



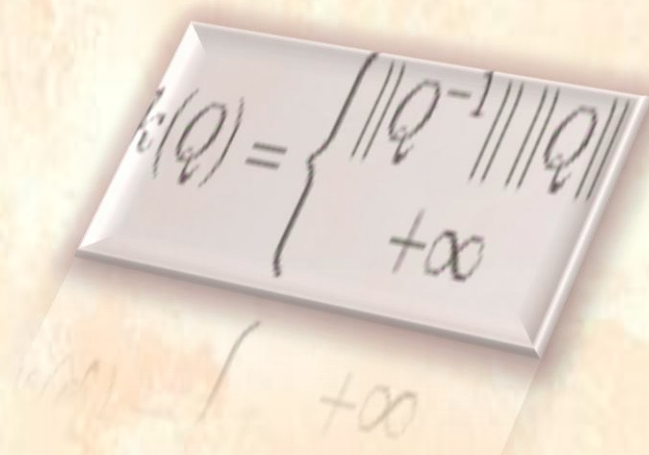
ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ
ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΔΕΙΚΤΕΣ ΚΑΤΑΣΤΑΣΗΣ ΙΔΙΟΤΙΜΩΝ
ΠΙΝΑΚΑ ΚΑΙ ΑΠΟΜΑΚΡΥΝΣΗ ΑΠΟ ΤΗΝ
ΚΑΝΟΝΙΚΟΤΗΤΑ

Διπλωματική Εργασία

ΓΚΡΟΠΑ ΕΥΑΓΓΕΛΙΑ

Αριθμός Μητρώου: 09110023



Επιβλέπων Καθηγητής:

Ψαρράκος Παναγιώτης,
Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2015



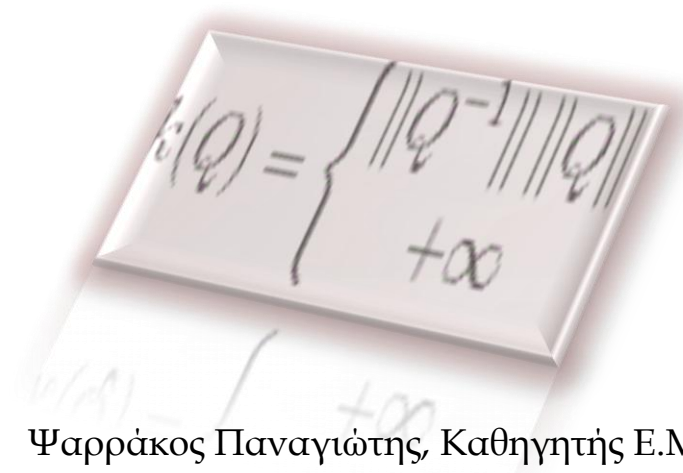
ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ
ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΔΕΙΚΤΕΣ ΚΑΤΑΣΤΑΣΗΣ ΙΔΙΟΤΙΜΩΝ
ΠΙΝΑΚΑ ΚΑΙ ΑΠΟΜΑΚΡΥΝΣΗ ΑΠΟ ΤΗΝ
ΚΑΝΟΝΙΚΟΤΗΤΑ

Διπλωματική Εργασία

ΓΚΡΟΠΑ ΕΥΑΓΓΕΛΙΑ

Αριθμός Μητρώου: 09110023



Τριμελής Επιτροπή: Ψαρράκος Παναγιώτης, Καθηγητής Ε.Μ.Π. (επιβλέπων)
Χρυσάφινος Κωνσταντίνος, Αναπλ. Καθηγητής Ε.Μ.Π.
Γιαννακάκης Νικόλαος, Επίκ. Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2015

