσμα $\sigma(A)$ του 5×5 πίνακα $A \geq 0$ είναι η ένωση $\sigma(A_1) \cup \sigma(A_2)$. Ο πίνακας A_1 έχει δύο απλές ιδιοτιμές μεγίστου μέτρου $\rho(A_1) = 1$ και ο πίνακας A_2 έχει τρεις απλές ιδιοτιμές μεγίστου μέτρου $\rho(A_2) = 1$. Συγκεκριμένα, οι ιδιοτιμές του A_1 είναι οι 1 και -1 , ενώ οι ιδιοτιμές του πίνακα A_2 είναι οι τρεις κυβικές ρίζες της μονάδας, 1, $(-1 + i\sqrt{3})/2$ και $(-1 - i\sqrt{3})/2$. Παρατηρούμε ότι οι ιδιοτιμές του πίνακα A δεν ισαπέχουν στην περιφέρεια του μοναδιαίου κύκλου, σε αντίθεση με τις ιδιοτιμές των πινάκων A_1 και A_2 οι οποίες βρίσκονται στις κορυφές κανονικών πολυγώνων.

7.6 Ένα Γενικό Οριακό Θεώρημα

Το σημαντικότερο ίσως αποτέλεσμα της θεωρίας Perron-Frobenius είναι η ύπαρξη του ορίου $\lim (A/\rho(A))^k$. Είδαμε ότι το όριο αυτό υπάρχει για κάθε θετικό πίνακα. Σε αυτή τη παράγραφο, θα δούμε ότι κάτι ανάλογο ισχύει για μη αρνητικούς, μη υποβιβάζσιμους πίνακες.

Αν $B \in \mathbb{C}^{\nu \times \nu}$ και ο πίνακας $I_\nu - B$ είναι αντιστρέψιμος, τότε μπορεί εύκολα να δει κανείς ότι για κάθε $N = 1, 2, \dots$, ισχύει

$$\sum_{k=1}^N B^k = B(I_\nu - B^N)(I_\nu - B)^{-1}. \quad (7.6.1)$$

Θεώρημα 7.6.1. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός, μη υποβιβάσιμος πίνακας και $x, y \in \mathbb{R}^\nu$ θετικά διανύσματα τέτοια ώστε $Ax = \rho(A)x$, $A^T y = \rho(A)y$, $L = xy^T$ και $x^T y = 1$. Τότε

$$\lim_{N \rightarrow \infty} \frac{1}{N} \sum_{k=1}^N \left(\frac{A}{\rho(A)} \right)^k = L.$$

Απόδειξη. Αν θέσουμε $\lambda = \rho(A)$ και επιλέξουμε για x και y το δεξιό και αριστερό διάνυσμα Perron του πίνακα A , αντίστοιχα, τότε ικανοποιούνται οι υποθέσεις (1)–(7) του Λήμματος 7.3.7 και έτσι ο πίνακας

$$I_\nu - \left(\frac{A}{\rho(A)} - L \right) = \frac{1}{\rho(A)} [\rho(A) I_\nu - (A - \rho(A) L)]$$

είναι αντιστρέψιμος. Από το (ε) του Λήμματος 7.3.7 και την ταυτότητα (7.6.1), έχουμε

$$\begin{aligned} \frac{1}{N} \sum_{k=1}^N \left(\frac{A}{\rho(A)} \right)^k &= \frac{1}{N} \sum_{k=1}^N \left(\frac{A}{\rho(A)} - L + L \right)^k \\ &= \frac{1}{N} \sum_{k=1}^N \left[\left(\frac{A}{\rho(A)} - L \right)^k + L \right] = L + \frac{1}{N} \sum_{k=1}^N \left(\frac{A}{\rho(A)} - L \right)^k \\ &= L + \frac{1}{N} \left(\frac{A}{\rho(A)} - L \right) \left[I_\nu - \left(\frac{A}{\rho(A)} - L \right)^N \right] \left[I_\nu - \left(\frac{A}{\rho(A)} - L \right) \right]^{-1} \\ &= L + \frac{1}{N} \left(\frac{A}{\rho(A)} - L \right) \left[I_\nu - \left(\frac{A}{\rho(A)} \right)^N + L \right] \left[I_\nu - \left(\frac{A}{\rho(A)} - L \right) \right]^{-1}. \end{aligned}$$

Οι μόνοι όροι του τελευταίου μέλους που εξαρτώνται από το N είναι ο $1/N$ και ο $(A/\rho(A))^N$. Αν $x = [x_i]$, τότε από το Πόρισμα 7.2.11, τα στοιχεία του πίνακα $(A/\rho(A))^N$ είναι μικρότερα ή ίσα του πηλίκου $\max_{k=1,2,\dots,\nu} x_k / \min_{k=1,2,\dots,\nu} x_k$. Η απόδειξη ολοκληρώθηκε. $\square$

7.7 Ασκήσεις

1. Να αποδείξετε το Πόρισμα 7.2.10.
2. Να αποδείξετε το Πόρισμα 7.2.11.
3. Να αποδείξετε το άνω φράγμα του Πορίσματος 7.2.12.
4. Να αποδείξετε το (e) του Λήμματος 7.3.7.
5. Να αποδείξετε το Λήμμα 7.5.2.
6. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός πίνακας με ένα θετικό ιδιοδιάνυσμα. Να αποδείξετε ότι ο A είναι όμοιος με ένα μη αρνητικό πίνακα ο οποίος έχει σταθερό άθροισμα γραμμών. Με τι ισούται αυτό το σταθερό άθροισμα;
7. Έστω $A = [a_{ij}] \in \mathbb{R}^{\nu \times \nu}$ ένας θετικός πίνακας με διάνυσμα Perron $x = [x_j]$. Να αποδείξετε ότι $\rho(A) = \sum_{i,j=1}^{\nu} a_{ij}x_j$.
8. Έστω δύο μη αρνητικοί πίνακες $A, B \in \mathbb{R}^{\nu \times \nu}$ τέτοιοι ώστε $AB = BA$. Αν ο A είναι μη υποβιβάζσιμος με διάνυσμα Perron x , να αποδείξετε ότι $Bx = \rho(B)x$.
9. Έστω $A \in \mathbb{R}^{\nu \times \nu}$ ένας μη αρνητικός, μη υποβιβάζσιμος πίνακας. Να αποδείξετε ότι για κάθε μη αρνητικό πίνακα $B \in \mathbb{R}^{\nu \times \nu}$, ισχύει $\rho(A+B) \geq \rho(A)$, ενώ αν επιπλέον $B \neq 0$, τότε $\rho(A+B) > \rho(A)$.

Κεφάλαιο 8

Διαταραχές Ιδιοτιμών και Ψευδοφάσμα Πίνακα

8.1 Διαταραχές Ιδιοτιμών

Το πρόβλημα των διαταραχών των ιδιοτιμών ενός πίνακα μπορεί να διατυπωθεί πολύ απλά: Δεδομένου ενός $n \times n$ πίνακα A και ενός $n \times n$ πίνακα διαταραχής E , ορίζουμε τον $\tilde{A} = A + E$ και μελετάμε τις σχέσεις που συνδέουν τις ιδιοτιμές των πινάκων A και $\tilde{A}$. Το πρώτο αποτέλεσμα που θα αποδείξουμε επιβεβαιώνει ουσιαστικά τη συνέχεια των ιδιοτιμών του πίνακα $\tilde{A} = A + E$ ως προς τα στοιχεία του E . Για την απόδειξη του, απαιτείται ένα κλασικό αποτέλεσμα της μιγαδικής ανάλυσης.

Θεώρημα 8.1.1. (Rouche) Έστω Ω ένα απλά συνεκτικό χωρίο του μιγαδικού επιπέδου. Αν για δύο αναλυτικές συναρτήσεις $\phi, \eta : \Omega \rightarrow \mathbb{C}$, ισχύει

$$|\eta(z)| \leq |\phi(z)|, \quad \forall z \in \partial\Delta,$$

όπου $\partial\Delta$ είναι το σύνορο ενός κλειστού (κυκλικού) δίσκου $\Delta \subset \Omega$, τότε οι συναρτήσεις $\phi(z)$ και $\phi(z) + \eta(z)$ έχουν το ίδιο ακριβώς πλήθος ριζών εντός του Δ , λαμβάνοντας υπόψη και τις πολλαπλότητες.

Θεώρημα 8.1.2. (Ostrowsky-Elsner) Έστω $\lambda \in \mathbb{C}$ μία ιδιοτιμή ενός τυχαίου πίνακα A , αλγεβρικής πολλαπλότητας m . Για κάθε νόρμα πινάκων $\|\cdot\|$ και για κάθε “αρκετά μικρό” $\varepsilon > 0$, υπάρχει ένας $\delta > 0$ τέτοιος ώστε για κάθε πίνακα E με $\|E\| < \delta$, ο δίσκος $\Delta(\lambda, \varepsilon) = \{z \in \mathbb{C} : |z - \lambda| \leq \varepsilon\}$ να περιέχει ακριβώς m ιδιοτιμές του πίνακα $\tilde{A} = A + E$, λαμβάνοντας υπόψη και τις πολλαπλότητες.

Απόδειξη. Έστω ένας αριθμός $\varepsilon > 0$ αρκετά μικρός ώστε η ιδιοτιμή λ να είναι η μοναδική ιδιοτιμή του πίνακα A εντός του δίσκου $\Delta(\lambda, \varepsilon)$. Ενδιαφερόμαστε για

διαταραχές της μορφής $\tilde{A} = A + E$ με $\|E\| \leq \delta$, όπου ο αριθμός $\delta > 0$ είναι “αρκετά μικρός”. Θεωρούμε τα χαρακτηριστικά πολυώνυμα των A και $\tilde{A}$, $\phi_A(z)$ και $\phi_{\tilde{A}}(z)$, καθώς και τη διαφορά τους, $\eta(z) = \phi_{\tilde{A}}(z) - \phi_A(z)$. Από τη συνέχεια των συντελεστών του χαρακτηριστικού πολυώνυμου ως προς τα στοιχεία του πίνακα, προκύπτει ότι για $\tilde{A} \rightarrow A$, η συνάρτηση $\eta(z)$ τείνει στο 0. Αφού λοιπόν το πολυώνυμο $\phi_A(z)$ δεν μηδενίζεται πουθενά στο σύνολο $\partial\Delta(\lambda, \varepsilon)$, υπάρχει $\delta > 0$ τέτοιος ώστε για κάθε E με $\|E\| \leq \delta$, ισχύει ότι $|\eta(z)| \leq |\phi(z)|$ για κάθε $z \in \partial\Delta$. Άρα από το Θεώρημα 8.1.1 (Rouche), τα πολυώνυμα $\phi_A(z)$ και $\phi_{\tilde{A}}(z) = \phi_A(z) + \eta(z)$ έχουν τον ίδιο αριθμό ριζών εντός του Δ . $\square$

Το επόμενο αποτέλεσμα είναι θεμελιώδες στη θεωρία φασματικών διαταραχών πινάκων.

Θεώρημα 8.1.3. (Bauer-Fike) Έστω $A \in \mathbb{C}^{n \times n}$ ένας πίνακας με φάσμα $\sigma(A)$ και $\|\cdot\|$ μία νόρμα πινάκων. Αν Q είναι ένας $n \times n$ αντιστρέψιμος πίνακας, τότε για κάθε ιδιοτιμή του διαταραγμένου πίνακα $\tilde{A} = A + E$, $\tilde{\lambda} \in \sigma(\tilde{A}) \setminus \sigma(A)$, ισχύει

$$\left\| Q^{-1}(A - \tilde{\lambda} I_n)^{-1} Q \right\|^{-1} \leq \|Q^{-1} E Q\|.$$

Απόδειξη. Έχουμε

$$\begin{aligned} Q^{-1}(\tilde{A} - \tilde{\lambda} I_n) Q &= Q^{-1}[(A - \tilde{\lambda} I_n) + E] Q \\ &= Q^{-1}(A - \tilde{\lambda} I_n) Q \left(I_n + Q^{-1}(A - \tilde{\lambda} I_n)^{-1} Q Q^{-1} E Q \right). \end{aligned}$$

Επειδή όμως ο πίνακας $\tilde{A} - \tilde{\lambda} I_n$ είναι μη αντιστρέψιμος, προκύπτει ότι

$$\begin{aligned} 0 &= \det \left(Q^{-1}(\tilde{A} - \tilde{\lambda} I_n) Q \right) \\ &= \det \left(Q^{-1}(A - \tilde{\lambda} I_n) Q \left[I_n + Q^{-1}(A - \tilde{\lambda} I_n)^{-1} Q Q^{-1} E Q \right] \right) \\ &= \det \left(Q^{-1}(A - \tilde{\lambda} I_n) Q \right) \det \left(I_n + Q^{-1}(A - \tilde{\lambda} I_n)^{-1} Q Q^{-1} E Q \right), \end{aligned}$$

όπου η πρώτη ορίζουσα του τελευταίου γινομένου είναι διαφορετική του 0. Επομένως, ο πίνακας $Q^{-1}(A - \tilde{\lambda} I_n)^{-1} Q Q^{-1} E Q$ έχει ιδιοτιμή το -1 . Γνωρίζουμε ότι κάθε ιδιοτιμή ενός τυχαίου πίνακα είναι μικρότερη ή ίση με τη νόρμα του πίνακα. Συνεπώς, από το Θεώρημα 1.4.1 έπεται

$$1 \leq \left\| Q^{-1}(A - \tilde{\lambda} I_n)^{-1} Q Q^{-1} E Q \right\| \leq \left\| Q^{-1}(A - \tilde{\lambda} I_n)^{-1} Q \right\| \|Q^{-1} E Q\|$$

και το ζητούμενο προκύπτει άμεσα. $\square$

Υπενθυμίζουμε ότι από το Λήμμα 1.4.2 του Schur, ένας πίνακας $A \in \mathbb{C}^{n \times n}$ γράφεται στη μορφή

$$A = U T U^* = U(\Lambda_U + R_U)U^*$$

όπου ο πίνακας $U \in \mathbb{C}^{\nu \times \nu}$ είναι ορθομοναδιαίος, ο $T = \Lambda_U + R_U$ είναι άνω τριγωνικός, ο Λ_U είναι διαγώνιος και ο R_U είναι γνήσια άνω τριγωνικός. Προφανώς, ο A είναι κανονικός πίνακας αν και μόνο αν $R_U = 0$. Σύμφωνα λοιπόν με τον Ορισμό 5.2.1 (i), η απομάκρυνση κατά Henrici του A από την κανονικότητα, ως προς μία νόρμα πινάκων $\|\cdot\|$, ορίζεται ως

$$\delta_{\|\cdot\|}(A) = \min \{\|R_U\| : U^* A U = \Lambda_U + R_U, U^* U = U U^* = I_\nu\}.$$

Ειδικά για τη νόρμα Frobenius, από την Πρόταση 5.2.1 γνωρίζουμε ότι για κάθε $A \in \mathbb{C}^{\nu \times \nu}$ με (όχι κατ' ανάγκη διακεκριμένες) ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_\nu$, ισχύει

$$\delta_{\|\cdot\|_F}(A) = \sqrt{\|A\|_F^2 - \sum_{i=1}^{\nu} |\lambda_i|^2}.$$

Η απομάκρυνση $\delta_{\|\cdot\|}(\cdot)$ ορίστηκε και χρησιμοποιήθηκε για πρώτη φορά από τον P. Henrici (1962) και συνδέεται άμεσα με την ευαισθησία των ιδιοτιμών.

Θεώρημα 8.1.4. (Henrici) Έστω $\|\cdot\|$ μία νόρμα πινάκων μεγαλύτερη της $\|\cdot\|_2$, δηλαδή $\|\cdot\| \geq \|\cdot\|_2$. Τότε για κάθε ιδιοτιμή $\tilde{\lambda}$ του διαταραγμένου πίνακα $\tilde{A} = A + E \in \mathbb{C}^{\nu \times \nu}$, υπάρχει μία ιδιοτιμή λ του A τέτοια ώστε

$$\frac{\left(\frac{|\tilde{\lambda} - \lambda|}{\delta_{\|\cdot\|}(A)}\right)^\nu}{1 + \left(\frac{|\tilde{\lambda} - \lambda|}{\delta_{\|\cdot\|}(A)}\right) + \left(\frac{|\tilde{\lambda} - \lambda|}{\delta_{\|\cdot\|}(A)}\right)^2 + \dots + \left(\frac{|\tilde{\lambda} - \lambda|}{\delta_{\|\cdot\|}(A)}\right)^{\nu-1}} \leq \frac{\|E\|_2}{\delta_{\|\cdot\|}(A)}.$$

Απόδειξη. Έστω $\tilde{\lambda}$ μία ιδιοτιμή του πίνακα $\tilde{A}$ και $U^* A U = \Lambda + R$ μία τριγωνοποίηση κατά Schur (Λήμμα 1.4.2) του A . Τότε από το Θεώρημα 8.1.3 (Bauer-Fike) και το γεγονός ότι η 2-νόρμα $\|\cdot\|_2$ είναι ορθομοναδιαία αναλλοίωτη και $\|U\|_2 = 1$, προκύπτει ότι

$$\left\|(\Lambda - \tilde{\lambda} I_\nu + R)^{-1}\right\|_2^{-1} \leq \|E\|_2. \quad (8.1.1)$$

Αφού ο πίνακας R είναι αυστηρά άνω τριγωνικός, ισχύει ότι $R^\nu = 0$. Κατά συνέπεια, μπορεί κανείς να επαληθεύσει ότι

$$\begin{aligned} & (\Lambda - \tilde{\lambda} I_\nu + R)^{-1} \\ &= \left\{ I_\nu - (\Lambda - \tilde{\lambda} I_\nu)^{-1} R + \dots + (-1)^{\nu-1} [(\Lambda - \tilde{\lambda} I_\nu)^{-1} R]^{\nu-1} \right\} (\Lambda - \tilde{\lambda} I_\nu)^{-1}. \end{aligned}$$

Έτσι αν θέσουμε $\delta = \min \left\{ |\tilde{\lambda} - \lambda| : \lambda \in \sigma(A) \right\}$, τότε

$$\left\|(\Lambda - \tilde{\lambda} I_\nu + R)^{-1}\right\|_2 \leq \delta^{-1} \left\{ 1 + \delta^{-1} \delta_{\|\cdot\|}(A) + \dots + [\delta^{-1} \delta_{\|\cdot\|}(A)]^{\nu-1} \right\}.$$

Από την (8.1.1), είναι φανερό ότι

$$\frac{\delta}{1 + \delta^{-1}\delta_{\|\circ\|}(A) + \cdots + [\delta^{-1}\delta_{\|\circ\|}(A)]^{\nu-1}} \leq \left\| (\Lambda - \tilde{\lambda} I_\nu + R)^{-1} \right\|_2^{-1} \leq \|E\|_2.$$

Η απόδειξη ολοκληρώνεται αν διαιρέσουμε την παραπάνω σχέση με $\delta_{\|\circ\|}(A)$ και πολλαπλασιάσουμε τον αριθμητή και τον παρονομαστή του πρώτου κλάσματος με την ποσότητα $(\delta/\delta_{\|\circ\|}(A))^{\nu-1}$. $\square$

Στο παραπάνω θεώρημα, είναι φανερό ότι αν ρ_0 είναι η μεγαλύτερη θετική ρίζα της πολυωνυμικής εξίσωσης

$$z^\nu - \frac{\|E\|_2}{\delta_{\|\circ\|}(A)} (z^{\nu-1} + \cdots + z + 1) = 0,$$

τότε για κάθε ιδιοτιμή $\tilde{\lambda}$ του $\tilde{A}$, υπάρχει ιδιοτιμή λ του A τέτοια ώστε

$$|\lambda - \tilde{\lambda}| \leq \rho_0 \delta_{\|\circ\|}(A).$$

Πρόταση 8.1.5. Έστω $A = Q J_A Q^{-1}$ η κανονική μορφή Jordan ενός πίνακα $A \in \mathbb{C}^{\nu \times \nu}$. Αν n είναι η τάξη του μεγαλύτερου Jordan block του J_A , τότε για κάθε ιδιοτιμή του διαταραγμένου πίνακα $\tilde{\lambda} \in \sigma(\tilde{A})$, υπάρχει ιδιοτιμή $\lambda \in \sigma(A)$ τέτοια ώστε

$$\frac{|\tilde{\lambda} - \lambda|^n}{1 + |\tilde{\lambda} - \lambda| + |\tilde{\lambda} - \lambda|^2 + \cdots + |\tilde{\lambda} - \lambda|^{n-1}} \leq \|Q^{-1} E Q\|_2.$$

Απόδειξη. Η απόδειξη είναι παρόμοια με την απόδειξη του Θεωρήματος 8.1.4. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Το Θεώρημα 8.1.3 (Bauer-Fike) οδηγεί άμεσα και στους γνωστούς δίσκους Gerschgorin.

Θεώρημα 8.1.6. (Gerschgorin) Έστω ένας πίνακας $A = [a_{ij}] \in \mathbb{C}^{\nu \times \nu}$. Για κάθε $i = 1, 2, \dots, \nu$, θεωρούμε το άθροισμα των μέτρων των μη διαγωνίων στοιχείων της i -γραμμής,

$$\alpha_i = \sum_{j \neq i} |a_{ij}|$$

και τον αντίστοιχο δίσκο Gerschgorin

$$\mathcal{G}_i(A) = \Delta(a_{ii}, \alpha_i) = \{z \in \mathbb{C} : |z - a_{ii}| \leq \alpha_i\}.$$

Τότε το φάσμα $\sigma(A)$ ανήκει στην ένωση $\bigcup_{i=1}^{\nu} \mathcal{G}_i(A)$.

Απόδειξη. Έστω $D = \text{diag} \{a_{11}, a_{22}, \dots, a_{nn}\}$. Στο Θεώρημα 8.1.3 (Bauer-Fike), στην τελευταία σχέση της απόδειξης του, θέτουμε:

- τον μοναδιαίο πίνακα I_n στη θέση του Q ,
- τον διαγώνιο πίνακα D στη θέση του A ,
- τον πίνακα A στη θέση του $\tilde{A}$,
- τη νόρμα $\|\cdot\|_\infty$ στη θέση της $\|\cdot\|$.

Τότε, από το τελευταίο μέρος της απόδειξης του Θεωρήματος 8.1.3, έχουμε

$$1 \leq \|(D - \lambda I_n)^{-1}(A - D)\|_\infty,$$

ή ισοδύναμα,

$$1 \leq \max_{1 \leq i \leq n} \left\{ \frac{1}{|a_{ii} - \lambda|} \sum_{j \neq i} |a_{ij}| \right\} = \max_{1 \leq i \leq n} \frac{\alpha_i}{|a_{ii} - \lambda|}.$$

Έτσι είναι εύκολο να δει κανείς ότι κάθε ιδιοτιμή του A βρίσκεται σε ένα δίσκο Gerschgorin. $\square$

Είναι χαρακτηριστικό ότι, καθώς οι απλές ιδιοτιμές ενός πίνακα διαταράσσονται, μπορούμε να έχουμε μία γραφή των διαταραχών τους ανάλογης του αναπτύγματος Taylor. Η απόδειξη του θεωρήματος που ακολουθεί είναι τεχνική και παραλείπεται.

Θεώρημα 8.1.7. Έστω λ μία απλή ιδιοτιμή του πίνακα A , με $x, y \in \mathbb{C}^n$ τα αντίστοιχα δεξιό και αριστερό ιδιοδιάνυσμα. Έστω $\tilde{A} = A + E$ μία “αρκετά μικρή” διαταραχή του A . Τότε υπάρχει μοναδική ιδιοτιμή $\tilde{\lambda}$ του $\tilde{A}$ τέτοια ώστε

$$\tilde{\lambda} = \lambda + \frac{y^* E x}{y^* x} + O(\|E\|_2^2). \quad (8.1.2)$$

Από τη σχέση (8.1.2), μπορεί κανείς να δεχθεί (όταν αυτό απαιτείται για πρακτικούς λόγους) ότι για κάθε απλή ιδιοτιμή του A , ισχύει

$$|\tilde{\lambda} - \lambda| \cong \frac{\|x\|_2 \|y\|_2}{|y^* x|} \|E\|_2,$$

όπου το δεύτερο μέλος είναι πάντα μεγαλύτερο ή ίσο του πρώτου. Ο συντελεστής

$$\kappa_A(\lambda) = \frac{\|x\|_2 \|y\|_2}{|y^* x|}$$

είναι γνωστός ως *βαθμός (δείκτης) κατάστασης* της ιδιοτιμής λ , ως προς τη 2-νόρμα, και προφανώς αποτελεί ένα μέτρο της ευαισθησίας της ιδιοτιμής αυτής.

Στη συνέχεια, στρέφουμε το ενδιαφέρον μας στους διαγωνοποιήσιμους πίνακες, ξεκινώντας από τους κανονικούς.

Πρόταση 8.1.8. Έστω $A \in \mathbb{C}^{\nu \times \nu}$ ένας κανονικός πίνακας με (όχι κατ' ανάγκη διακεκριμένες) ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_\nu$. Αν $\tilde{\lambda}$ είναι μία ιδιοτιμή του $\tilde{A} = A + E$, τότε υπάρχει ιδιοτιμή $\lambda_i \in \sigma(A)$ τέτοια ώστε $|\tilde{\lambda} - \lambda_i| \leq \|E\|_2$.

Απόδειξη. Το αποτέλεσμα προκύπτει άμεσα από το Θεώρημα 8.1.3 (Bauer-Fike), λαμβάνοντας υπόψη το ότι οι κανονικοί πίνακες είναι ορθομοναδιαία όμοιοι με διαγώνιους πίνακες και η 2-νόρμα κάθε ορθομοναδιαίου πίνακα ισούται με 1. $\square$

Το αποτέλεσμα που ακολουθεί, παρουσιάζεται συχνά στην βιβλιογραφία ως το Θεώρημα Bauer-Fike αντί του Θεωρήματος 8.1.3.

Θεώρημα 8.1.9. Έστω ότι ο πίνακας A είναι διαγωνοποιήσιμος με $A = Q \Lambda Q^{-1}$ και $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_\nu\}$. Θεωρούμε επίσης μία νόρμα $\|\circ\|$ που ικανοποιεί τη σχέση $\|\text{diag}\{\mu_1, \mu_2, \dots, \mu_\nu\}\| = \max_{1 \leq i \leq \nu} |\mu_i|$, για κάθε διαγώνιο πίνακα. Τότε για κάθε ιδιοτιμή $\tilde{\lambda}$ του $\tilde{A}$, υπάρχει μία ιδιοτιμή $\lambda_i \in \sigma(A)$ τέτοια ώστε

$$|\tilde{\lambda} - \lambda_i| \leq \|Q\| \|Q^{-1}\| \|E\| = \kappa(Q) \|E\|,$$

όπου $\kappa(Q) = \|Q\| \|Q^{-1}\|$ είναι ο γνωστός βαθμός κατάστασης του πίνακα Q .

Απόδειξη. Από το Θεώρημα 8.1.3 (Bauer-Fike), έχουμε

$$\begin{aligned} 1 &\leq \|(\tilde{\lambda} I_\nu - \Lambda)^{-1} Q E Q^{-1}\| \leq \|Q E Q^{-1}\| \|(\tilde{\lambda} I_\nu - \Lambda)^{-1}\| \\ &= \|Q E Q^{-1}\| \max_{1 \leq i \leq \nu} \frac{1}{|\tilde{\lambda} - \lambda_i|} = \frac{\|Q E Q^{-1}\|}{\min_{1 \leq i \leq \nu} |\tilde{\lambda} - \lambda_i|}. \end{aligned}$$

Εύκολα βλέπει κανείς ότι

$$\min_{1 \leq i \leq \nu} |\tilde{\lambda} - \lambda_i| \leq \|Q E Q^{-1}\| \leq \|Q\| \|Q^{-1}\| \|E\| = \kappa(Q) \|E\|.$$

Η απόδειξη ολοκληρώθηκε. $\square$

Θεώρημα 8.1.10. (Rayleigh-Ritz) Αν A είναι ένας $\nu \times \nu$ ερμιτιανός πίνακας με ιδιοτιμές $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_\nu$, τότε

$$\lambda_\nu x^* x \leq x^* A x \leq \lambda_1 x^* x, \quad \text{για κάθε } x \in \mathbb{C}^\nu,$$

$$\lambda_1 = \max_{x \neq 0} \frac{x^* A x}{x^* x} = \max_{x^* x = 1} x^* A x$$

και

$$\lambda_\nu = \min_{x \neq 0} \frac{x^* A x}{x^* x} = \min_{x^* x = 1} x^* A x.$$

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

8.2 Ψευδοφάσμα Πίνακα

Σε αυτό το εδάφιο, θα ορίσουμε και θα μελετήσουμε το ψευδοφάσμα ενός πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, $\sigma_\varepsilon(A)$, για δεδομένο $\varepsilon > 0$. Θα δούμε πως το σύνολο αυτό “διαστέλεται” καθώς η παράμετρος ε αυξάνει, αλλά και τις ιδιότητες του σε σχέση με τις ιδιότητες του A . Σημειώνουμε ότι, όπου απαιτείται, το ψευδοφάσμα ορίζεται για νόρμες πινάκων που επάγονται από νόρμες διανυσμάτων.

Για ένα δεδομένο $\varepsilon > 0$, το ψευδοφάσμα ενός $\nu \times \nu$ μιγαδικού πίνακα A είναι το σύνολο όλων των ιδιοτιμών όλων των πινάκων που απέχουν από τον A απόσταση (κατά νόρμα) μικρότερη ή ίση του ε .

Ορισμός 8.2.1. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$ και μία νόρμα πινάκων $\|\cdot\|$ η οποία επάγεται από μία νόρμα διανυσμάτων. Τότε οι ακόλουθοι ορισμοί του ψευδοφάσματος του A είναι ισοδύναμοι:

- (1) $\sigma_\varepsilon(A) = \{\lambda \in \mathbb{C} : \|(\lambda I_\nu - A)^{-1}\| \geq \varepsilon^{-1}\}.$
- (2) $\sigma_\varepsilon(A) = \{\lambda \in \mathbb{C} : \lambda \in \sigma(A + E) \text{ για κάποιον } E \in \mathbb{C}^{\nu \times \nu} \text{ με } \|E\| \leq \varepsilon\}.$
- (3) $\sigma_\varepsilon(A) = \{\lambda \in \mathbb{C} : \|(\lambda I_\nu - A)v\| \leq \varepsilon \text{ για κάποιο } v \in \mathbb{C}^\nu \text{ με } \|v\| = 1\}.$
- (4) Όταν αναφερόμαστε στη νόρμα $\|\cdot\|_2$, συμβολίζοντας με $s_{\min}(\cdot)$ την ελάχιστη ιδιάζουσα τιμή ενός πίνακα, ο ακόλουθος ορισμός είναι επίσης ισοδύναμος:

$$\sigma_\varepsilon(A) = \{\lambda \in \mathbb{C} : s_{\min}(\lambda I_\nu - A) \leq \varepsilon\}.$$

Στην περίπτωση που $\lambda \in \sigma(A)$, θεωρούμε ότι $\|(\lambda I_\nu - A)^{-1}\| = \infty$.

Απόδειξη της ισοδυναμίας των ορισμών. Αρχικά θα αποδείξουμε τις συνεπαγωγές $(1) \Rightarrow (3) \Rightarrow (2) \Rightarrow (1)$ και στη συνέχεια, την ισοδυναμία $(1) \Leftrightarrow (4)$.

$[(1) \Rightarrow (3)]$ Υποθέτουμε ότι $\|(\lambda I_\nu - A)^{-1}\| \geq \varepsilon^{-1}$. Τότε υπάρχει μοναδιαίο διάνυσμα $u \in \mathbb{C}^\nu$ τέτοιο ώστε $\|(\lambda I_\nu - A)^{-1}\| = \|(\lambda I_\nu - A)^{-1}u\|$. Αν ορίσουμε $w = (\lambda I_\nu - A)^{-1}u$, τότε

$$\frac{1}{\varepsilon} \leq \|(\lambda I_\nu - A)^{-1}\| = \frac{\|(\lambda I_\nu - A)^{-1}u\|}{\|u\|} = \frac{\|w\|}{\|(\lambda I_\nu - A)w\|}.$$

Συνεπώς, το διάνυσμα $v = w / \|w\|$ είναι μοναδιαίο και ικανοποιεί τη σχέση

$$\|(\lambda I_\nu - A)v\| \leq \varepsilon.$$

$[(3) \Rightarrow (2)]$ Υποθέτουμε ότι υπάρχει ένα μοναδιαίο διάνυσμα $v \in \mathbb{C}^\nu$ τέτοιο ώστε $\|(A - \lambda I_\nu)v\| \leq \varepsilon$. Τότε θα υπάρχει ακόμη ένα μοναδιαίο διάνυσμα $u \in \mathbb{C}^\nu$ που

ικανοποιεί τη σχέση $(A - \lambda I_\nu) v = \hat{\varepsilon} u$, για κάποιον θετικό αριθμό $\hat{\varepsilon} \leq \varepsilon$. Από το Θεώρημα Αναπαράστασης του Riesz και το Θεώρημα Hahn-Banach, υπάρχει ένα διάνυσμα $w \in \mathbb{C}^\nu$ τέτοιο ώστε $w^* v = 1$ και $\|w\| = 1$. Κατά συνέπεια, $\lambda v = A v - \hat{\varepsilon} u w^* v = (A - \hat{\varepsilon} u w^*) v$. Δηλαδή, $\lambda \in \sigma(A + E)$, όπου $E = -\hat{\varepsilon} u w^*$ και $\|E\| \leq \varepsilon$.

[(2) $\Rightarrow$ (1)] Υποθέτουμε ότι $\lambda \in \sigma(A + E)$ για κάποιον πίνακα E με $\|E\| \leq \varepsilon$. Τότε υπάρχει ένα μοναδιαίο διάνυσμα $v \in \mathbb{C}^\nu$ για το οποίο ισχύει $(A + E) v = \lambda v$. Με απλές πράξεις επαληθεύουμε ότι $v = (\lambda I_\nu - A)^{-1} E v$ και

$$1 = \|v\| = \|(\lambda I_\nu - A)^{-1} E v\| \leq \|(\lambda I_\nu - A)^{-1}\| \|E v\| \leq \|(\lambda I_\nu - A)^{-1}\| \varepsilon.$$

Συνεπώς, ισχύει ότι $\|(\lambda I_\nu - A)^{-1}\| \geq \varepsilon^{-1}$.

[(1) $\Leftrightarrow$ (4)] Ας υποθέσουμε τώρα ότι $\|\circ\| = \|\circ\|_2$. Αν

$$s_1(A) (= s_{\max}(A)) \geq s_2(A) \geq s_3(A) \geq \cdots \geq s_{\nu-1}(A) \geq s_\nu(A) (= s_{\min}(A))$$

είναι οι θετικές ιδιάζουσες τιμές ενός $\nu \times \nu$ αντιστρέψιμου πίνακα A , τότε οι ιδιάζουσες τιμές του A^{-1} είναι οι

$$\frac{1}{s_1(A)} \leq \frac{1}{s_2(A)} \leq \frac{1}{s_3(A)} \leq \cdots \leq \frac{1}{s_{\nu-1}(A)} \leq \frac{1}{s_\nu(A)}.$$

Επιπλέον, $\|A\|_2 = s_1(A)$ και $\|A^{-1}\|_2 = 1/s_\nu(A)$. Εύκολα λοιπόν βλέπει κανείς ότι

$$\|A^{-1}\|_2 = s_{\max}(A^{-1}) = \frac{1}{s_{\min}(A)}.$$

Επομένως, έχουμε ότι

$$\|(\lambda I_\nu - A)^{-1}\| \geq \frac{1}{\varepsilon}$$

αν και μόνο αν

$$s_{\max}((\lambda I_\nu - A)^{-1}) = \frac{1}{s_{\min}(\lambda I_\nu - A)} \geq \frac{1}{\varepsilon},$$

ή ισοδύναμα, αν και μόνο αν

$$s_{\min}(\lambda I_\nu - A) \leq \varepsilon.$$

□

Στη συνέχεια, αποδεικνύουμε κάποιες βασικές ιδιότητες του ψευδοφάσματος $\sigma_\varepsilon(A)$. Η πρώτη από αυτές τις ιδιότητες είναι η συμπαγεια του ψευδοφάσματος και προκύπτει εύκολα από τη συνέχεια των ιδιοτιμών.

Πρόταση 8.2.1. Το ψευδοφάσμα $\sigma_\varepsilon(A)$ είναι ένα συμπαγές (δηλαδή, κλειστό και φραγμένο) υποσύνολο του $\mathbb{C}$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πρόταση 8.2.2. (i) Για κάθε $a, b \in \mathbb{C}$, ισχύει $\sigma_{\varepsilon|b|}(a I_\nu + b A) = a + b \sigma_\varepsilon(A)$.
(ii) Αν $A = S B S^{-1}$ και $\kappa(S) = \|S\| \|S^{-1}\|$, τότε $\sigma_\varepsilon(A) \subseteq s_{\varepsilon \kappa(S)}(B)$.

Απόδειξη. (i) Αν $b = 0$ ή $\varepsilon = 0$, τότε το πρόβλημα είναι τετριμμένο. Υποθέτουμε ότι $b \neq 0$ και $\varepsilon > 0$. Από τον ορισμό (1) του ψευδοφάσματος, έχουμε ότι $z \in \sigma_{\varepsilon|b|}(a I_\nu + b A)$ αν και μόνο αν

$$\frac{1}{\varepsilon|b|} \leq \|(z I_\nu - (a I_\nu + b A))^{-1}\| = \frac{1}{|b|} \left\| \left(\frac{z - a}{b} I_\nu - A \right)^{-1} \right\|,$$

ή ισοδύναμα, αν και μόνο αν $(z - a)/b \in \sigma_\varepsilon(A)$.

(ii) Υποθέτουμε ότι $z \in \sigma_\varepsilon(A)$. Από τον ορισμό (1) του ψευδοφάσματος και τις ιδιότητες των νορμών, έχουμε ότι

$$\begin{aligned} \frac{1}{\varepsilon} &\leq \|(z I_\nu - A)^{-1}\| = \|(z S S^{-1} - S B S^{-1})^{-1}\| \\ &\leq \|S\| \|S^{-1}\| \|(z I_\nu - B)^{-1}\|. \end{aligned}$$

Επομένως, αν $1/\varepsilon \leq \|(z I_\nu - A)^{-1}\|$, τότε $1/(\varepsilon \kappa(S)) \leq \|(z I_\nu - B)^{-1}\|$. $\square$

Το ψευδοφάσμα ενός πίνακα A δεν είναι απαραίτητα συνεκτικό. Για το λόγο αυτό, παρουσιάζει ενδιαφέρον η κατανομή των ιδιοτιμών του A στα συνεκτικά τμήματα του $\sigma_\varepsilon(A)$.

Θεώρημα 8.2.3. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Κάθε συνεκτικό τμήμα του ψευδοφάσματος $\sigma_\varepsilon(A)$ περιέχει τουλάχιστον μία ιδιοτιμή του A .