ΠΕΡΙΕΧΟΜΕΝΑ

	Σελ.
Πρόλογος	4
Ευχαριστίες	6
Συμβολισμοί	7
1) Εισαγωγή	
1.1) Γενικοί Ορισμοί	9
1.2) Στοιχεία Ανάλυσης Πινάκων	12
1.3) Παραγοντοποίηση QR	16
1.4) Παραγοντοποίηση SVD	18
1.5) Ευστάθεια Τετραγωνικών Γραμμικών Συστημάτων	19
1.6) Κανονικοί Πίνακες	21
1.7) Αριθμητικό Πεδίο - Αριθμητική Ακτίνα	22
1.8) Ψευδοφάσμα Πίνακα και Διαταραχές Ιδιοτιμών	25
2) Οι Δείκτες Κατάστασης του Προβλήματος Ιδιοτιμών Πίνακα	
2.1) Οι Jordan Δείκτες Κατάστασης	32
2.2) Οι Δείκτες Κατάστασης Ιδιοτιμών Πίνακα	35
2.3) Οπτικοποίηση του σημαντικού ρόλου του Δείκτη Κατάστασης της Ιδιοτιμής για μικρές Διαταραχές	40
2.4) Άνω Φράγμα του Δείκτη Κατάστασης Ιδιοτιμών Πίνακα	54
2.5) Gram Ορίζουσα και Διακρίνουσα	56
2.6) Κάτω Φράγμα του Δείκτη κ_2	58
2.7) Παραδείγματα Υπολογισμού του Δείκτη Κατάστασης- Παρατηρήσεις σχετικά με τα Αποτελέσματα	60
3) Φράγματα για Πίνακες Υψωμένους σε Δύναμη και για Αντίστροφους Πίνακες, Φασματική Μεταβολή και Αριθμητικά Πεδία για Μη - Κανονικούς Πίνακες	
3.1) Ένα Μέτρο για τη Μη-Κανονικότητα κατά Henrici	63
3.2) Δυνάμεις Πινάκων	72
3.3) Προσεγγιστικές Λύσεις για Γραμμικά Συστήματα	74
3.4) Φασματική Μεταβολή και Μεταβολή Ιδιοτιμών	75
3.5) Αριθμητικό Πεδίο	82
4) Πίνακες με Κακής-Κατάστασης Πρόβλημα Ιδιοτιμών	
4.1) Παρατηρήσεις σε Πίνακες με Ιδιοπρόβλημα Κακής-Κατάστασης	89
4.2) Ιδιότητες ενός Πίνακα με ένα Κακής-Κατάστασης Ιδιοπρόβλημα	93
Βιβλιογραφία	98

ΠΡΟΛΟΓΟΣ

Η συγγραφή της παρούσας διπλωματικής εργασίας με τίτλο *Δείκτες Κατάστασης Ιδιοτιμών Πίνακα και Απομάκρυνση από την Κανονικότητα* πραγματοποιήθηκε στο πλαίσιο του Προγράμματος Σπουδών της Σχολής Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών του Εθνικού Μετσόβιου Πολυτεχνείου. Αποτελεί μια πρώτη προσπάθεια δημιουργίας γνωστικού υποβάθρου αλλά και κινήτρου για τη μελέτη του δείκτη κατάστασης και του κατά πόσο η απομάκρυνση από την κανονικότητα επηρεάζει τη συμπεριφορά ενός (τυχαίου) τετραγωνικού πίνακα.

Η επιλογή του συγκεκριμένου θέματος έγινε με κριτήριο το προσωπικό ενδιαφέρον για τη μελέτη του δείκτη κατάστασης και της συμπεριφοράς του καθώς και της απομάκρυνσης από την κανονικότητα. Παραθέτουμε τους κατάλληλους ορισμούς και κάποια θεωρήματα με τις αποδείξεις τους, που αποτελούν χρήσιμα εργαλεία για την κατανόηση του αντικειμένου. Ο δείκτης κατάστασης αποτελεί χρήσιμο εργαλείο για τη διερεύνηση της ευστάθειας (ή αστάθειας) ενός πίνακα. Επιπλέον καθορίζει και το πώς οι διάφορες διαταραχές του πίνακα επηρεάζουν τη λύση του αντίστοιχου ιδιοπροβλήματος. Στη συνέχεια εστιάζουμε την προσοχή μας σε μια ειδική μορφή πινάκων, τους κανονικούς πίνακες. Με τους συγκεκριμένους πίνακες ασχολήθηκαν κατά καιρούς γνωστοί μαθηματικοί όπως ο P. Henrici, ο A. Ruhe, ο H. Wielandt και ο G.E. Forsythe, η επιστημονική έρευνα των οποίων αποτέλεσε το κίνητρο για τη σύνταξη της εργασίας αυτής. Επίσης ορίζουμε το ψευδοφάσμα και παρατηρούμε τη συμπεριφορά του με την παρουσία τυχαίων διαταραχών. Επιλέγουμε διάφορους τετραγωνικούς πίνακες και σχεδιάζουμε το ψευδοφάσμα τους. Τέλος, δίνουμε ιδιαίτερη προσοχή σε πίνακες που έχουν ιδιοπρόβλημα κακής - κατάστασης.