Απόδειξη. Από τον ορισμό (2) του ψευδοφάσματος, για κάθε E με $\|E\| < \varepsilon$, ισχύει $\sigma(A + E) \subset \sigma_\varepsilon(A)$. Αν G είναι ένα συνεκτικό τμήμα του $\sigma_\varepsilon(A)$ και $z \in G$, τότε υπάρχει πίνακας E με $\|E\| \leq \varepsilon$ και $z \in \sigma(A + E)$. Επιπλέον, για το φάσμα κάθε πίνακα της οικογένειας

$$A + t E, \quad t \in [0, 1],$$

ισχύει $\sigma(A + t E) \subset \sigma_\varepsilon(A)$. Από την συνέχεια των ιδιοτιμών (βλέπε το Θεώρημα 8.1.2) του $A + t E$ ως προς τα στοιχεία του $t E$, άρα ως προς t , και καθώς το t αυξάνει από το 0 προς το 1, οι ιδιοτιμές του A και του $A + t E$ συνδέονται με συνεχείς καμπύλες εντός του $\sigma_\varepsilon(A)$. Άρα κάποια ιδιοτιμή $\lambda_0 \in \sigma(A)$ συνδέεται με συνεχή καμπύλη εντός του $\sigma_\varepsilon(A)$ με το z . Επομένως, $\lambda_0 \in G$, και δεν είναι δυνατό ένα συνεκτικό τμήμα του ψευδοφάσματος $\sigma_\varepsilon(A)$ να μην περιέχει τουλάχιστον μία ιδιοτιμή του A . $\square$

Από το θεώρημα αυτό είναι προφανές ότι το ψευδοφάσμα ενός $\nu \times \nu$ πίνακα, αποτελείται από το πολύ ν συνεκτικά τμήματα.

Πόρισμα 8.2.4. Αν το ψευδοφάσμα $\sigma_\varepsilon(A)$ ενός πίνακα $A \in \mathbb{C}^{\nu \times \nu}$ αποτελείται από ν διακεκριμένα συνεκτικά τμήματα, τότε κάθε πίνακας $A + E$ με $\|E\| \leq \varepsilon$ είναι διαγωνοποιήσιμος.

Απόδειξη. Αν το $\sigma_\varepsilon(A)$ αποτελείται από ν διακεκριμένα συνεκτικά τμήματα, τότε από το προηγούμενο θεώρημα, κάθε πίνακας $A + E$ με $\|E\| \leq \varepsilon$, έχει μία ιδιοτιμή σε κάθε συνεκτικό τμήμα. Άρα κάθε $A + E$ με $\|E\| \leq \varepsilon$, έχει ν διακεκριμένες ιδιοτιμές και είναι διαγωνοποιήσιμος. $\square$

Από το σημείο αυτό και στο υπόλοιπο της παραγράφου, θεωρούμε αποκλειστικά ορθομοναδιαία αναλλοίωτες νόρμες πινάκων που ικανοποιούν τη σχέση

$$\|\text{diag}\{\mu_1, \mu_2, \dots, \mu_\nu\}\| = \max_{1 \leq i \leq \nu} |\mu_i|$$

για κάθε διαγώνιο πίνακα.

Πρόταση 8.2.5. Έστω $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ ένας διαγώνιος πίνακας και $\Delta(0, \varepsilon) = \{z \in \mathbb{C} : |z| \leq \varepsilon\}$. Τότε ισχύει

$$\sigma_\varepsilon(\Lambda) = \sigma(\Lambda) + \Delta(0, \varepsilon) = \bigcup_{i=1}^{\nu} \Delta(\lambda_i, \varepsilon).$$

Απόδειξη. Από τον ορισμό (1) του ψευδοφάσματος, ένας $\lambda \in \mathbb{C} \setminus \sigma(\Lambda)$ ανήκει στο ψευδοφάσμα $\sigma_\varepsilon(\Lambda)$ αν και μόνο αν

$$\frac{1}{\|(\Lambda - \lambda I_\nu)^{-1}\|} \leq \varepsilon,$$

ή ισοδύναμα,

$$\frac{1}{\|\text{diag}\{\lambda_1 - \lambda, \lambda_2 - \lambda, \dots, \lambda_\nu - \lambda\}^{-1}\|} \leq \varepsilon,$$

ή ισοδύναμα,

$$\frac{1}{\max_{1 \leq i \leq \nu} |\lambda_i - \lambda|^{-1}} \leq \varepsilon,$$

ή ισοδύναμα,

$$\min_{1 \leq i \leq \nu} |\lambda_i - \lambda| \leq \varepsilon.$$

Η απόδειξη ολοκληρώθηκε. $\square$

Πόρισμα 8.2.6. Έστω $A \in \mathbb{C}^{\nu \times \nu}$ ένας διαγωνοποιήσιμος πίνακας, δηλαδή $A = S \Lambda S^{-1}$ για κάποιον διαγώνιο $\Lambda \in \mathbb{C}^{\nu \times \nu}$ και κάποιον αντιστρέψιμο $S \in \mathbb{C}^{\nu \times \nu}$. Τότε

$$\sigma_\varepsilon(A) \subset \sigma(A) + \Delta(0, \varepsilon \kappa(S)) = \bigcup_{\lambda_i \in \sigma(A)} \Delta(\lambda_i, \varepsilon \kappa(S)).$$

Απόδειξη. Από την Πρόταση 8.2.2 (ii), έχουμε ότι $\sigma_\varepsilon(A) \subset \sigma_{\varepsilon \kappa(S)}(\Lambda)$. Επιπλέον, από την προηγούμενη πρόταση, γνωρίζουμε ότι $s_\delta(\Lambda) = \sigma(\Lambda) + \Delta(0, \delta)$ για κάθε $\delta > 0$. Αφού λοιπόν $\sigma(A) = \sigma(\Lambda)$, έπεται ότι $\sigma_\varepsilon(A) \subset \sigma(A) + \Delta(0, \varepsilon \kappa(S))$. $\square$

Από το πόρισμα αυτό, γίνεται αντιληπτό ότι το ψευδοφάσμα ενός πίνακα A μπορεί να μεταβληθεί όταν εφαρμόσουμε στον A ένα μετασχηματισμό ομοιότητας. Εξάιρεση αποτελούν οι ορθομοναδιαίοι μετασχηματισμοί ομοιότητας.

Πρόταση 8.2.7. Το ψευδοφάσμα $\sigma_\varepsilon(A)$ παραμένει αναλλοίωτο στους ορθομοναδιαίους μετασχηματισμούς ομοιότητας.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πόρισμα 8.2.8. Αν ο πίνακας A είναι κανονικός, τότε $\sigma_\varepsilon(A) = \sigma(A) + \Delta(0, \varepsilon)$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Είναι γνωστό ότι ο πίνακας A είναι μη αντιστρέψιμος αν και μόνο αν $0 \in \sigma(A)$. Κάτι ανάλογο συμβαίνει και με το ψευδοφάσμα.

Πόρισμα 8.2.9. Ισχύει ότι $\|A^{-1}\| \geq 1/\varepsilon$ αν και μόνο αν $0 \in \sigma_\varepsilon(A)$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Τα αποτελέσματα που ακολουθούν αποδεικνύονται εύκολα, και αναφέρονται ουσιαστικά στο μέγεθος που μπορεί να έχει το ψευδοφάσμα ενός πίνακα.

Πρόταση 8.2.10. Για κάθε $A, B \in \mathbb{C}^{\nu \times \nu}$, ισχύει $\sigma_\varepsilon(A + B) \subset \sigma_{\varepsilon + \|B\|}(A)$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πρόταση 8.2.11. Για κάθε $\nu \times \nu$ πίνακα A , ισχύει $\sigma(A) + \Delta(0, \varepsilon) \subseteq \sigma_\varepsilon(A)$. Ισοδύναμα, αν η απόσταση $\text{dist}(z, \sigma(A))$ ενός $z \in \mathbb{C}$ από το φάσμα του A είναι μικρότερη ή ίση του ε , τότε $z \in \sigma_\varepsilon(A)$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Από την παραπάνω πρόταση, παρατηρούμε ότι καθώς αυξάνει η παράμετρος $\varepsilon > 0$, το ψευδοφάσμα αυξάνει τουλάχιστον “γραμμικά” γύρω από τις ιδιοτιμές του πίνακα. Επιπλέον, από το Πρόγραμμα 8.2.8, γίνεται φανερό ότι οι λιγότερο ευαίσθητοι πίνακες είναι οι κανονικοί. Αυτό ήταν κάτι το αναμενόμενο, λόγω της ορθογωνιότητας των ιδιοχώρων των κανονικών πινάκων.

Πρόγραμμα 8.2.12. Έστω ένας πίνακας $A \in \mathbb{C}^{n \times n}$. Αν $z \notin \sigma(A)$, τότε

$$\text{dist}(z, \sigma(A)) \geq \frac{1}{\|(zI_n - A)^{-1}\|}.$$

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πρόταση 8.2.13. Έστω ένας πίνακας $A \in \mathbb{C}^{n \times n}$. Για κάθε $\varepsilon \geq 0$, ισχύει $\sigma_\varepsilon(A) \subseteq \Delta(0, \|A\| + \varepsilon)$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Είναι γνωστό πως όταν ισχύει η συνθήκη $\max_{\lambda \in \sigma(A)} |\lambda| > 1$, τότε

$$\sup_{t \geq 0} \|A^t\| = +\infty.$$

Τώρα θα δούμε κάτι ανάλογο για το ψευδοφάσμα.

Θεώρημα 8.2.14. Έστω ένας πίνακας $A \in \mathbb{C}^{n \times n}$. Αν $\max\{|\lambda_\varepsilon| : \lambda_\varepsilon \in \sigma_\varepsilon(A)\} > 1 + c\varepsilon$, $c > 0$, τότε υπάρχει φυσικός αριθμός k (συμπεριλαμβανομένου και του 0) τέτοιος ώστε $\|A^k\| > c$.

Απόδειξη. Για $c < 1$, το αποτέλεσμα είναι τετριμμένο αφού $\|A^0\| = 1$. Έτσι υποθέτουμε ότι $c \geq 1$.

Αν $\max\{|\lambda| : \lambda \in \sigma(A)\} > 1$, τότε εύκολα μπορεί κανείς να δει ότι το ζητούμενο συμπέρασμα ισχύει. Η απόδειξη αυτής της περίπτωσης αφήνεται στον αναγνώστη, ως άσκηση.

Υποθέτουμε λοιπόν ότι $\max\{|\lambda| : \lambda \in \sigma(A)\} \leq 1$. Στην περίπτωση αυτή, έχουμε την ισότητα

$$(zI_n - A)^{-1} = z^{-1}(I_n + z^{-1}A + z^{-2}A^2 + \dots),$$

για όλα τα $z \in \mathbb{C}$ με $|z| > 1$. Αν δεχθούμε ότι $\|A^k\| \leq c$, για κάθε $k \geq 0$, τότε

$$\frac{1}{\varepsilon} \leq \|(zI_n - A)^{-1}\| \leq \frac{|z|^{-1}c}{1 - |z|^{-1}} = \frac{c}{|z| - 1},$$

συνεπώς,

$$|z| \leq 1 + c\varepsilon.$$

Αυτό όμως σημαίνει ότι το $\sigma_\varepsilon(A)$ περιέχεται σε ένα δίσκο με κέντρο την αρχή των αξόνων και ακτίνα $1 + c\varepsilon$. Δηλαδή, $\max\{|\lambda_\varepsilon| : \lambda_\varepsilon \in \sigma_\varepsilon(A)\} \leq 1 + c\varepsilon$, το οποίο είναι άτοπο. Η απόδειξη ολοκληρώθηκε. $\square$

8.3 Ψευδοφάσμα Κατά τη 2-Νόρμα

Συνεχίζουμε τη μελέτη του ψευδοφάσματος πινάκων, περιορίζοντας το ενδιαφέρον μας αποκλειστικά στη νόρμα $\| \circ \|_2$. Οι ιδιαίτερες ιδιότητες του ψευδοφάσματος κατά τη 2-νόρμα είναι πραγματικά αξιοσημείωτες και έχουν υποχρεώσει πολλούς ερευνητές να επικεντρώσουν την έρευνα τους στην περίπτωση αυτή.

Είναι γνωστό από το Θεώρημα 8.1.6 (Gerschgorin) ότι

$$\sigma(A) \subset \bigcup_{i=1}^{\nu} \mathcal{G}_i(A),$$

όπου

$$\mathcal{G}_i(A) = \Delta(a_{ii}, \alpha_i) = \{z \in \mathbb{C} : |z - a_{ii}| \leq \alpha_i\}.$$

και

$$\alpha_i = \sum_{j \neq i} |a_{ij}|, \quad i = 1, 2, \dots, \nu.$$

Θεώρημα 8.3.1. Για κάθε $A \in \mathbb{C}^{\nu \times \nu}$ και $\varepsilon > 0$, ισχύει

$$\sigma_\varepsilon(A) \subseteq \bigcup_{i=1}^{\nu} \mathcal{G}_i(A) + \Delta(0, \sqrt{\nu} \varepsilon).$$

Απόδειξη. Εφαρμόζουμε το Θεώρημα 8.1.6 (Gerschgorin) στον πίνακα $A + E = [a_{ij} + \varepsilon_{ij}] \in \mathbb{C}^{\nu \times \nu}$, με $\|E\|_2 \leq \varepsilon$. Δηλαδή, θεωρούμε τους δίσκους με κέντρα τα διαγώνια στοιχεία $a_{ii} + \varepsilon_{ii}$ του $A + E$ και αντίστοιχες ακτίνες $r_i = \sum_{j \neq i} |a_{ij} + \varepsilon_{ij}| \leq \alpha_i + \varepsilon_i$, όπου $\varepsilon_i = \sum_{j \neq i} |\varepsilon_{ij}|$. Κάθε τέτοιος δίσκος περιέχεται σε ένα δίσκο με κέντρο a_{ii} και ακτίνα $\rho_i = \alpha_i + \sum_{j=1}^{\nu} |\varepsilon_{ij}| = \alpha_i + \|E_i\|_\infty$, όπου E_i , είναι ο πίνακας που προκύπτει αν μηδενίσουμε τα στοιχεία του E εκτός της i -γραμμής. Τέλος, εύκολα μπορεί κανείς να επαληθεύσει ότι

$$\|E_i\|_\infty \leq \sqrt{\nu} \|E_i\|_2 \leq \sqrt{\nu} \|E\|_2,$$

απ' όπου προκύπτει ο συντελεστής $\sqrt{\nu} \varepsilon$. □

Το αριθμητικό πεδίο

$$F(A) = \{x^* A x \in \mathbb{C} : x \in \mathbb{C}^\nu, x^* x = 1\},$$

που μελετήσαμε στο Κεφάλαιο 6, μας δίνει ένα χωρίο το οποίο υπολογίζεται εύκολα και περιέχει επίσης το κατά τη 2-νόρμα ψευδοφάσμα ενός πίνακα. Επιπλέον, το χωρίο αυτό αποτελεί ένα ικανοποιητικό αρχικό σύνολο για την προσέγγιση του ψευδοφάσματος.

Θεώρημα 8.3.2. Για κάθε πίνακα $A \in \mathbb{C}^{\nu \times \nu}$ και $\varepsilon > 0$, ισχύει

$$\sigma_\varepsilon(A) \subset F(A) + \Delta(0, \varepsilon).$$

Απόδειξη. Έστω $z \in \sigma_\varepsilon(A)$. Από τον ορισμό (2) του ψευδοφάσματος υπάρχει ένας πίνακας E με $\|E\|_2 \leq \varepsilon$ τέτοιος ώστε $z \in \sigma(A + E)$. Επομένως, υπάρχει ένα μη μηδενικό διάνυσμα $v \in \mathbb{C}^\nu$ τέτοιο ώστε $(A + E)v = zv$. Τότε

$$z = \frac{v^*(A + E)v}{v^*v} = \frac{v^*Av}{v^*v} + \frac{v^*Ev}{v^*v},$$

όπου ο πρώτος όρος ανήκει στο $F(A)$ και ο δεύτερος έχει μέτρο το πολύ ε . $\square$

Μια σημαντική ιδιότητα του αριθμητικού πεδίου $F(A)$ ενός πίνακα A είναι η κυρτότητά του. Επομένως, από το παραπάνω θεώρημα, έχουμε ότι

$$\text{Co}(\sigma_\varepsilon(A)) \subset F(A) + \Delta(0, \varepsilon).$$

Όπως μπορεί κανείς να παρατηρήσει στα παραδείγματα του επόμενου εδαφίου, το πλήθος των μη διαφορίσιμων σημείων του συνόρου του ψευδοφάσματος ενός πίνακα είναι πεπερασμένο. Αυτό δικαιολογείται από το γεγονός ότι το σύνολο του ψευδοφάσματος περιέχεται πάντα σε μία αλγεβρική καμπύλη.

Θεώρημα 8.3.3. Έστω ένας πίνακας $A \in \mathbb{C}^{\nu \times \nu}$. Το σύνολο του ψευδοφάσματος $\sigma_\varepsilon(A)$ είναι τμήμα αλγεβρικής καμπύλης (δηλαδή, μιας καμπύλης της οποίας η αναλυτική εξίσωση είναι πολυωνυμική).

Απόδειξη. Αρχικά πρέπει να παρατηρήσουμε ότι λόγω της συνέχειας των ιδιοτιμών ενός πίνακα ως προς τα στοιχεία του, το εσωτερικό του $\sigma_\varepsilon(A)$ περιέχει το σύνολο $\{z \in \mathbb{C} : s_{\min}(zI_\nu - A) < \varepsilon\}$. Η απόδειξη αυτής της παρατήρησης αφήνεται στον αναγνώστη, ως άσκηση.

Γράφοντας τώρα $z = a + ib$ ($a, b \in \mathbb{R}$), έχουμε

$$\begin{aligned} \partial\sigma_\varepsilon(A) &\subseteq \{z \in \mathbb{C} : s_{\min}(zI_\nu - A) = \varepsilon\} \\ &\subseteq \{z \in \mathbb{C} : \text{το } \varepsilon \text{ είναι ιδιάζουσα τιμή του } zI_\nu - A\} \\ &= \left\{ a + ib \in \mathbb{C} : \det \begin{pmatrix} \varepsilon I_\nu & (a + ib)I_\nu - A \\ ((a + ib)I_\nu - A)^* & \varepsilon I_\nu \end{pmatrix} = 0 \right\} \\ &= \left\{ a + ib \in \mathbb{C} : \det \left(\varepsilon I_{2\nu} - \begin{bmatrix} 0 & (a + ib)I_\nu - A \\ ((a + ib)I_\nu - A)^* & 0 \end{bmatrix} \right) = 0 \right\}. \end{aligned}$$

Στο τελευταίο σύνολο, η ορίζουσα παριστάνει τη χαρακτηριστική εξίσωση ενός ερμιτιανού πίνακα. Επομένως, είναι ένα πραγματικό πολυώνυμο των $a, b \in \mathbb{R}$. Δηλαδή, το τελευταίο σύνολο παριστάνει μία αλγεβρική καμπύλη που περιέχει το σύνολο $\partial\sigma_\varepsilon(A)$. $\square$

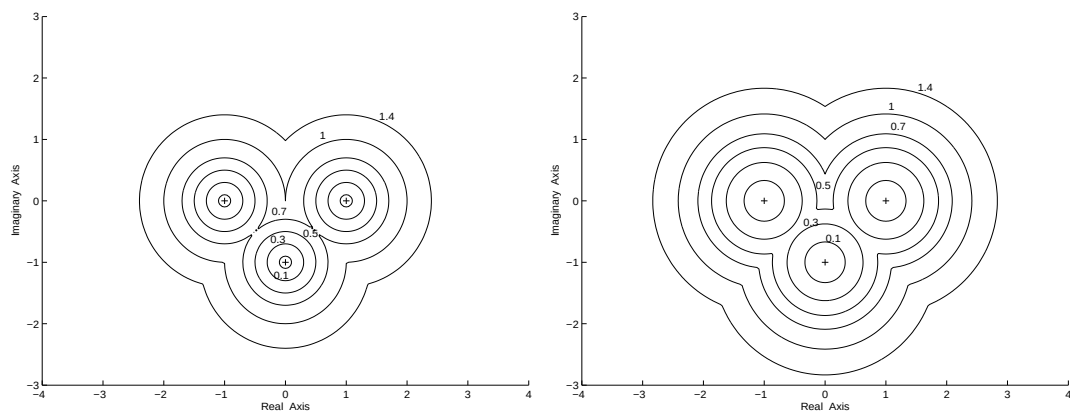
8.4 Παραδείγματα

Ολοκληρώνοντας το κεφάλαιο, παρουσιάζουμε παραδείγματα ψευδοφασμάτων κατά τη 2-νόρμα, με τα οποία επαληθεύουμε τα αποτελέσματα των δύο προηγούμενων παραγράφων. Λόγω της τριγωνποίησης κατά Schur (Λήμμα 1.4.2) και του αναλλοίωτου του ψευδοφάσματος στους ορθομοναδιαίους μετασχηματισμούς ομοιότητας (βλέπε την Πρόταση 8.2.7), χωρίς βλάβη της γενικότητας, επιλέγουμε τους πίνακες των παραδειγμάτων μας να είναι άνω τριγωνικοί. Σε όλα τα σχήματα, οι ιδιοτιμές σημειώνονται με “+”.

Αρχικά θεωρούμε τους πίνακες

$$M = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 0 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix} \quad \text{και} \quad \hat{M} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -i & 1 & 0 & 0 \\ 0 & 0 & 0 & -i & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 \end{bmatrix},$$

οι οποίοι έχουν τις ίδιες ιδιοτιμές, $1, i, -1$, όλες με αλγεβρική πολλαπλότητα 2. Τα ψευδοφάσματα του διαγώνιου πίνακα M , για $\varepsilon = 0.1, 0.3, 0.5, 0.7, 1, 1.4$, εικονίζονται στο αριστερό μέρος του Σχήματος 8.1. Προφανώς, επαληθεύεται ότι το ψευδοφάσμα ενός διαγώνιου πίνακα είναι η ένωση των (κυκλικών) δίσκων με κέντρα τις ιδιοτιμές του πίνακα και ακτίνες ίσες με ε . Στο δεξιό μέρος του ίδιου σχήματος μπορεί κανείς να δει τα ψευδοφάσματα του πίνακα Jordan $\hat{M}$, για τις ίδιες τιμές της παραμέτρου ε . Παρατηρούμε πως κάθε ψευδοφάσμα του $\hat{M}$ περιέχει το αντίστοιχο ψευδοφάσμα του M , επαληθεύοντας την Πρόταση 8.2.11 (δηλαδή, επαληθεύοντας ουσιαστικά ότι οι κανονικοί πίνακες είναι οι λιγότερο ευαίσθητοι πίνακες).



Σχήμα 8.1: Τα ψευδοφάσματα του διαγώνιου πίνακα M και του πίνακα Jordan $\hat{M}$.

Επιπλέον, τα δύο μέρη του Σχήματος 8.1 επιβεβαιώνουν πως το ψευδοφάσμα ενός $\nu \times \nu$ πίνακα είναι συμπαγές υποσύνολο του $\mathbb{C}$ (βλέπε την Πρόταση 8.2.1) και αποτελείται από ν το πολύ συνεκτικά τμήματα με το καθένα τους να περιέχει τουλάχιστον μία ιδιοτιμή του πίνακα (βλέπε το Θεώρημα 8.2.3). Τέλος, για την τιμή $\varepsilon = 1.4$, είναι φανερό ότι (βλέπε την Πρόταση 8.2.13)

$$s_{1.4}(M) \subseteq \Delta(0, \|M\|_2 + 1.4) = \Delta(0, 2.4)$$

και

$$s_{1.4}(\hat{M}) \subseteq \Delta(0, \|\hat{M}\|_2 + 1.4) = \Delta(0, 3.0180).$$

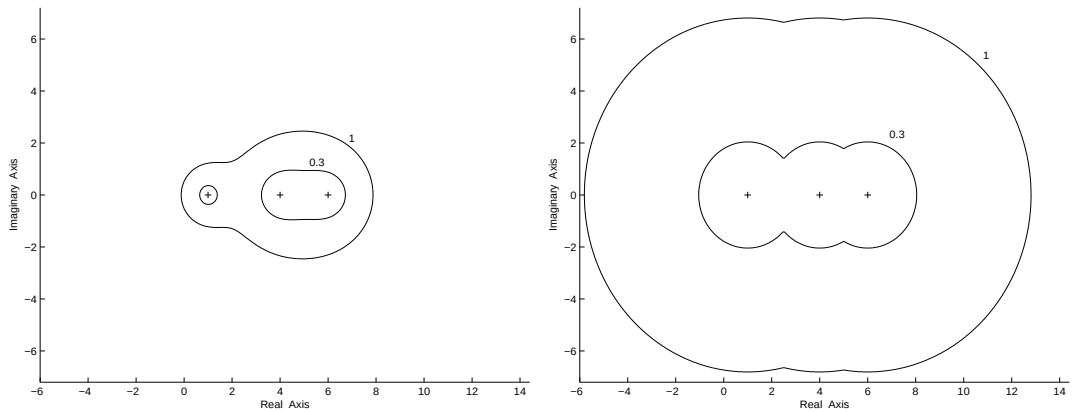
Έστω τώρα ο διαγωνοποιήσιμος πίνακας

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} = S\Lambda S^{-1} = S \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix} S^{-1},$$

όπου ο πίνακας μετασχηματισμού

$$S = \begin{bmatrix} 1 & 0.5547 & 0.5108 \\ 0 & 0.8321 & 0.7982 \\ 0 & 0 & 0.3193 \end{bmatrix}$$

έχει βαθμό κατάστασης $\kappa_2(S) = \|S\|_2 \|S^{-1}\|_2 = 6.8066$. Στο αριστερό μέρος του Σχήματος 8.2 εικονίζονται τα ψευδοφάσματα του A για $\varepsilon = 0.3, 1$, ενώ στο δεξιό μέρος του ίδιου σχήματος εικονίζονται τα ψευδοφάσματα του Λ για $\varepsilon = 0.3, \kappa_2(S) = 2.0420$ και $\varepsilon = \kappa_2(S) = 6.8066$. Συγκρίνοντας τα ψευδοφάσματα, βλέπουμε ότι $s_{0.3}(A) \subset s_{2.0420}(\Lambda)$ και $s_1(A) \subset s_{6.8066}(\Lambda)$, επαληθεύοντας την Πρόταση 8.2.2 (ii) και το Πόρισμα 8.2.6. Σημειώνουμε πως αν και δεν φαίνεται



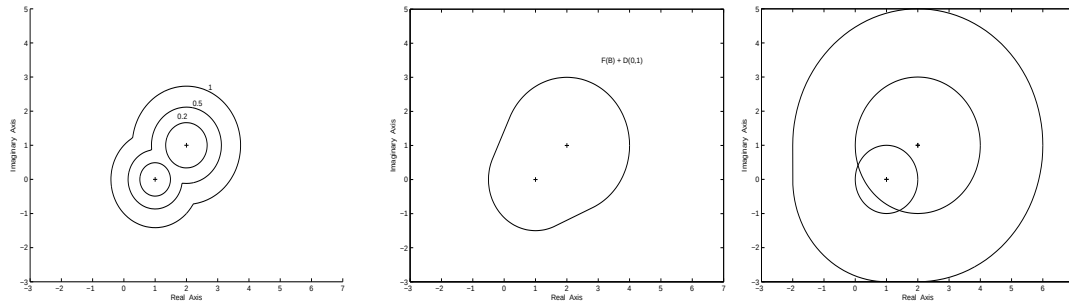
Σχήμα 8.2: Τα ψευδοφάσματα του $A = S\Lambda S^{-1}$ και του Λ .

ξεκάθαρα με την πρώτη ματιά, τα ψευδοφάσματα του διαγώνιου πίνακα Λ είναι η ένωση τριών κυκλικών δίσκων.

Τέλος, για να επαληθεύσουμε τα Θεωρήματα 8.3.1 και 8.3.2, θεωρούμε τον πίνακα

$$B = \begin{bmatrix} 1 & -1 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 2+i & 2 \\ 0 & 0 & 0 & 2+i \end{bmatrix},$$

του οποίου τα ψευδοφάσματα, για $\varepsilon = 0.2, 0.5, 1$, έχουν σχεδιαστεί στο αριστερό μέρος του Σχήματος 8.3. Στο μεσαίο μέρος του σχήματος, έχουμε το σύνολο $F(B) + \Delta(0, 1)$, το οποίο προφανώς περιέχει το ψευδοφάσμα $\sigma_1(B)$ (δηλαδή, για $\varepsilon = 1$). Στο δεξιό μέρος του Σχήματος 8.3, μπορεί κανείς να δει τους δίσκους



Σχήμα 8.3: Τα ψευδοφάσματα του B , το $F(B) + \Delta(0, 1)$, και οι δίσκοι Gersgorin.

Gersgorin $\Delta(1, 1)$ και $\Delta(2 + i, 2)$, καθώς και το σύνολο του Θεωρήματος 8.3.1, $(\Delta(1, 1) \cup \Delta(2 + i, 2)) + \Delta(0, 2)$, το οποίο επίσης περιέχει το $\sigma_1(B)$.

8.5 Ασκήσεις

1. Να αποδείξετε την Πρόταση 8.1.5.
2. Να αποδείξετε το Θεώρημα 8.1.10.
3. Έστω $\lambda \in \mathbb{C}$ μία ιδιοτιμή ενός $\nu \times \nu$ πίνακα A , με $u \in \mathbb{C}^\nu$ ένα αριστερό ιδιοδιάνυσμα της και $v \in \mathbb{C}^\nu$ ένα δεξιό ιδιοδιάνυσμα της. Να αποδειχθεί ότι αν $u^*v = 0$, τότε η λ είναι πολλαπλή ιδιοτιμή του A .
4. Έστω ένας $\nu \times \nu$ κανονικός πίνακας A με διακεκριμένες ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_\nu$ τέτοιες ώστε $|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_\nu|$. Κατασκευάστε ένα άνω φράγμα για την απόσταση

$$d_k(A) = \min \{ \|E\|_2 : \text{o } A + E \text{ έχει ιδιοτιμή πολλαπλότητας } k \text{ το } 0 \}.$$

και υπολογίστε την απόσταση

$$\Delta_2(A) = \min \{ \|E\|_2 : \text{o } A + E \text{ έχει μία διπλή ιδιοτιμή} \}.$$

5. Να αποδείξετε την Πρόταση 8.2.1.
6. Να αποδείξετε την Πρόταση 8.2.7.
7. Να αποδείξετε την Πρόταση 8.2.8.
8. Να αποδείξετε την Πρόταση 8.2.10.
9. Να αποδείξετε την Πρόταση 8.2.11.
10. Να αποδείξετε το Πόρισμα 8.2.12.
11. Να αποδείξετε την Πρόταση 8.2.13.
12. Να αποδείξετε ότι το εσωτερικό του ψευδοφάσματος $\sigma_\varepsilon(A)$, κατά τη 2-νόρμα, περιέχει το σύνολο $\{z \in \mathbb{C} : s_{\min}(z I_\nu - A) < \varepsilon\}$.
13. Να αποδείξετε ότι το ψευδοφάσμα ενός πραγματικού ή ερμιτιανού πίνακα είναι πάντα συμμετρικό ως προς τον άξονα των πραγματικών αριθμών.
14. Κατασκευάστε έναν αλγόριθμο και το αντίστοιχο πρόγραμμα Matlab για το σχεδιασμό του ψευδοφάσματος, κατά τη 2-νόρμα, ενός πίνακα.

Κεφάλαιο 9

Πολυωνυμικοί Πίνακες

9.1 Ορισμός Πολυωνυμικού Πίνακα

Πολυωνυμικό πίνακα ονομάζουμε κάθε $\nu \times \mu$ πίνακα

$$P(\lambda) = \begin{bmatrix} p_{11}(\lambda) & p_{12}(\lambda) & \cdots & p_{1\mu}(\lambda) \\ p_{21}(\lambda) & p_{22}(\lambda) & \cdots & p_{2\mu}(\lambda) \\ \vdots & \vdots & & \vdots \\ p_{\nu 1}(\lambda) & p_{\nu 2}(\lambda) & \cdots & p_{\nu\mu}(\lambda) \end{bmatrix},$$

του οποίου τα στοιχεία $p_{ij}(\lambda)$ είναι (μιγαδικά) πολυώνυμα της μεταβλητής λ . Ένα απλό παράδειγμα πολυωνυμικού πίνακα αποτελεί η γνωστή παράσταση¹ $I_\nu \lambda - A$ που εμφανίζεται στην κατασκευή του χαρακτηριστικού πολυωνύμου ενός πίνακα $A \in \mathbb{C}^{\nu \times \nu}$.

Ένας τετραγωνικός πολυωνυμικός πίνακας $P(\lambda)$ λέγεται *ομαλός* (*regular*) αν η ορίζουσα του, $\det(P(\lambda))$, δεν είναι ταυτοτικά μηδενική. Ο *βαθμός* ενός πολυωνυμικού πίνακα $P(\lambda)$ είναι ο μεγαλύτερος βαθμός των πολυωνύμων που εμφανίζονται ως στοιχεία του $P(\lambda)$ και συμβολίζεται με $\deg P(\lambda)$, δηλαδή $\deg(I_\nu \lambda - A) = 1$ και $\deg A = 0$.

Εύκολα γίνεται αντιληπτό ότι κάθε πολυωνυμικός πίνακας $P(\lambda)$ βαθμού m γράφεται και στη μορφή

$$P(\lambda) = A_m \lambda^m + A_{m-1} \lambda^{m-1} + \cdots + A_1 \lambda + A_0, \quad \text{με } A_m \neq 0.$$

Για παράδειγμα,

$$\begin{bmatrix} \lambda^2 + \lambda + 1 & \lambda^2 - \lambda + 2 \\ 2\lambda & \lambda^2 - 3\lambda - 1 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \lambda^2 + \begin{bmatrix} 1 & -1 \\ 2 & -3 \end{bmatrix} \lambda + \begin{bmatrix} 1 & 2 \\ 0 & -1 \end{bmatrix}.$$

¹Στα προηγούμενα κεφάλαια, η παράμετρος λ εμφανιζόταν πάντα αριστερά (μπροστά) από τον αντίστοιχο πίνακα, ως συντελεστής. Στο κεφάλαιο αυτό, η παράμετρος λ γράφεται πάντα δεξιά (μετά) από τον αντίστοιχο πίνακα, ως μεταβλητή.

Τετραγωνικοί πολυωνυμικοί πίνακες ίδιων διαστάσεων μπορούν να προστεθούν και να πολλαπλασιαστούν μεταξύ τους με το συνήθη τρόπο, και σε κάθε περίπτωση το αποτέλεσμα είναι ένας πολυωνυμικός πίνακας. Ειδικότερα, ένας $\nu \times \nu$ πολυωνυμικός πίνακας $P(\lambda)$ λέγεται *αντιστρέψιμος* (*unimodular*) αν υπάρχει ένας $\nu \times \nu$ πολυωνυμικός πίνακας $Q(\lambda)$ τέτοιος ώστε $P(\lambda)Q(\lambda) = Q(\lambda)P(\lambda) = I_\nu$.

Πρόταση 9.1.1. Ένας τετραγωνικός πολυωνυμικός πίνακας $P(\lambda)$ είναι αντιστρέψιμος αν και μόνο αν η ορίζουσα $\det(P(\lambda))$ είναι σταθερή και μη μηδενική.

Απόδειξη. Αν $\det(P(\lambda)) = c \neq 0$ ($c \in \mathbb{C}$), τότε ο $P(\lambda)^{-1} = c^{-1} \text{adj}(P(\lambda))$ είναι ένας πολυωνυμικός πίνακας. Αντίστροφα, αν ο $P(\lambda)$ είναι αντιστρέψιμος, τότε η ισότητα $P(\lambda)Q(\lambda) = I_\nu$ οδηγεί στη σχέση

$$\det(P(\lambda)Q(\lambda)) = \det(P(\lambda))\det(Q(\lambda)) = 1.$$

Επομένως, το γινόμενο των πολυωνύμων $\det(P(\lambda))$ και $\det(Q(\lambda))$ ταυτίζεται με το σταθερό πολυώνυμο $p(\lambda) = 1$. Αυτό είναι δυνατό αν μόνο αν τα δύο πολυώνυμα είναι σταθερά. $\square$

Για παράδειγμα, οι πολυωνυμικοί πίνακες

$$P(\lambda) = \begin{bmatrix} -1 & \lambda & -2\lambda^2 \\ 0 & 1 & \lambda^4 \\ 0 & 0 & 1 \end{bmatrix} \quad \text{και} \quad Q(\lambda) = \begin{bmatrix} (\lambda-1)^2 & \lambda \\ \lambda-2 & 1 \end{bmatrix}$$

είναι αντιστρέψιμοι με $\det(P(\lambda)) = -1$ και $\det(Q(\lambda)) = 1$. Οι αντιστρέψιμοι πολυωνυμικοί πίνακες παίζουν σημαντικό ρόλο στο μετασχηματισμό ενός τυχαίου πολυωνυμικού πίνακα σε απλούστερη μορφή.

Πρέπει να σημειωθεί ότι η έλλειψη της αντιμεταθετικότητας των πινάκων-συντελεστών των πολυωνυμικών πινάκων δημιουργεί αρκετές δυσκολίες στις διαδικασίες μετασχηματισμών τους. Παρ' όλα αυτά, τις τελευταίες τέσσερις δεκαετίες έχει αναπτυχθεί μία πλήρης θεωρία επί των πολυωνυμικών πινάκων, με σημαντικές εφαρμογές. Εισαγωγικά στοιχεία αυτής της θεωρίας θα παρουσιάσουμε στο υπόλοιπο του κεφαλαίου. Για να διατηρήσουμε τη ζωή μας απλή, σε αυτή τη γνωριμία με τους πολυωνυμικούς πίνακες, παραλείπουμε τις τεχνικές και περίπλοκες αποδείξεις (βλέπε τα Θεωρήματα 9.3.6 και 9.4.5, την Πρόταση 9.3.7 και τα αποτελέσματα της Παραγράφου 9.5).

9.2 Διαίρεση Πολυωνυμικών Πινάκων

Έστω ένας $\nu \times \nu$ πολυωνυμικός πίνακας $P(\lambda) = A_m\lambda^m + A_{m-1}\lambda^{m-1} + \dots + A_1\lambda + A_0$ βαθμού m . Αν $A_m = I_\nu$, τότε ο πολυωνυμικός πίνακας $P(\lambda)$

λέγεται *μόνικος* (*monic*). Αν $B(\lambda) = B_l \lambda^l + B_{l-1} \lambda^{l-1} + \cdots + B_1 \lambda + B_0$ είναι ένας $\nu \times \nu$ πολυωνυμικός πίνακας βαθμού l , τότε ο βαθμός του γινομένου

$$P(\lambda) B(\lambda) = A_m B_l \lambda^{m+l} + \cdots + (A_1 B_0 + A_0 B_1) \lambda + A_0 B_0$$

δεν υπερβαίνει το $m + l$. Προφανώς, αν ο A_m ή ο B_l είναι αντιστρέψιμος, τότε ο βαθμός του γινομένου είναι ακριβώς $m + l$.

Υποθέτουμε ότι ο $B(\lambda)$ είναι ένας πολυωνυμικός πίνακας βαθμού l με αντιστρέψιμο μεγιστοβάθμιο συντελεστή και ότι υπάρχουν πολυωνυμικοί πίνακες $Q(\lambda)$ και $R(\lambda)$ με $\deg R(\lambda) < l$, τέτοιοι ώστε

$$P(\lambda) = Q(\lambda) B(\lambda) + R(\lambda).$$

Τότε ο $Q(\lambda)$ καλείται *δεξιό πηλίκο* της διαίρεσης του $P(\lambda)$ με τον $B(\lambda)$, ενώ ο $R(\lambda)$ καλείται *δεξιό υπόλοιπο* της διαίρεσης του $P(\lambda)$ με τον $B(\lambda)$. Ομοίως, οι $\tilde{Q}(\lambda)$ και $\tilde{R}(\lambda)$ λέγονται *αριστερό πηλίκο* και *αριστερό υπόλοιπο* του $P(\lambda)$ στην διαίρεσή του με τον $B(\lambda)$, αν

$$P(\lambda) = B(\lambda) \tilde{Q}(\lambda) + \tilde{R}(\lambda)$$

και $\deg \tilde{R}(\lambda) < l$. Αν $R(\lambda) \equiv 0$ (ή $\tilde{R}(\lambda) \equiv 0$), τότε ο $B(\lambda)$ λέγεται *δεξιός* *διαίρετης* (αντίστοιχα, *αριστερός διαίρετης*) του $P(\lambda)$.

Θεώρημα 9.2.1. Έστω οι $\nu \times \nu$ πολυωνυμικοί πίνακες

$$P(\lambda) = A_m \lambda^m + A_{m-1} \lambda^{m-1} + \cdots + A_1 \lambda + A_0$$

και

$$B(\lambda) = B_l \lambda^l + B_{l-1} \lambda^{l-1} + \cdots + B_1 \lambda + B_0$$

με $\det(B_l) \neq 0$. Τότε υπάρχει δεξιό (αριστερό) πηλίκο και δεξιό (αριστερό) υπόλοιπο της διαίρεσης $P(\lambda)$ με τον $B(\lambda)$.

Απόδειξη. Αν $m < l$, τότε αρκεί μόνο να θέσουμε $Q(\lambda) = 0$ και $R(\lambda) = P(\lambda)$.

Αν $m \geq l$, τότε παρατηρούμε ότι ο μεγιστοβάθμιος όρος του πολυωνυμικού πίνακα $A_m B_l^{-1} \lambda^{m-l} B(\lambda)$ είναι $A_m \lambda^m$. Επομένως,

$$P(\lambda) = A_l B_l^{-1} \lambda^{m-l} B(\lambda) + P_1(\lambda),$$

όπου ο $P_1(\lambda)$ είναι ένας πολυωνυμικός πίνακας του οποίου ο βαθμός, έστω m_1 , δεν υπερβαίνει το $m - 1$. Γράφοντας τον $P_1(\lambda)$ σε φθίνουσες δυνάμεις, έχουμε

$$P_1(\lambda) = A_{m_1}^{(1)} \lambda^{m_1} + \cdots + A_1^{(1)} \lambda + A_0^{(1)} \quad (A_{m_1}^{(1)} \neq 0, \quad m_1 < m).$$

Αν $m_1 \geq l$, τότε επαναλαμβάνουμε τη διαδικασία στον $P_1(\lambda)$ αντί του $P(\lambda)$ και παίρνουμε

$$P_1(\lambda) = A_{m_1}^{(1)} B_l^{-1} \lambda^{m_1-l} B(\lambda) + P_2(\lambda),$$

όπου

$$P_2(\lambda) = A_{m_2}^{(2)} \lambda^{m_2} + \cdots + A_1^{(2)} \lambda + A_0^{(2)} \quad (A_{m_2}^{(2)} \neq 0, \quad m_2 < m_1 < l).$$

Επαναλαμβάνουμε την ίδια διαδικασία μέχρι να καταλήξουμε σε μία πεπερασμένη ακολουθία από πολυωνυμικούς πίνακες $P(\lambda), P_1(\lambda), P_2(\lambda), \dots, P_r(\lambda)$ με φθίνοντες βαθμούς και $\deg P_{r-1}(\lambda) \geq l > \deg P_r(\lambda)$. Αν συμβολίσουμε $P(\lambda) = P_0(\lambda)$, τότε προκύπτει ότι

$$P_{s-1}(\lambda) = A_{m_{s-1}}^{(s-1)} B_l^{-1} \lambda^{m_{s-1}-l} B(\lambda) + P_s(\lambda), \quad s = 1, 2, \dots, r.$$

Συνδυάζοντας αυτές τις εξισώσεις, παίρνουμε

$$P(\lambda) = (A_m B_l^{-1} \lambda^{m-l} + A_{m_1}^{(1)} B_l^{-1} \lambda^{m_1-l} + \cdots + A_{m_{r-1}}^{(r-1)} B_l^{-1} \lambda^{m_{r-1}-l}) B(\lambda) + P_r(\lambda).$$

Ο πολυωνυμικός πίνακας στην παρένθεση είναι το δεξιό πηλίκο της διαίρεσης του $P(\lambda)$ με τον $B(\lambda)$, ενώ ο $P_r(\lambda)$ είναι το δεξιό υπόλοιπο.

Με ανάλογο τρόπο αποδεικνύεται η ύπαρξη του αριστερού πηλίκου και υπολοίπου. $\square$

Πρόταση 9.2.2. Κάτω από τις υποθέσεις του Θεωρήματος 9.2.1, το δεξιό πηλίκο, το δεξιό υπόλοιπο, το αριστερό πηλίκο και το αριστερό υπόλοιπο είναι μοναδικά.

Απόδειξη. Υποθέτουμε ότι υπάρχουν $\nu \times \nu$ πολυωνυμικοί πίνακες $Q(\lambda), R(\lambda)$ και $Q_1(\lambda), R_1(\lambda)$ τέτοιοι ώστε

$$P(\lambda) = Q(\lambda) B(\lambda) + R(\lambda) \quad \text{και} \quad P(\lambda) = Q_1(\lambda) B(\lambda) + R_1(\lambda),$$

όπου καθένας από τους $R(\lambda)$ και $R_1(\lambda)$ έχει βαθμό μικρότερο από l . Τότε

$$(Q(\lambda) - Q_1(\lambda)) B(\lambda) = R_1(\lambda) - R(\lambda).$$

Αν $Q(\lambda) \neq Q_1(\lambda)$, τότε το αριστερό μέλος της εξίσωσης είναι ένας πολυωνυμικός πίνακας του οποίου ο βαθμός είναι τουλάχιστον l . Όμως, το δεξιό μέλος είναι ένας πολυωνυμικός πίνακας βαθμού μικρότερου από l και καταλήγουμε σε άτοπο. Συνεπώς, $Q_1(\lambda) = Q(\lambda)$ και $R_1(\lambda) = R(\lambda)$.

Η μοναδικότητα του αριστερού πηλίκου και υπολοίπου αποδεικνύεται με τον ίδιο τρόπο. $\square$

Παράδειγμα 9.2.1. Αν θεωρήσουμε τους πολυωνυμικούς πίνακες

$$P(\lambda) = \begin{bmatrix} \lambda^4 + \lambda^2 + \lambda - 1 & \lambda^3 + \lambda^2 + \lambda + 2 \\ 2\lambda^3 - \lambda & 2\lambda^2 + 2\lambda \end{bmatrix} \quad \text{και} \quad B(\lambda) = \begin{bmatrix} \lambda^2 + 1 & 1 \\ \lambda & \lambda^2 + \lambda \end{bmatrix},$$

τότε ισχύει

$$\begin{aligned} P(\lambda) &= \begin{bmatrix} \lambda^2 - 1 & \lambda - 1 \\ 2\lambda & 2 \end{bmatrix} \begin{bmatrix} \lambda^2 + 1 & 1 \\ \lambda & \lambda^2 + \lambda \end{bmatrix} + \begin{bmatrix} 2\lambda & 2\lambda + 3 \\ -5\lambda & -2\lambda \end{bmatrix} \\ &= Q(\lambda) B(\lambda) + R(\lambda) \end{aligned}$$

και

$$\begin{aligned} P(\lambda) &= \begin{bmatrix} \lambda^2 + 1 & 1 \\ \lambda & \lambda^2 + \lambda \end{bmatrix} \begin{bmatrix} \lambda^2 & \lambda + 1 \\ \lambda - 1 & 1 \end{bmatrix} \\ &= B(\lambda) \tilde{Q}(\lambda). \end{aligned}$$

Δηλαδή, ο $B(\lambda)$ είναι αριστερός, αλλά όχι δεξιός, διαιρέτης του $P(\lambda)$.

Στη συνέχεια, μελετάμε την περίπτωση των γραμμικών διαιρετών, δηλαδή των διαιρετών πρώτου βαθμού. Αν ο $P(\lambda)$ είναι ένας $\nu \times \nu$ πολυωνυμικός πίνακας βαθμού m όπως παραπάνω και B ένας $\nu \times \nu$ πίνακας, τότε ορίζουμε τη δεξιά τιμή $P(B)$ του $P(\lambda)$ στον B

$$P(B) = A_m B^m + A_{m-1} B^{m-1} + \cdots + A_1 B + A_0$$

και την αριστερή τιμή $\tilde{P}(B)$ του $P(\lambda)$ στον B

$$\tilde{P}(B) = B^m A_m + B^{m-1} A_{m-1} + \cdots + B A_1 + A_0.$$

Θεώρημα 9.2.3. Το δεξιό και αριστερό υπόλοιπο ενός $\nu \times \nu$ πολυωνυμικού πίνακα $P(\lambda)$, όταν αυτός διαιρείται με το μόνικο γραμμικό πολυωνυμικό πίνακα $I_\nu \lambda - B$, είναι $P(B)$ και $\tilde{P}(B)$, αντίστοιχα.

Απόδειξη. Η παραγοντοποίηση

$$I_\nu \lambda^j - B^j = (I_\nu \lambda^{j-1} + B \lambda^{j-2} + \cdots + B^{j-2} \lambda + B^{j-1})(I_\nu \lambda - B)$$

μπορεί να επαληθευθεί, εκτελώντας το γινόμενο στο δεξιό μέλος. Πολλαπλασιάζουμε από αριστερά και τα δύο μέλη της εξίσωσης με A_j και αθροίζουμε τις εξισώσεις που προκύπτουν για $j = 1, 2, \dots, m$. Το δεξιό μέλος της εξίσωσης που λαμβάνεται είναι της μορφής $C(\lambda)(I_\nu \lambda - B)$, όπου ο $C(\lambda)$ είναι ένας πολυωνυμικός πίνακας. Το αριστερό μέλος είναι

$$\sum_{j=1}^m A_j \lambda^j - \sum_{j=1}^m A_j B^j = \sum_{j=0}^m A_j \lambda^j - \sum_{j=0}^m A_j B^j = A(\lambda) - P(B).$$

Επομένως, $P(\lambda) = C(\lambda)(I_\nu \lambda - B) + P(B)$. Το συμπέρασμα έπεται από τη μοναδικότητα του δεξιού υπολοίπου της διαίρεσης του $P(\lambda)$ με τον $I_\nu \lambda - B$.

Το συμπέρασμα για το αριστερό υπόλοιπο λαμβάνεται αντιστρέφοντας τους παράγοντες στην αρχική παραγοντοποίηση, πολλαπλασιάζοντας από δεξιά με τον πίνακα A_j και αθροίζοντας. $\square$

Ένας $\nu \times \nu$ πίνακας X τέτοιος ώστε $P(X) = 0$ (ή $\tilde{P}(X) = 0$) καλείται δεξιά ρίζα (αντίστοιχα, αριστερή ρίζα) του πολυωνυμικού πίνακα $P(\lambda)$.

Πόρισμα 9.2.4. Ένας $\nu \times \nu$ πολυωνυμικός πίνακας $P(\lambda)$ διαιρείται ακριβώς από δεξιά (ή από αριστερά) από τον $I_\nu \lambda - B$ αν και μόνο αν ο πίνακας B είναι δεξιά (αντίστοιχα, αριστερή) ρίζα του $P(\lambda)$.

9.3 Ισοδυναμίες και Κανονική Μορφή Smith

Σε αυτή την παράγραφο, θα μελετήσουμε τη δυνατότητα αναγωγής ενός πολυωνυμικού πίνακα σε απλούστερη μορφή, μέσω συγκεκριμένων μετασχηματισμών, επεκτείνοντας τη θεωρία που γνωρίζουμε για τους πίνακες. Αρχικά ορίζουμε τους στοιχειώδεις μετασχηματισμούς γραμμών και στηλών ενός πολυωνυμικού πίνακα:

- (1) Πολλαπλασιασμός μιας γραμμής (ή στήλης) με ένα μη μηδενικό αριθμό $c \in \mathbb{C}$.
- (2) Εναλλαγή δύο γραμμών (ή στηλών).
- (3) Πρόσθεση σε μία γραμμή (ή στήλη) μιας οποιαδήποτε άλλης γραμμής (αντίστοιχα, στήλης) πολλαπλασιασμένης με ένα αυθαίρετο πολώνυμο $b(\lambda)$.

Είναι εύκολο να επαληθεύσουμε ότι η εφαρμογή των στοιχειωδών μετασχηματισμών (1)–(3) πάνω σε ένα $\nu \times \nu$ πολυωνυμικό πίνακα είναι ισοδύναμη με τον πολλαπλασιασμό, από αριστερά για τις γραμμές και από δεξιά για τις στήλες, του πολυωνυμικού πίνακα με έναν από τους ακόλουθους $\nu \times \nu$ πίνακες:

$$E^{(1)} = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & & c & \\ & & & & 1 & \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix}, \quad E^{(2)} = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 0 & \cdots & 1 \\ & & \vdots & & \vdots \\ & & 1 & \cdots & 0 \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix},$$

$$E^{(3)}(\lambda) = \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & \cdots & b(\lambda) \\ & & & \ddots & \vdots \\ & & & & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix} \quad \text{ή} \quad \begin{bmatrix} 1 & & & & \\ & \ddots & & & \\ & & 1 & & \\ & & \vdots & \ddots & \\ & & b(\lambda) & \cdots & 1 \\ & & & & & \ddots \\ & & & & & & 1 \end{bmatrix}.$$

Οι παραπάνω πίνακες λέγονται *στοιχειώδεις πίνακες* πρώτου, δεύτερου και τρίτου τύπου, αντίστοιχα, και προφανώς παράγονται με την εφαρμογή των αντιστοιχών μετασχηματισμών στο μοναδιαίο πίνακα. Επιπλέον, εύκολα μπορεί κανείς να επαληθεύσει την ακόλουθη πρόταση.

Πρόταση 9.3.1. *Οι στοιχειώδεις πίνακες είναι αντιστρέψιμοι και οι αντίστροφοι τους είναι επίσης στοιχειώδεις πίνακες.*

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Δύο $\nu \times \mu$ πολυωνυμικοί πίνακες $P(\lambda)$ και $Q(\lambda)$ καλούνται *ισοδύναμοι* αν ο $Q(\lambda)$ μπορεί να παραχθεί από τον $P(\lambda)$ μέσω μιας πεπερασμένης ακολουθίας στοιχειωδών μετασχηματισμών. Από την ισοδυναμία των μετασχηματισμών αυτών και των πολλαπλασιασμών με τους στοιχειώδεις πίνακες που περιγράφονται παραπάνω, προκύπτει ότι οι $P(\lambda)$ και $Q(\lambda)$ είναι ισοδύναμοι αν και μόνο αν υπάρχουν στοιχειώδεις πίνακες $E_1(\lambda), E_2(\lambda), \dots, E_k(\lambda), E_{k+1}(\lambda), \dots, E_s(\lambda)$ τέτοιοι ώστε

$$Q(\lambda) = E_k(\lambda) E_{k-1}(\lambda) \cdots E_1(\lambda) P(\lambda) E_{k+1}(\lambda) E_{k+2}(\lambda) \cdots E_s(\lambda),$$

ή

$$Q(\lambda) = M(\lambda) P(\lambda) N(\lambda), \quad (9.3.1)$$

όπου οι πολυωνυμικοί πίνακες

$$M(\lambda) = E_k(\lambda) E_{k-1}(\lambda) \cdots E_1(\lambda) \quad \text{και} \quad N(\lambda) = E_{k+1}(\lambda) E_{k+2}(\lambda) \cdots E_s(\lambda)$$

είναι αντιστρέψιμοι από την Πρόταση 9.3.1. Έτσι, προκύπτουν οι επόμενες προτάσεις.

Πρόταση 9.3.2. *Δύο $\nu \times \mu$ πολυωνυμικοί πίνακες $P(\lambda)$ και $Q(\lambda)$ είναι ισοδύναμοι αν και μόνο αν υπάρχουν αντιστρέψιμοι πίνακες $M(\lambda)$ και $N(\lambda)$ τέτοιοι ώστε να ισχύει η (9.3.1) .*

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πόρισμα 9.3.3. *Αν $\mu \in P(\lambda) \sim Q(\lambda)$ συμβολίσουμε τη σχέση ισοδυναμίας του πολυωνυμικού πίνακα $P(\lambda)$ με τον $Q(\lambda)$, τότε ισχύουν τα ακόλουθα:*

- (a) $P(\lambda) \sim P(\lambda)$.
- (b) Αν $P(\lambda) \sim Q(\lambda)$, τότε $Q(\lambda) \sim P(\lambda)$.
- (c) Αν $P(\lambda) \sim Q(\lambda)$ και $Q(\lambda) \sim R(\lambda)$, τότε $P(\lambda) \sim R(\lambda)$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πόρισμα 9.3.4. Κάθε αντιστρέψιμος πολυωνυμικός πίνακας είναι ισοδύναμος με το μοναδιαίο πίνακα.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Πόρισμα 9.3.5. Ένας τετραγωνικός πολυωνυμικός πίνακας είναι αντιστρέψιμος αν και μόνο αν μπορεί να αναλυθεί σε γινόμενο στοιχειωδών πινάκων.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Παράδειγμα 9.3.1. Θεωρούμε τους πολυωνυμικούς πίνακες

$$P(\lambda) = \begin{bmatrix} \lambda & \lambda+1 \\ \lambda^2 - \lambda & \lambda^2 - 1 \end{bmatrix} \quad \text{και} \quad Q(\lambda) = Q = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}.$$

Στον $P(\lambda)$ προσθέτουμε $\lambda - 1$ φορές την πρώτη γραμμή στη δεύτερη και έπειτα αφαιρούμε την πρώτη στήλη από τη δεύτερη. Τελικά, προσθέτοντας $-\lambda$ φορές τη δεύτερη στήλη του νέου πολυωνυμικού πίνακα στην πρώτη και εναλλάσσοντας στήλες, προκύπτει ο Q . Χρησιμοποιώντας τους στοιχειώδεις πίνακες και υπολογίζοντας τα γινόμενα τους, είναι εύκολο να βρούμε τους

$$M(\lambda) = \begin{bmatrix} -1 & 1+\lambda \\ 1 & -\lambda \end{bmatrix} \quad \text{και} \quad N(\lambda) = \begin{bmatrix} 1 & 0 \\ 1-\lambda & 1 \end{bmatrix}$$

οι οποίοι ικανοποιούν την (9.3.1).

Ένας $\nu \times \nu$ πολυωνυμικός πίνακας ονομάζεται *κανονικός (canonical)* αν είναι της μορφής

$$A_0(\lambda) = \text{diag} \{ \alpha_1(\lambda), \alpha_2(\lambda), \dots, \alpha_\nu(\lambda) \},$$

όπου το κάθε $\alpha_j(\lambda)$ είναι μηδενικό ή μόνικο πολυώνυμο και διαιρείται ακριβώς από το $\alpha_{j-1}(\lambda)$ ($j = 1, 2, \dots, \nu$). Παρατηρούμε ότι αν υπάρχουν μηδενικά πολυώνυμα μεταξύ των $\alpha_j(\lambda)$, τότε αυτά πρέπει να συναντώνται στις τελευταίες θέσεις. Επιπλέον, αν κάποια από τα πολυώνυμα $\alpha_j(\lambda)$ είναι (μη μηδενικά) σταθερά, τότε πρέπει να είναι ίσα με 1 και πρέπει να τοποθετηθούν στις πρώτες θέσεις του κανονικού πολυωνυμικού πίνακα. Επομένως, ένας κανονικός πολυωνυμικός πίνακας είναι γενικά της μορφής

$$A_0(\lambda) = \text{diag} \{ 1, \dots, 1, \alpha_1(\lambda), \dots, \alpha_k(\lambda), 0, \dots, 0 \}, \quad (9.3.2)$$

όπου κάθε $\alpha_j(\lambda)$ είναι ένα μόνικο πολυώνυμο βαθμού τουλάχιστον 1 και διαιρείται ακριβώς από το $\alpha_{j-1}(\lambda)$ ($j = 1, 2, \dots, k$). Η τάξη ενός τυχαίου πολυωνυμικού πίνακα ορίζεται ως ο βαθμός της μεγαλύτερης υποορίζουσάς του που δεν είναι το μηδενικό πολυώνυμο. Για παράδειγμα, η τάξη του πολυωνυμικού πίνακα $A_0(\lambda) = \text{diag} \{ \alpha_1(\lambda), \dots, \alpha_r(\lambda), 0, \dots, 0 \}$ είναι ίση με r .

Θεώρημα 9.3.6. Κάθε $\nu \times \nu$ πολωνυμικός πίνακας είναι ισοδύναμος με έναν κανονικό πολωνυμικό πίνακα της μορφής (9.3.2), ο οποίος είναι μοναδικός. Επιπλέον, η τάξη ενός πολωνυμικού πίνακα διατηρείται αναλλοίωτη στους μετασχηματισμούς ισοδυναμίας.

Παρατηρούμε ότι ένας πολωνυμικός πίνακας είναι ομαλός αν και μόνο αν όλα τα διαγώνια στοιχεία της κανονικής του μορφής (9.3.2) είναι μη μηδενικά πολυώνυμα.

Ας θεωρήσουμε τώρα ένα $\nu \times \nu$ πολωνυμικό πίνακα $P(\lambda)$ τάξης r και ας συμβολίσουμε με $d_j(\lambda)$ το μέγιστο κοινό διαιρέτη όλων των υποοριζουσών του $P(\lambda)$ τάξης j ($j = 1, 2, \dots, r$). Αφού κάθε υποορίζουσα τάξης $j \geq 2$ μπορεί να εκφραστεί ως γραμμικός συνδυασμός των υποοριζουσών τάξεως $j - 1$, το $d_{j-1}(\lambda)$ διαιρεί ακριβώς το $d_j(\lambda)$. Επομένως, αν θέσουμε $d_0(\lambda) \equiv 1$, τότε στην ακολουθία

$$d_0(\lambda), d_1(\lambda), d_2(\lambda), \dots, d_r(\lambda),$$

το $d_j(\lambda)$ διαιρείται ακριβώς από το $d_{j-1}(\lambda)$ ($j = 1, 2, \dots, r$).

Παρατήρηση 9.3.1. Τα πολυώνυμα $d_j(\lambda)$, όπως αυτά ορίστηκαν παραπάνω, του κανονικού πολωνυμικού πίνακα $A_0(\lambda) = \text{diag} \{\alpha_1(\lambda), \alpha_2(\lambda), \dots, \alpha_r(\lambda), 0, \dots, 0\}$ είναι

$$d_1(\lambda) = \alpha_1(\lambda), \quad d_2(\lambda) = \alpha_1(\lambda) \alpha_2(\lambda), \quad \dots, \quad d_r(\lambda) = \alpha_1(\lambda) \alpha_2(\lambda) \cdots \alpha_r(\lambda).$$

Τα πολυώνυμα $d_0(\lambda), d_1(\lambda), \dots, d_r(\lambda)$ έχουν τη σημαντική ιδιότητα να μένουν αναλλοίωτα υπό την εφαρμογή των μετασχηματισμών ισοδυναμίας.

Πρόταση 9.3.7. Έστω δύο ισοδύναμοι $\nu \times \nu$ πολωνυμικοί πίνακες $P(\lambda)$ και $Q(\lambda)$ τάξης k . Θεωρούμε $d_j(\lambda)$ και $\delta_j(\lambda)$ τους μεγαλύτερους μόνιμους κοινούς διαιρέτες όλων των υποοριζουσών βαθμού j ($j = 1, 2, \dots, r$) των $P(\lambda)$ και $Q(\lambda)$, αντίστοιχα. Τότε τα πολυώνυμα $d_j(\lambda)$ και $\delta_j(\lambda)$ ($j = 1, 2, \dots, r$) ταυτίζονται ένα προς ένα.

Ας θεωρήσουμε τώρα τα πηλίκα

$$i_1(\lambda) = \frac{d_1(\lambda)}{d_0(\lambda)}, \quad i_2(\lambda) = \frac{d_2(\lambda)}{d_1(\lambda)}, \quad \dots, \quad i_r(\lambda) = \frac{d_r(\lambda)}{d_{r-1}(\lambda)}.$$

Λόγω της διαιρετότητας του πολωνύμου $d_j(\lambda)$ από το $d_{j-1}(\lambda)$, τα πηλίκα $i_j(\lambda)$ ($j = 1, 2, \dots, r$) είναι πολυώνυμα και καλούνται αναλλοίωτα πολυώνυμα (*invariant polynomials*) του $P(\lambda)$, όπου το αναλλοίωτο αναφέρεται στους μετασχηματισμούς ισοδυναμίας. Αξίζει να σημειωθεί ότι

$$d_j(\lambda) = i_1(\lambda) i_2(\lambda) \cdots i_j(\lambda), \quad j = 1, 2, \dots, r \quad (9.3.3)$$

και κάθε αναλλοίωτο πολυώνυμο $i_j(\lambda)$ διαιρείται ακριβώς από το $i_{j-1}(\lambda)$ ($j = 2, 3, \dots, r$).

Παράδειγμα 9.3.2. Τα αναλλοίωτα πολυώνυμα του 3×3 πολυωνυμικού πίνακα

$$P(\lambda) = \begin{bmatrix} 0 & 1 & \lambda \\ \lambda & \lambda & 1 \\ \lambda^2 - \lambda & \lambda^2 - 1 & \lambda^2 - 1 \end{bmatrix}$$

είναι $i_1(\lambda) = i_2(\lambda) = 1$, ενώ τα αναλλοίωτα πολυώνυμα του

$$Q(\lambda) = \begin{bmatrix} \lambda & \lambda^2 & 0 \\ \lambda^3 & \lambda^5 & 0 \\ 0 & 0 & 2\lambda \end{bmatrix}$$

είναι $i_1(\lambda) = i_2(\lambda) = \lambda$ και $i_3(\lambda) = \lambda^5 - \lambda^4$.

Σύμφωνα με την Πρόταση 9.3.7, και παρατηρώντας ότι τα αναλλοίωτα πολυώνυμα ενός ομαλού πολυωνυμικού πίνακα ταυτίζονται με τα πολυώνυμα $\alpha_j(\lambda)$ της κανονικής του μορφής, παίρνουμε άμεσα το ακόλουθο αποτέλεσμα.

Θεώρημα 9.3.8. Ένας $\nu \times \nu$ πολυωνυμικός πίνακας $P(\lambda)$ τάξης r είναι ισοδύναμος με τον πολυωνυμικό πίνακα $\text{diag}\{i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda), 0, \dots, 0\}$, όπου τα $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$ είναι τα αναλλοίωτα πολυώνυμα του $P(\lambda)$.

Η κανονική μορφή που περιγράφεται στο παραπάνω θεώρημα είναι γνωστή ως κανονική μορφή *Smith* του $P(\lambda)$.

Πόρισμα 9.3.9. Δύο $\nu \times \nu$ πολυωνυμικοί πίνακες είναι ισοδύναμοι αν και μόνο αν έχουν ακριβώς τα ίδια αναλλοίωτα πολυώνυμα.

Απόδειξη. Το ευθύ αποτελεί την Πρόταση 9.3.7. Αντίστροφα, αν δύο πολυωνυμικοί πίνακες έχουν τα ίδια αναλλοίωτα πολυώνυμα, τότε από το Θεώρημα 9.3.8 έπεται ότι θα έχουν και την ίδια κανονική μορφή *Smith*. Η μεταβατική ιδιότητα των σχέσεων ισοδυναμίας (βλέπε το Πόρισμα 9.3.3 (c)) μας δίνει ότι οι πίνακες είναι ισοδύναμοι. $\square$

Η *Smith* κανονική μορφή των πινάκων του Παραδείγματος 9.3.2 δίνεται από τις σχέσεις

$$P(\lambda) \sim \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \quad \text{και} \quad Q(\lambda) \sim \begin{bmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & \lambda^5 - \lambda^4 \end{bmatrix}.$$

Έστω ένας $\nu \times \nu$ πολυωνυμικός πίνακας $P(\lambda)$ τάξης r ($\leq \nu$), με αναλλοίωτα πολυώνυμα $i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda)$. Ορίζουμε ιδιοτιμές του $P(\lambda)$ τις ρίζες

της $\det(P(\lambda))$, και αφού η ορίζουσα $\det(P(\lambda))$ είναι ένα (μιγαδικό) πολυώνυμο, μπορούμε να γράψουμε

$$\det(P(\lambda)) = c \prod_{j=1}^s (\lambda - \lambda_j)^{m_j},$$

όπου $c \neq 0$, $\lambda_1, \lambda_2, \dots, \lambda_s$ είναι οι διακεκριμένες ιδιοτιμές του $P(\lambda)$ και $m_j \geq 1$ για κάθε $j = 1, 2, \dots, s$. Από την κανονική μορφή Smith προκύπτει ότι

$$\det(P(\lambda)) = c' \prod_{j=1}^r i_j(\lambda),$$

όπου $c' \neq 0$. Αφού τα αναλλοίωτα πολυώνυμα είναι μόνικα, έπεται ότι $c = c'$ και

$$\prod_{j=1}^r i_j(\lambda) = \prod_{k=1}^s (\lambda - \lambda_k)^{m_k}.$$

Επιπλέον, αφού κάθε $i_j(\lambda)$ διαιρεί ακριβώς το $i_{j+1}(\lambda)$ ($j = 1, 2, \dots, r-1$), υπάρχουν ακέραιοι α_{jk} ($1 \leq j \leq r$, $1 \leq k \leq s$) τέτοιοι ώστε

$$\begin{aligned} i_1(\lambda) &= (\lambda - \lambda_1)^{\alpha_{11}} (\lambda - \lambda_2)^{\alpha_{12}} \dots (\lambda - \lambda_s)^{\alpha_{1s}}, \\ i_2(\lambda) &= (\lambda - \lambda_1)^{\alpha_{21}} (\lambda - \lambda_2)^{\alpha_{22}} \dots (\lambda - \lambda_s)^{\alpha_{2s}}, \\ &\vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \quad \quad \quad \vdots \\ i_r(\lambda) &= (\lambda - \lambda_1)^{\alpha_{r1}} (\lambda - \lambda_2)^{\alpha_{r2}} \dots (\lambda - \lambda_s)^{\alpha_{rs}}, \end{aligned} \quad (9.3.4)$$

όπου για κάθε $k = 1, 2, \dots, s$, ισχύουν

$$0 \leq \alpha_{1k} \leq \alpha_{2k} \leq \dots \leq \alpha_{rk} \leq m_k \quad \text{και} \quad \alpha_{1k} + \alpha_{2k} + \dots + \alpha_{rk} = m_k.$$

Κάθε παράγοντας $(\lambda - \lambda_k)^{\alpha_{jk}}$ που εμφανίζεται στη σχέση (9.3.4) με $\alpha_{jk} > 0$ καλείται *στοιχειώδης διαιρέτης* του $P(\lambda)$. Από το Πρόσχημα 9.3.9 και τη στενή σχέση των στοιχειωδών διαιρέτων με τα αναλλοίωτα πολυώνυμα, προκύπτει ότι οι στοιχειώδεις διαιρέτες διατηρούνται αμετάβλητοι στους μετασχηματισμούς ισοδυναμίας των πολυωνυμικών πινάκων.

Θεώρημα 9.3.10. Δύο $\nu \times \nu$ πολυωνυμικοί πίνακες είναι ισοδύναμοι αν και μόνο αν έχουν τους ίδιους στοιχειώδεις διαιρέτες.

Παράδειγμα 9.3.3. Έστω ένας πολυωνυμικός πίνακας βαθμού 5 και τάξης 4, με στοιχειώδεις διαιρέτες λ , λ , λ^3 , $\lambda - 2$, $(\lambda - 2)^2$, $\lambda + 5$. Τότε αφού ο βαθμός του είναι 5, υπάρχουν 5 αναλλοίωτα πολυώνυμα $i_k(\lambda)$ ($k = 1, 2, \dots, 5$). Επιπλέον, αφού η τάξη είναι 4, έχουμε $i_5(\lambda) = d_5(\lambda)/d_4(\lambda) \equiv 0$. Ξέρουμε ακόμη ότι το αναλλοίωτο πολυώνυμο $i_j(\lambda)$ διαιρείται με το $i_{j-1}(\lambda)$ ($j \geq 2$). Επομένως, $i_4(\lambda) = \lambda^5(\lambda - 2)^2(\lambda + 5)$. Ομοίως για το $i_3(\lambda)$ και τους υπόλοιπους στοιχειώδεις διαιρέτες, παίρνουμε $i_3(\lambda) = \lambda(\lambda - 2)$ και μετά $i_2(\lambda) = \lambda$. Οι στοιχειώδεις διαιρέτες έχουν εξαντληθεί και καταλήγουμε στο ότι $i_1(\lambda) \equiv 1$.

Θεώρημα 9.3.11. Αν $P(\lambda)$ και $Q(\lambda)$ είναι δύο τετραγωνικοί πολυωνυμικοί πίνακες, τότε οι στοιχειώδεις διαιρέτες του

$$R(\lambda) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & Q(\lambda) \end{bmatrix}$$

είναι η ένωση των στοιχειωδών διαιρετών των $P(\lambda)$ και $Q(\lambda)$.

Απόδειξη. Έστω $D_P(\lambda)$ και $D_Q(\lambda)$ οι κανονικές μορφές Smith των $P(\lambda)$ και $Q(\lambda)$, αντίστοιχα. Δηλαδή, ισχύει

$$R(\lambda) = E(\lambda) \begin{bmatrix} D_P(\lambda) & 0 \\ 0 & D_Q(\lambda) \end{bmatrix} F(\lambda)$$

για κάποιους αντιστρέψιμους πολυωνυμικούς πίνακες $E(\lambda)$ και $F(\lambda)$. Έστω $(\lambda - \lambda_0)^{\alpha_1}, (\lambda - \lambda_0)^{\alpha_2}, \dots, (\lambda - \lambda_0)^{\alpha_s}$ και $(\lambda - \lambda_0)^{\beta_1}, (\lambda - \lambda_0)^{\beta_2}, \dots, (\lambda - \lambda_0)^{\beta_t}$ οι στοιχειώδεις διαιρέτες των $D_P(\lambda)$ και $D_Q(\lambda)$, αντίστοιχα, της ιδιοτιμής λ_0 . Τοποθετούμε τους εκθέτες $\alpha_1, \alpha_2, \dots, \alpha_s, \beta_1, \beta_2, \dots, \beta_t$ σε αύξουσα σειρά

$$\{\alpha_1, \dots, \alpha_s, \beta_1, \dots, \beta_t\} = \{\gamma_1, \gamma_2, \dots, \gamma_{s+t}\}, \quad 0 \leq \gamma_1 \leq \gamma_2 \leq \dots \leq \gamma_{s+t}.$$

Από τον ορισμό των αναλλοίωτων πολυωνύμων και τη μοναδικότητα της κανονικής μορφής Smith $D = \text{diag}\{i_1(\lambda), i_2(\lambda), \dots, i_r(\lambda), 0, \dots, 0\}$ του πολυωνυμικού πίνακα $\text{diag}\{D_P(\lambda), D_Q(\lambda)\}$, ολοκληρώνεται η απόδειξη. $\square$

Πόρισμα 9.3.12. Οι αντιστρέψιμοι πολυωνυμικοί πίνακες, και μόνο αυτοί, δεν έχουν στοιχειώδεις διαιρέτες.

Πόρισμα 9.3.13. Για κάθε διαγώνιο πίνακα $D \in \mathbb{C}^{\nu \times \nu}$, οι στοιχειώδεις διαιρέτες του $I_\nu \lambda - D$ είναι γραμμικοί.

9.4 Standard Ζεύγη και Τριάδες

Για ένα (μιγαδικό) μόνικο πολώνυμο $p(\lambda) = \lambda^m + a_{m-1}\lambda^{m-1} + \dots + a_1\lambda + a_0$, ο $m \times m$ πίνακας

$$C_p = \begin{bmatrix} 0 & 1 & \dots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & & 1 \\ -a_0 & -a_1 & \dots & -a_{m-1} \end{bmatrix} \quad (9.4.1)$$

είναι ο συνοδεύοντας πίνακας (*companion matrix*) του $p(\lambda)$. Είναι γνωστό ότι οι ρίζες του $p(\lambda)$ ταυτίζονται με τις ιδιοτιμές του C_p (λαμβάνοντας υπόψη και τις αλγεβρικές πολλαπλότητες). Για ένα $\nu \times \nu$ πολυωνυμικό πίνακα

$$P(\lambda) = A_m \lambda^m + A_{m-1} \lambda^{m-1} + \dots + A_1 \lambda + A_0$$

με $\det(A_m) \neq 0$, ορίζουμε τη γενίκευση του πίνακα της (9.4.1),

$$C_P = \begin{bmatrix} 0 & I_\nu & \cdots & 0 \\ 0 & 0 & \ddots & \vdots \\ \vdots & \vdots & & I_\nu \\ -A_m^{-1}A_0 & -A_m^{-1}A_1 & \cdots & -A_m^{-1}A_{m-1} \end{bmatrix}. \quad (9.4.2)$$

Ο $m\nu \times m\nu$ πίνακας C_P καλείται (πρώτος) συνοδεύοντας πίνακας (*first companion matrix*) του $P(\lambda)$.

Από το σημείο αυτό και στο υπόλοιπο του κεφαλαίου, θεωρούμε πάντα (χωρίς να χρειάζεται να το αναφέρουμε κάθε φορά) ότι ο πίνακας-συντελεστής A_m του μεγιστοβάθμιου όρου του $\nu \times \nu$ πολυωνυμικού πίνακα $P(\lambda)$ είναι αντιστρέψιμος.

Η πρώτη παρατήρηση είναι ότι τα χαρακτηριστικά πολυώνυμα ικανοποιούν την ταυτότητα

$$\det(P(\lambda)) = \det(I_{m\nu}\lambda - C_P) \det(A_m). \quad (9.4.3)$$

Αυτό σημαίνει ότι οι ιδιοτιμές του $P(\lambda)$ συμπίπτουν με τις ιδιοτιμές του C_P . Όπως και στην περίπτωση των πινάκων, το σύνολο των ιδιοτιμών του $P(\lambda)$ καλείται *φάσμα* του $P(\lambda)$ και συμβολίζεται $\sigma(P)$. Έτσι, σύμφωνα με τα όσα έχουμε αναφέρει, η σχέση (9.4.3) μας δίνει ότι ο $P(\lambda)$ και ο $I_{m\nu}\lambda - C_P$ έχουν τα ίδια ακριβώς αναλλοίωτα πολυώνυμα μεγίστου βαθμού.

Θεώρημα 9.4.1. Οι $m\nu \times m\nu$ πολυωνυμικοί πίνακες

$$\begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(m-1)\nu} \end{bmatrix} \quad \text{και} \quad I_{m\nu}\lambda - C_P$$

είναι ισοδύναμοι.

Απόδειξη. Ορίζουμε αρχικά τους $m\nu \times m\nu$ πολυωνυμικούς πίνακες

$$F(\lambda) = \begin{bmatrix} I_\nu & 0 & \cdots & 0 & 0 \\ -I_\nu\lambda & I_\nu & \cdots & 0 & 0 \\ 0 & -I_\nu\lambda & \ddots & & \vdots \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & -I_\nu\lambda & I_\nu \end{bmatrix}$$

και

$$E(\lambda) = \begin{bmatrix} B_{m-1}(\lambda) & B_{m-2}(\lambda) & \cdots & B_1(\lambda) & B_0(\lambda) \\ -I_\nu & 0 & \cdots & 0 & 0 \\ 0 & -I_\nu & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & -I_\nu & 0 \end{bmatrix}$$

όπου $B_0(\lambda) = A_m$ και $B_{s+1}(\lambda) = \lambda B_s(\lambda) + A_{m-s-1}$ για $s = 0, 1, \dots, m-2$. Προφανώς,

$$\det(F(\lambda)) = 1 \quad \text{και} \quad \det(E(\lambda)) = \pm \det(A_m), \quad (9.4.4)$$

δηλαδή οι πολυωνυμικοί πίνακες $F(\lambda)$ και $E(\lambda)$ είναι αντιστρέψιμοι. Άρα ο $F(\lambda)^{-1}$ είναι επίσης ένας πολυωνυμικός πίνακας. Αποδεικνύεται εύκολα ότι

$$E(\lambda) (I_{m\nu}\lambda - C_P) = \begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(m-1)\nu} \end{bmatrix} F(\lambda) \quad (9.4.5)$$

και έτσι

$$\begin{bmatrix} P(\lambda) & 0 \\ 0 & I_{(m-1)\nu} \end{bmatrix} = E(\lambda) (I_{m\nu}\lambda - C_P) F(\lambda)^{-1},$$

ολοκληρώνοντας την απόδειξη. $\square$

Το Θεώρημα 9.4.1 (καθώς και το Θεώρημα 9.3.8 με το πόρισμα του) δείχνει ότι όλα τα αναλλοίωτα πολυώνυμα, άρα και όλοι οι στοιχειώδεις διαιρέτες, των $P(\lambda)$ και $I_{m\nu}\lambda - C_P$ συμπίπτουν.

Ας θεωρήσουμε τώρα έναν $m\nu \times m\nu$ πίνακα A όμοιο με τον C_L , δηλαδή $A = T C_P T^{-1}$ για κάποιον αντιστρέψιμο πίνακα T . Τότε $I_{m\nu}\lambda - A = T(I_{m\nu}\lambda - C_P)T^{-1}$ και οι πρωτοβάθμιοι πολυωνυμικοί πίνακες $I_{m\nu}\lambda - A$ και $I_{m\nu}\lambda - C_P$ είναι ισοδύναμοι. Έτσι, σύμφωνα με τη μεταβατική ιδιότητα της ισοδυναμίας προκύπτει επίσης ότι οι $I_{m\nu}\lambda - A$ και $\text{diag}\{P(\lambda), I_{(m-1)\nu}\}$ είναι ισοδύναμοι. Κάθε $m\nu \times m\nu$ πίνακας A που ικανοποιεί αυτήν τη τελευταία ιδιότητα καλείται *γραμμικοποιητής* του $P(\lambda)$.

Θεώρημα 9.4.2. Ένας $m\nu \times m\nu$ πίνακας A είναι γραμμικοποιητής του $P(\lambda)$ αν και μόνο αν ο A είναι όμοιος με το συνοδεύοντα πίνακα C_P του $P(\lambda)$.

Απόδειξη. Αφήνεται στον αναγνώστη, ως άσκηση. $\square$

Από το θεώρημα αυτό γίνεται φανερό ότι υπάρχει ένας πίνακας Jordan ο οποίος είναι γραμμικοποιητής του $P(\lambda)$ (βλέπε το Θεώρημα 6.7.1) και περιέχει πληροφορίες που αφορούν τις ιδιοτιμές του $P(\lambda)$.

Παράδειγμα 9.4.1. Έστω ο πολυωνυμικός πίνακας $P(\lambda) = \begin{bmatrix} \lambda^2 & -\lambda \\ 0 & \lambda^2 \end{bmatrix}$. Ο πίνακας Jordan που αποτελεί γραμμικοποίηση του $P(\lambda)$ είναι

$$\begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix} \quad \text{ή} \quad \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}.$$

Στη φασματική θεωρία ενός πίνακα $A \in \mathbb{C}^{\nu \times \nu}$, σε κάποιες περιπτώσεις, η επιλύουσα συνάρτηση $(I_\nu \lambda - A)^{-1}$ παίζει σημαντικό ρόλο. Για έναν πολωνυμικό πίνακα $P(\lambda)$, η συνάρτηση $P(\lambda)^{-1}$, $\lambda \notin \sigma(P)$, καλείται *επιλύουσα συνάρτηση* (*resolvent*) του $P(\lambda)$.

Θεώρημα 9.4.3. Για κάθε μιγαδικό $\lambda \notin \sigma(P)$, ισχύει

$$P(\lambda)^{-1} = P_1(I_{m\nu}\lambda - C_P)^{-1}R_1, \quad (9.4.6)$$

όπου οι P_1 και R_1 είναι πίνακες διαστάσεων $\nu \times m\nu$ και $m\nu \times \nu$, αντίστοιχα, και δίνονται από τις σχέσεις

$$P_1 = [I_\nu \ 0 \ \cdots \ 0] \quad \text{και} \quad R_1 = \begin{bmatrix} 0 \\ \vdots \\ 0 \\ A_m^{-1} \end{bmatrix}. \quad (9.4.7)$$

Απόδειξη. Αρχικά παρατηρούμε ότι λόγω της (9.4.3), ο $(I_{m\nu}\lambda - C_P)^{-1}$ ορίζεται ακριβώς στα λ όπου ορίζεται ο $P(\lambda)^{-1}$. Η σχέση (9.4.5) συνεπάγεται ότι

$$\begin{bmatrix} P(\lambda)^{-1} & 0 \\ 0 & I_{(m-1)\nu} \end{bmatrix} = F(\lambda)(I_{m\nu}\lambda - C_P)^{-1}E(\lambda)^{-1}. \quad (9.4.8)$$

Χρησιμοποιώντας τον ορισμό του $E(\lambda)$ διαπιστώνουμε ότι

$$E(\lambda) \begin{bmatrix} 0 \\ \vdots \\ 0 \\ A_m^{-1} \end{bmatrix} = \begin{bmatrix} I_\nu \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

το οποίο σημαίνει ότι οι πρώτες ν στήλες του $E(\lambda)^{-1}$ σχηματίζουν τον R_1 . Εφόσον οι πρώτες ν γραμμές του $F(\lambda)$ σχηματίζουν τον P_1 , αν εξισώσουμε τους αρχικούς $\nu \times \nu$ υποπίνακες στα αριστερά και δεξιά της (9.4.8), τότε λαμβάνουμε την (9.4.6). $\square$

Μια τριάδα πινάκων (U, T, V) λέγεται *εφαρμόσιμη* (*admissible*) για το $\nu \times \nu$ πολωνυμικό πίνακα $P(\lambda)$ βαθμού m εάν οι διαστάσεις των U , T και V είναι $\nu \times m\nu$, $m\nu \times m\nu$ και $m\nu \times \nu$, αντίστοιχα. Δύο εφαρμόσιμες τριάδες του $P(\lambda)$, (U_1, T_1, V_1) και (U_2, T_2, V_2) λέγονται *όμοιες* αν υπάρχει αντιστρέψιμος $m\nu \times m\nu$ πίνακας S τέτοιος ώστε

$$U_1 = U_2S, \quad T_1 = S^{-1}T_2S \quad \text{και} \quad V_1 = S^{-1}V_2. \quad (9.4.9)$$

Κάθε τριάδα όμοια με τη τριάδα (P_1, C_P, R_1) που ορίστηκε στο Θεώρημα 9.4.3 λέγεται *standard τριάδα* (*standard triple*) του πολωνυμικού πίνακα $P(\lambda)$. Για

$S = I_{m\nu}$, η τριάδα (P_1, C_P, R_1) είναι προφανώς μία standard τριάδα. Επιπλέον, από το Θεώρημα 9.4.2, έπεται ότι το δεύτερο μέλος μιας standard τριάδας του $P(\lambda)$ είναι πάντα ένας γραμμικοποιητής του $P(\lambda)$. Θα προχωρήσουμε τώρα στη γενίκευση του Θεωρήματος 9.4.3.

Θεώρημα 9.4.4. *Αν (U, T, V) είναι μία standard τριάδα του $P(\lambda)$, τότε*

$$P(\lambda)^{-1} = U(I_{m\nu}\lambda - T)^{-1}V, \quad \lambda \notin \sigma(P).$$

Απόδειξη. Αφού η τριάδα (U, T, V) είναι μία standard τριάδα του $P(\lambda)$, υπάρχει ένας $m\nu \times m\nu$ αντιστρέψιμος πίνακας S τέτοιος ώστε

$$P_1 = US, \quad C_L = S^{-1}TS \quad \text{και} \quad R_1 = S_{-1}V.$$

Παρατηρούμε ότι

$$(I_{m\nu}\lambda - C_P)^{-1} = (S^{-1}(I_{m\nu}\lambda - T)S)^{-1} = S^{-1}(I_{m\nu}\lambda - T)^{-1}S$$

και αντικαθιστώντας στην (9.4.6), παίρνουμε το επιθυμητό αποτέλεσμα. $\square$

Θεώρημα 9.4.5. *Έστω (U, T, V) μία εφαρμόσιμη τριάδα του $P(\lambda)$ η οποία ικανοποιεί τη σχέση*

$$P(\lambda)^{-1} = U(I_{m\nu}\lambda - T)^{-1}V. \quad (9.4.10)$$

Τότε, η (U, T, V) είναι μία standard τριάδα του $P(\lambda)$.

Από τα Θεωρήματα 9.4.4 και 9.4.5, και στοιχειώδεις ιδιότητες των ανάστροφων και αναστροφοσυζυγών πινάκων, προκύπτει άμεσα το ακόλουθο πόρισμα.

Πόρισμα 9.4.6. *Αν η τριάδα (U, T, V) μία standard τριάδα του πολυωνυμικού πίνακα $P(\lambda) = \sum_{j=0}^m A_j \lambda^j$, τότε η (V^T, T^T, U^T) είναι μία standard τριάδα του $P^T(\lambda) = \sum_{j=0}^m A_j^T \lambda^j$ και η (V^*, T^*, U^*) είναι μία standard τριάδα του $P^*(\lambda) = \sum_{j=0}^m A_j^* \lambda^j$.*

9.5 Αλυσίδες Jordan και Εφαρμογές

Θεωρούμε τον $\nu \times \nu$ πολυωνυμικό πίνακα $P(\lambda) = A_m \lambda^m + \dots + A_1 \lambda + A_0$ και με $P^{(r)}(\lambda)$ συμβολίζουμε την r -τάξεως παράγωγο του $P(\lambda)$ ως προς λ , η οποία είναι προφανώς ένας $\nu \times \nu$ πολυωνυμικός πίνακας βαθμού $m-r$. Μια πεπερασμένη ακολουθία από διανύσματα $x_0, x_1, \dots, x_k$, με $x_0 \neq 0$, είναι μία αλυσίδα Jordan μήκους $k+1$ του $P(\lambda)$ που αντιστοιχεί στην ιδιοτιμή λ_0 αν ισχύουν οι ακόλουθες $k+1$ σχέσεις:

$$\begin{aligned}
P(\lambda_0) x_0 &= 0, \\
P(\lambda_0) x_1 + \frac{1}{1!} P^{(1)}(\lambda_0) x_0 &= 0, \\
&\vdots \\
P(\lambda_0) x_k + \frac{1}{1!} P^{(1)}(\lambda_0) x_{k-1} + \cdots + \frac{1}{k!} P^{(k)}(\lambda_0) x_0 &= 0.
\end{aligned} \tag{9.5.1}$$

Παρατηρούμε ότι το x_0 , δηλαδή το πρώτο διάνυσμα της αλυσίδας, είναι ένα ιδιοδιάνυσμα του $P(\lambda)$ που αντιστοιχεί στην ιδιοτιμή λ_0 . Όσον αφορά τα υπόλοιπα διανύσματα της αλυσίδας Jordan, το καθένα ικανοποιεί μία μη ομογενή εξίσωση με τον μη αντιστρέψιμο πίνακα $P(\lambda_0)$ και καλείται γενικευμένο ιδιοδιάνυσμα του $P(\lambda)$ για την ιδιοτιμή λ_0 .

Οι αλυσίδες Jordan και οι τριάδες Jordan ενός (τετραγωνικού) πολυωνυμικού πίνακα ικανοποιούν πολλές και ενδιαφέρουσες ιδιότητες από τις οποίες οι πιο σημαντικές συνοψίζονται, χωρίς απόδείξεις, στο θεώρημα που ακολουθεί.

Θεώρημα 9.5.1. Έστω (X, J, Y) μία τριάδα Jordan του $n \times n$ πολυωνυμικού πίνακα $P(\lambda) = A_m \lambda^m + A_{m-1} \lambda^{m-1} + \cdots + A_1 \lambda + A_0$. Υποθέτουμε ότι οι πίνακες J και X έχουν διαμερίσεις της μορφής $J = \text{diag}\{J_1, J_2, \dots, J_s\}$ όπου κάθε J_i είναι ένα $\nu_i \times \nu_i$ Jordan block ($i = 1, 2, \dots, s$), και $X = [X_1 \ X_2 \ \cdots \ X_s]$ όπου $X_i = [x_{i,1} \ x_{i,2} \ \cdots \ x_{i,\nu_i}]$ είναι ένας $\nu \times \nu_i$ πίνακας ($i = 1, 2, \dots, s$). Τότε ισχύουν τα ακόλουθα:

- (a) Για κάθε $i = 1, 2, \dots, s$, η ακολουθία $x_{i,1}, x_{i,2}, \dots, x_{i,\nu_i}$ αποτελεί μία αλυσίδα Jordan του $P(\lambda)$ που αντιστοιχεί στην ιδιοτιμή του J_i .
- (b) Για κάθε $i = 1, 2, \dots, s$, $A_m X_i J_i^m + A_{m-1} X_i J_i^{m-1} + \cdots + A_1 X_i J_i + A_0 X_i = 0$.
- (c) $A_m X J^m + A_{m-1} X J^{m-1} + \cdots + A_1 X J + A_0 X = 0$.

$$\begin{aligned}
(d) \text{ Ο } m\nu \times m\nu \text{ πίνακας } Q = \begin{bmatrix} X \\ X J \\ \vdots \\ X J^{m-1} \end{bmatrix} \text{ είναι αντιστρέψιμος και ικανοποιεί} \\
\text{τις σχέσεις } C_P = Q J Q^{-1} \text{ και } Y = Q \begin{bmatrix} 0 \\ \vdots \\ 0 \\ A_m^{-1} \end{bmatrix}.
\end{aligned}$$

Αντίστροφα, αν για ένα εφαρμόσιμο ζεύγος (X, J) του $P(\lambda)$ (με τον $m\nu \times m\nu$ πίνακα J να είναι πίνακας Jordan) ισχύουν οι ιδιότητες (c) και (d), τότε αυτό είναι ένα ζεύγος Jordan του $P(\lambda)$.

Στο σημείο αυτό πρέπει να σημειώσουμε ότι γενικά, ένα εφαρμόσιμο ζεύγος (U, T) του πολυωνυμικού πίνακα $P(\lambda)$ είναι standard ζεύγος του $P(\lambda)$ αν και μόνο αν ικανοποιεί τις ιδιότητες (c) και (d) του παραπάνω θεωρήματος.

Το σύνολο των αλυσίδων Jordan που ταυτίζεται με τις $m\nu$ στήλες του X λέγεται *κανονικό σύνολο* (canonical set) των αλυσίδων Jordan του $P(\lambda)$.

Παράδειγμα 9.5.1. Έστω ο 2×2 μόνικος πολυωνυμικός πίνακας

$$P(\lambda) = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \lambda^2 + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \lambda + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} \lambda^2 & 0 \\ \lambda + 1 & \lambda(\lambda - 1) \end{bmatrix}.$$

Παρατηρούμε ότι $\det(P(\lambda)) = \lambda^3(\lambda - 1)$, δηλαδή ο $P(\lambda)$ έχει διακεκριμένες ιδιοτιμές $\lambda_1 = 0$ και $\lambda_2 = 1$. Αφού

$$P(0) = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} \quad \text{και} \quad P(1) = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix},$$

κάθε διάνυσμα της μορφής $[0 \ a]^T$ ($a \neq 0$) μπορεί να χρησιμοποιηθεί ως ιδιοδιάνυσμα για τον $P(\lambda)$ που αντιστοιχεί σε καθεμία από τις ιδιοτιμές.

Παρατηρούμε ότι η αλυσίδα Jordan του $P(\lambda)$ που αντιστοιχεί στη ιδιοτιμή $\lambda_2 = 1$ έχει μήκος 1. Πράγματι, αν $x_0 = [0 \ a]^T$ με $a \neq 0$, τότε η εξίσωση (ως προς x_1)

$$P(1)x_1 + P^{(1)}(1)x_0 = \begin{bmatrix} 1 & 0 \\ 2 & 0 \end{bmatrix} x_1 + \begin{bmatrix} 2 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

δεν έχει λύση. Η αντίστοιχη εξίσωση της ιδιοτιμής $\lambda_1 = 0$ είναι

$$P(0)x_1 + P^{(1)}(0)x_0 = \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x_1 + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} 0 \\ a \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

η οποία έχει τη γενική λύση $x_1 = [a \ \beta]^T$ με $\beta \in \mathbb{C}$ αυθαίρετο.

Ένα τρίτο μέλος της ίδιας αλυσίδας λαμβάνεται από την εξίσωση

$$\begin{aligned} & P(0)x_2 + P^{(1)}(0)x_1 + \frac{1}{2}P^{(2)}(0)x_0 \\ &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x_2 + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} a \\ \beta \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 \\ a \end{bmatrix}, \\ &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x_2 + \begin{bmatrix} 0 \\ 2a - \beta \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned}$$

η οποία δίνει $x_2 = [-2a + \beta \ \gamma]^T$ για αυθαίρετο γ . Για το τέταρτο μέλος της

αλυσίδας, θεωρούμε την εξίσωση

$$\begin{aligned} & P(0)x_3 + P^{(1)}(0)x_2 + \frac{1}{2}L^{(2)}(0)x_1 \\ &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x_3 + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} \begin{bmatrix} -2a + \beta \\ \gamma \end{bmatrix} + \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} a \\ \beta \end{bmatrix} \\ &= \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x_3 + \begin{bmatrix} a \\ -2a + 2\beta - \gamma \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \end{aligned}$$

η οποία δεν έχει λύση, διότι $a \neq 0$ για οποιαδήποτε τιμή του γ . Θέτοντας $a = 1$ και $\beta = \gamma = 0$, έχουμε την αλυσίδα Jordan

$$x_0 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad x_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad x_2 = \begin{bmatrix} -2 \\ 0 \end{bmatrix}$$

που αντιστοιχεί στην ιδιοτιμή $\lambda_1 = 0$. Είναι εύκολο να επαληθεύσουμε ότι οι πίνακες

$$X = \begin{bmatrix} 0 & 0 & 1 & -2 \\ 1 & 1 & 0 & 0 \end{bmatrix} \quad \text{και} \quad J = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

σχηματίζουν ένα ζεύγος Jordan του $P(\lambda)$.

Όπως έχουμε δει, ένας πολυωνυμικός πίνακας $P_1(\lambda)$ είναι ο δεξιός διαιρέτης ενός πολυωνυμικού πίνακα $P(\lambda)$, αν $P(\lambda) = P_2(\lambda)P_1(\lambda)$ για κάποιον άλλο πολυωνυμικό πίνακα $P_2(\lambda)$. Ο λόγος για τον οποίο απαιτούμε ο πολυωνυμικός πίνακας να έχει αντιστρέψιμο μεγιστοβάθμιο συντελεστή είναι για να αποφύγουμε τις πολλαπλότητες που θα προκύψουν διαφορετικά. Οι standard τριάδες, κι επομένως και οι φασματικές ιδιότητες (μέσω των τριάδων Jordan) των γινομένων των μόνικων πολυωνυμικών πινάκων, είναι απλές στην περιγραφή τους. Παρατηρούμε αρχικά ότι η σχέση μόνικων πολυωνυμικών πινάκων $P(\lambda) = P_2(\lambda)P_1(\lambda)$ συνεπάγεται τα εξής:

$$(a) \sigma(P_1) \cup \sigma(P_2) = \sigma(P),$$

$$(b) \text{ τα ιδιοδιανύσματα των } P_1(\lambda) \text{ και } P_2^T(\lambda) \text{ θα είναι και ιδιοδιανύσματα των } P(\lambda) \text{ και } P^T(\lambda), \text{ αντίστοιχα.}$$

Επιπλέον, έχουμε το ακόλουθο αποτέλεσμα.

Θεώρημα 9.5.2. Έστω $P(\lambda)$, $P_1(\lambda)$ και $P_2(\lambda)$ μόνικοι πολυωνυμικοί πίνακες τέτοιοι ώστε $P(\lambda) = P_2(\lambda)P_1(\lambda)$ και έστω (U_1, T_1, V_1) και (U_2, T_2, V_2) δύο standard τριάδες των $P_1(\lambda)$ και $P_2(\lambda)$, αντίστοιχα. Τότε η τριάδα

$$\begin{bmatrix} U_1 & 0 \end{bmatrix}, \quad \begin{bmatrix} T_1 & V_1 U_2 \\ 0 & T_2 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ V_2 \end{bmatrix} \quad (9.5.2)$$

είναι μία standard τριάδα του $P(\lambda)$.

Απόδειξη. Έστω m_1 και m_2 οι βαθμοί των $P_1(\lambda)$ και $P_2(\lambda)$, αντίστοιχα. Για τον $m\nu \times m\nu$ πίνακα $T = \begin{bmatrix} T_1 & V_1 U_2 \\ 0 & T_2 \end{bmatrix}$, είναι εύκολο να δει κανείς ότι

$$(I_{m\nu}\lambda - T)^{-1} = \begin{bmatrix} (I_{m_1\nu}\lambda - T_1)^{-1} & (I_{m_1\nu}\lambda - T_1)^{-1} V_1 U_2 (I_{m_2\nu}\lambda - T_2)^{-1} \\ 0 & (I_{m_2\nu}\lambda - T_2)^{-1} \end{bmatrix}.$$

Τότε οι ορισμοί των U και V δίνουν

$$\begin{aligned} U(I_{m\nu}\lambda - T)^{-1}V &= U_1(I_{m_1\nu}\lambda - T_1)^{-1}V_1 U_2(I_{m_2\nu}\lambda - T_2)^{-1}V_2 \\ &= P_1(\lambda)^{-1}P_2(\lambda)^{-1} = (P_2(\lambda)P_1(\lambda))^{-1} \\ &= P(\lambda)^{-1}, \end{aligned}$$

όπου έχει χρησιμοποιηθεί το Θεώρημα 9.4.4. Το συμπέρασμα προκύπτει από το Θεώρημα 9.4.5. $\square$

Για να διαπιστώσουμε τη σημασία του συμπεράσματος αυτού για τις φασματικές ιδιότητες των $P(\lambda)$, $P_1(\lambda)$ και $P_2(\lambda)$, αντικαθιστούμε τις standard τριάδες των $P_1(\lambda)$ και $P_2(\lambda)$ με τις τριάδες Jordan (X_1, J_1, Y_1) και (X_2, J_2, Y_2) , αντίστοιχα. Τότε ο $P(\lambda)$ έχει μία standard τριάδα (η οποία δεν είναι αναγκαία τριάδα Jordan),

$$\begin{bmatrix} X_1 & 0 \end{bmatrix}, \quad \begin{bmatrix} J_1 & Y_2 X_1 \\ 0 & J_2 \end{bmatrix}, \quad \begin{bmatrix} 0 \\ Y_2 \end{bmatrix}.$$

Αν και ο δεύτερος πίνακας (δηλαδή, ο γραμμικοποιητής) δεν είναι σε μορφή Jordan, μας δίνει σημαντική πληροφορία για το φάσμα του πολυωνυμικού πίνακα $P(\lambda)$ λόγω της τριγωνικής του μορφής.

Πόρισμα 9.5.3. Σύμφωνα με τις υποθέσεις του Θεωρήματος 9.5.2, $\sigma(P_1) \subseteq \sigma(P)$, και αν $\lambda_0 \in \sigma(L_1)$, τότε κάθε αλυσίδα Jordan του πολυωνυμικού πίνακα $P_1(\lambda)$ που αντιστοιχεί στο λ_0 είναι επίσης μία αλυσίδα Jordan του $P(\lambda)$ που αντιστοιχεί στο λ_0 . Ομοίως, $\sigma(P_2) \subseteq \sigma(P)$ και κάθε αλυσίδα Jordan του $P_2^T(\lambda)$ είναι μία αλυσίδα του $P^T(\lambda)$.

Στο τέλος του κεφαλαίου, παρουσιάζουμε, χωρίς αποδείξεις, δύο σημαντικές εφαρμογές των αλυσίδων Jordan στην επίλυση γραμμικών συστημάτων διαφορικών εξισώσεων και εξισώσεων διαφορών ανώτερης τάξης. Θεωρούμε το ομογενές σύστημα διαφορικών εξισώσεων

$$A_m x^{(m)}(t) + A_{m-1} x^{(m-1)}(t) + \cdots + A_1 x^{(1)}(t) + A_0 x(t) = 0 \quad (9.5.3)$$

ως προς την άγνωστη διανυσματική συνάρτηση $x(t) \in \mathbb{C}^\nu$ ($t \in \mathbb{R}$) και το αντίστοιχο ομογενές σύστημα εξισώσεων διαφορών

$$A_m x_{j+m} + A_{m-1} x_{j+m-1} + \cdots + A_1 x_{j+1} + A_0 x_j = 0, \quad j = 0, 1, \dots \quad (9.5.4)$$

ως προς την άγνωστη ακολουθία $x_0, x_1, x_2, \dots \in \mathbb{C}^\nu$.

Θεώρημα 9.5.4. Έστω $x_0, x_1, \dots, x_{k-1}$ μία αλυσίδα Jordan του πολυωνυμικού πίνακα $P(\lambda)$ για την ιδιοτιμή $\lambda_0 \in \sigma(P)$. Τότε οι k συναρτήσεις

$$\begin{aligned} u_0(t) &= x_0 e^{\lambda_0 t}, \\ u_1(t) &= (t x_0 + x_1) e^{\lambda_0 t}, \\ &\vdots \\ u_{k-1}(t) &= \left(\sum_{j=0}^{k-1} \frac{t^j}{j!} x_{k-1-j} \right) e^{\lambda_0 t} \end{aligned}$$

είναι γραμμικά ανεξάρτητες λύσεις της εξίσωσης (9.5.3).

Ισχύει επίσης ένα αντίστροφο αποτέλεσμα.

Θεώρημα 9.5.5. Έστω $x_0, x_1, \dots, x_{k-1} \in \mathbb{C}^n$ με $x_0 \neq 0$. Αν η διανυσματική συνάρτηση

$$u_{k-1}(t) = \left(\frac{t^{k-1}}{(k-1)!} x_0 + \dots + \frac{t}{1!} x_{k-2} + x_{k-1} \right) e^{\lambda_0 t}$$

είναι μία λύση της εξίσωσης (9.5.3), τότε ο αριθμός λ_0 είναι ιδιοτιμή του $P(\lambda)$ και η ακολουθία $x_0, x_1, \dots, x_{k-1}$ είναι μία αλυσίδα Jordan του $P(\lambda)$ που αντιστοιχεί στη λ_0 .

Παράδειγμα 9.5.2. Έστω το ομογενές σύστημα διαφορικών εξισώσεων

$$\left. \begin{aligned} x_1^{(2)}(t) &= 0 \\ x_2^{(2)}(t) + x_1^{(1)}(t) - x_2^{(1)}(t) + x_1(t) &= 0 \end{aligned} \right\}.$$

Γράφοντας το σύστημα αυτό στη μορφή (9.5.3) για $x = [x_1 \ x_2]^T$, έχουμε

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x^{(2)}(t) + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} x^{(1)}(t) + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x(t) = 0. \quad (9.5.5)$$

Επομένως ο αντίστοιχος πολυωνυμικός πίνακας είναι ο $P(\lambda)$ του Παραδείγματος 9.5.1, ο οποίος έχει μία απλή ιδιοτιμή $\lambda_1 = 1$ με ιδιοδιάνυσμα $[0 \ 1]^T$. Επομένως, μία λύση του συστήματος (9.5.5) είναι

$$v_0(t) = \begin{bmatrix} 0 \\ 1 \end{bmatrix} e^t.$$

Επιπλέον, ο $P(\lambda)$ έχει και την πολλαπλή ιδιοτιμή $\lambda_2 = 0$ με αλυσίδα Jordan $[0 \ 1]^T, [1 \ 0]^T, [-2 \ 0]^T$. Συνεπώς, τρεις γραμμικά ανεξάρτητες λύσεις του συστήματος (9.5.5) είναι

$$u_0(t) = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad u_1(t) = \begin{bmatrix} 1 \\ t \end{bmatrix} \quad \text{και} \quad u_2(t) = \begin{bmatrix} -2 + t \\ \frac{1}{2} t^2 \end{bmatrix}.$$

Θεώρημα 9.5.6. Αν $x_0, x_1, \dots, x_{k-1}$ είναι μία αλυσίδα Jordan του $P(\lambda)$ για την ιδιοτιμή $\lambda_0 \in \sigma(P)$, τότε οι k ακολουθίες $u_0^{(s)}, u_1^{(s)}, \dots \in \mathbb{C}^\nu$, $s = 0, 1, \dots, k-1$, με

$$\begin{aligned} u_j^{(0)} &= \lambda_0^j x_0, \\ u_j^{(1)} &= \begin{pmatrix} j \\ 1 \end{pmatrix} \lambda_0^{j-1} x_0 + \lambda_0^j x_1, \\ &\vdots \\ u_j^{(k-1)} &= \begin{pmatrix} j \\ k-1 \end{pmatrix} \lambda_0^{j-k+1} x_0 + \begin{pmatrix} j \\ k-2 \end{pmatrix} \lambda_0^{j-k+2} x_1 + \dots + \lambda_0^j x_{k-1}, \end{aligned}$$

για $j = 0, 1, \dots$, είναι γραμμικά ανεξάρτητες λύσεις της του συστήματος (9.5.4).

Θεώρημα 9.5.7. Έστω $x_0, x_1, \dots, x_{k-1} \in \mathbb{C}^\nu$, με $x_0 \neq 0$. Αν η ακολουθία $u_0^{(k-1)}, u_1^{(k-1)}, \dots \in \mathbb{C}^\nu$ με

$$u_j^{(k-1)} = \sum_{s=0}^{k-1} \begin{pmatrix} j \\ s \end{pmatrix} \lambda_0^{j-s} x_{k-1-s}$$

είναι μία λύση της εξίσωσης (9.5.4), τότε ο αριθμός λ_0 είναι μία ιδιοτιμή του $P(\lambda)$ και η ακολουθία $x_0, x_1, \dots, x_{k-1}$ είναι μία αλυσίδα Jordan του $P(\lambda)$ που αντιστοιχεί στη λ_0 .

Παράδειγμα 9.5.3. Θεωρούμε το ομογενές σύστημα εξισώσεων διαφορών

$$\left. \begin{aligned} x_{j+2}^{(2)} + x_{j+1}^{(1)} - x_{j+1}^{(2)} + x_j^{(1)} &= 0 \\ x_{j+2}^{(1)} &= 0 \end{aligned} \right\}, \quad j = 0, 1, \dots$$

Γράφουμε το συγκεκριμένο σύστημα στη μορφή

$$\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} x_{j+2} + \begin{bmatrix} 0 & 0 \\ 1 & -1 \end{bmatrix} x_{j+1} + \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix} x_j = 0, \quad j = 0, 1, \dots,$$

όπου $x_j = [x_j^{(1)} \ x_j^{(2)}]^T$ για κάθε j . Από το Θεώρημα 9.5.6, λαμβάνουμε τις ακόλουθες τέσσερις γραμμικά ανεξάρτητες λύσεις του συστήματος:

$$\begin{aligned} u_0^{(0)} &= u_1^{(0)} = u_2^{(0)} = \dots = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \\ v_0^{(0)} &= \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad v_1^{(0)} = v_2^{(0)} = \dots = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ v_0^{(1)} &= \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad v_1^{(1)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad v_2^{(1)} = v_3^{(1)} = \dots = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \\ v_0^{(2)} &= \begin{bmatrix} -2 \\ 0 \end{bmatrix}, \quad v_1^{(2)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad v_2^{(2)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad v_3^{(2)} = v_4^{(2)} = \dots = \begin{bmatrix} 0 \\ 0 \end{bmatrix}. \end{aligned}$$

9.6 Ασκήσεις

1. Να αποδείξετε την Πρόταση 9.3.1.
2. Να αποδείξετε την Πρόταση 9.3.2.
3. Να αποδείξετε το Πόρισμα 9.3.3.
4. Να αποδείξετε το Πόρισμα 9.3.4.
5. Να αποδείξετε το Πόρισμα 9.3.5.
6. Να αποδείξετε το Θεώρημα 9.4.2.
7. Να αποδείξετε ότι η σχέση ομοιότητας των εφαρμόσιμων τριάδων είναι μία σχέση ισοδυναμίας.
8. Αν (U, T) είναι ένα standard ζεύγος ενός τετραγωνικού πολυωνυμικού πίνακα $P(\lambda) = A_m \lambda^m + \cdots + A_1 \lambda + A_0$, $\det(A_m) \neq 0$, τότε να αποδείξετε ότι είναι επίσης standard ζεύγος για κάθε πολυωνυμικό πίνακα $B P(\lambda)$ με $\det(B) \neq 0$.
9. Κατασκευάστε ένα ζεύγος Jordan για τον πολυωνυμικό πίνακα

$$P(\lambda) = \begin{bmatrix} \lambda^2 - 2\lambda - 2 & \lambda + 2 \\ \lambda + 2 & \lambda^2 - 2\lambda - 2 \end{bmatrix}.$$

10. Κατασκευάστε ένα ζεύγος Jordan για τον πολυωνυμικό πίνακα

$$P(\lambda) = \begin{bmatrix} \lambda^2 & -\lambda \\ 0 & \lambda^2 \end{bmatrix}.$$

Βιβλιογραφία

- [1] F.L. Bauer and C.T. Fike, Norms and exclusion theorems, *Numerische Mathematik*, 2 (1960) 137–144.
- [2] E.F. Beckenbach and R. Bellman, *Inequalities*, Springer, New York, 1965.
- [3] A. Berman and R.J. Plemmons, *Nonnegative Matrices in Mathematical Sciences*, SIAM, Philadelphia, 1994.
- [4] R. Bhatia, *Matrix Analysis*, Springer, New York, 1997.
- [5] M. Eiermann, Fields of values and iterative methods, *Linear Algebra and its Applications*, 180 (1988) 167–196.
- [6] L. Elsner and Kh. Ikramov, Normal matrices: An update, *Linear Algebra and its Applications*, 285 (1998) 291–303.
- [7] L. Elsner and M.H. Paardekooper, On measures of nonnormality of matrices, *Linear Algebra and its Applications*, 92 (1987) 107–124.
- [8] M. Embree and L.N. Trefethen, Generalizing eigenvalue problems to pseudospectra problems, *SIAM Journal of Scientific Computing*, 23 (2001) 583–590.
- [9] W. Givens, Field of values of a matrix, *Proceedings of the American Mathematical Society*, 3 (1952) 206–209.
- [10] G.H. Golub and C.F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, 1996.
- [11] I. Gohberg, P. Lancaster and L. Rodman, *Matrix Polynomials*, Academic Press, New York, 1982.
- [12] M. Goldberg and E. Tadmor, On the numerical radius and its applications, *Linear Algebra and its Applications*, 42 (1982) 263–284.
- [13] R. Grone, C.R. Johnson, E. Sa and H. Wolkowicz, Normal matrices, *Linear Algebra and its Applications*, 87 (1987) 213–225.

- [14] K.E. Gustafson and D.K.M. Rao, *Numerical Range. The Field of Values of Linear Operators and Matrices*, Springer-Verlag, New York, 1997.
- [15] P. Henrici, Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices, *Numerische Mathematik*, 4 (1962) 24–40.
- [16] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.
- [17] R.A. Horn and C.R. Johnson, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.
- [18] A. Housholder, *The Theory of Matrices in Numerical Analysis*, Blaisdel, New York, 1964.
- [19] C.R. Johnson, Normality and the numerical range, *Linear Algebra and its Applications*, 15 (1976) 89–94.
- [20] C.R. Johnson, Numerical determination of the field of values of a general complex matrix, *SIAM Journal on Numerical Analysis*, 15 (1978) 595–602.
- [21] T. Kato, *Perturbation Theory of Linear Operators*, Springer-Verlag, Berlin, 1976.
- [22] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Academic Press, New York, 1985.
- [23] C.-K. Li, A simple proof of the elliptical range theorem, *Proceedings of the American Mathematical Society*, 124 (1996) 1985–1986.
- [24] M. Marcus and H. Minc, *A Survey of Matrix Theory and Matrix Inequalities*, Dover Publications, New York 1969.
- [25] A.S. Markus, *Introduction to the Spectral Theory of Polynomial Operator Pencils*, American Mathematical Society, Providence, Translations of Mathematical Monographs, Vol. 71, 1988.
- [26] A. Ruhe, On the closeness of eigenvalues and singular values for almost normal matrices, *Linear Algebra and its Applications*, 11 (1978) 87–94.
- [27] R.A. Smith, The condition numbers of the matrix eigenvalue problem, *Numerische Mathematik*, 10 (1967) 232–240.
- [28] G.W. Stewart and J.-G. Sun, *Matrix Perturbation Theory*, Academic Press, 1990.
- [29] L.N. Trefethen, Pseudospectra of linear operators, *SIAM Review*, 39 (1997) 383–406.

- [30] L.N. Trefethen and D. Bau, *Numerical Linear Algebra*, SIAM, 1997.
- [31] L.N. Trefethen and M. Embree, *Spectra and Pseudospectra: The Behavior of Nonnormal Matrices and Operators*, Princeton University Press, 2005.
- [32] R. Varga, *Geršgorin and His Circles*, Springer-Verlag, Heidelberg, 2004.
- [33] J.H. Wilkinson, *The Algebraic Eigenvalue Problem*, Clarendon Press, 1965.
- [34] Σ. Αργυρός, *Σημειώσεις Παραδόσεων Συναρτησιακής Ανάλυσης*, Ε.Μ. Πολυτεχνείο, Αθήνα, 2004.
- [35] Α. Γιαννόπουλος, *Σημειώσεις Συναρτησιακής Ανάλυσης*, Πανεπιστήμιο Κρήτης, 2003.
- [36] Ν. Καδιανάκης και Σ. Καρανάσιος, *Γραμμική Άλγεβρα – Αναλυτική Γεωμετρία και Εφαρμογές*, Αθήνα, 2001.
- [37] Ι. Μαρουλάς, *Γραμμική Άλγεβρα*, Αθήνα, 2000.
- [38] Ι. Μαρουλάς, *Σημειώσεις Ανάλυσης Πινάκων*, Αθήνα, 2005.
- [39] Ν. Παπαθανασίου, *Θεωρία Διαταραχής Ιδιοτιμών Πίνακα και Ψευδοφάσμα*, Μεταπτυχιακή Διπλωματική Εργασία, Τομέας Μαθηματικών Σ.Ε.Μ.Φ.Ε., Ε.Μ. Πολυτεχνείο, 2005.
- [40] Μ. Ροκίδης, *Κανονικότητα Πινάκων και Συγγενείς Αποστάσεις*, Μεταπτυχιακή Διπλωματική Εργασία, Τομέας Μαθηματικών Σ.Ε.Μ.Φ.Ε., Ε.Μ. Πολυτεχνείο, 2011.
- [41] Σ. Σαμαράς, *Πολυωνυμικοί Πίνακες και Εφαρμογές*, Μεταπτυχιακή Διπλωματική Εργασία, Τομέας Μαθηματικών Σ.Ε.Μ.Φ.Ε., Ε.Μ. Πολυτεχνείο, 2008.
- [42] Α. Φελλούρης, *Γραμμική Άλγεβρα και Αναλυτική Γεωμετρία*, Αθήνα, 2009.
- [43] Χ. Χωριανόπουλος, *Αριθμητικό Πεδίο και Αριθμητική Πολυωνυμική Θήκη Πίνακα (Θεωρία και Εφαρμογές)*, Μεταπτυχιακή Διπλωματική Εργασία, Τομέας Μαθηματικών Σ.Ε.Μ.Φ.Ε., Ε.Μ. Πολυτεχνείο, 2006.