Το Κεφάλαιο 1 ουσιαστικά έχει κυρίως εισαγωγικό και βοηθητικό χαρακτήρα. Παρουσιάζουμε βασικούς ορισμούς της Γραμμικής Άλγεβρας και της Ανάλυσης Πινάκων προκειμένου ο αναγνώστης να κατανοήσει τα κεφάλαια που ακολουθούν. Γίνεται ιδιαίτερη αναφορά στις διανυσματικές νόρμες και στις νόρμες πινάκων που θα χρησιμοποιηθούν ευρύτατα στα επόμενα κεφάλαια. Ακολουθεί η παραγοντοποίηση QR και η παραγοντοποίηση ιδιαζουσών τιμών. Στη συνέχεια παραθέτουμε τον ορισμό του δείκτη κατάστασης ενός πίνακα και γίνεται αναφορά στους κανονικούς πίνακες και στις διάφορες ιδιότητές τους. Υπενθυμίζονται, επιπλέον, οι ορισμοί του αριθμητικού πεδίου και της αριθμητικής ακτίνας καθώς και το ελλειπτικό θεώρημα. Στο τέλος του κεφαλαίου παρουσιάζουμε τον ορισμό του ψευδοφάσματος και αναφέρουμε κάποια χρήσιμα θεωρήματα, απαραίτητα για τη μελέτη της συμπεριφοράς του ψευδοφάσματος με την ύπαρξη διαταραχής.

Στο Κεφάλαιο 2, εξετάζονται εκτενέστερα ο δείκτης κατάστασης πίνακα και ο δείκτης κατάστασης απλής ιδιοτιμής και διάφορα θεωρήματα με τις αποδείξεις τους όπου χρησιμοποιείται ο δείκτης κατάστασης. Ακόμη παρουσιάζουμε παραδείγματα προκειμένου να παρατηρήσουμε τη συμπεριφορά του. Στο σημείο αυτό αξίζει να

σημειωθεί ότι η μελέτη των δεικτών κατάστασης γίνεται αποκλειστικά με τη χρήση της 2-νόρμας και της νόρμας Frobenius (συμβ. $\| \cdot \|_2$ και $\| \cdot \|_F$, αντίστοιχα). Το κεφάλαιο αυτό στηρίζεται στη δημοσίευση του R.A. Smith με τίτλο *The Condition Numbers of the Matrix Eigenvalue Problem* στο περιοδικό *Numerische Mathematik*, 10, 232-240 (1967). Παραθέτουμε παραδείγματα που μας βοηθούν στο να συνοψίσουμε τα συμπεράσματα μας για το ψευδοφάσμα. Επίσης παρουσιάζουμε γραφήματα του ψευδοφάσματος τα οποία μας βοηθούν να επιβεβαιώσουμε τη σημασία του δείκτη κατάστασης.

Στο Κεφάλαιο 3 επικεντρωνόμαστε στην απομάκρυνση από την κανονικότητα κατά Henrici. Αρχικά γίνεται μια αναφορά στο Θεώρημα του Schur που είναι ένα εύχρηστο εργαλείο για τη τριγωνοποίηση πινάκων και θα μας χρειαστεί σε αποδείξεις θεωρημάτων. Στη συνέχεια δίνουμε ιδιαίτερη προσοχή σε δυνάμεις πινάκων μαζί με βασικές ιδιότητές τους και περιγράφουμε τις προσεγγιστικές λύσεις γραμμικών συστημάτων. Έπειτα αναφερόμαστε στη φασματική μεταβολή και στη μεταβολή ιδιοτιμών. Πρόκειται για δύο έννοιες που συσχετίζουν τις ιδιοτιμές δύο διαφορετικών πινάκων. Εν συνεχεία, δίνονται φράγματα για τη φασματική μεταβολή και για τη μεταβολή ιδιοτιμών καθώς επίσης και παραδείγματα για την καλύτερη κατανόησή τους. Στο τελευταίο μέρος του κεφαλαίου μελετάμε την απόσταση του αριθμητικού πεδίου ενός τετραγωνικού πίνακα από την κυρτή θήκη των ιδιοτιμών του και την απομάκρυνση του πίνακα από την κανονικότητα, με τη χρήση της νόρμας Frobenius. Παραθέτουμε τρία σημαντικά θεωρήματα που μας βοηθούν να κατανοήσουμε ότι η απόσταση του αριθμητικού πεδίου από την κυρτή θήκη του φάσματος (η οποία περιέχεται πάντα στο αριθμητικό πεδίο) εξαρτάται από την απομάκρυνση από την κανονικότητα και ότι το αριθμητικό πεδίο περιέχεται στην ένωση συγκεκριμένων ελλειπτικών δίσκων με εστίες τις ιδιοτιμές. Το κεφάλαιο αυτό βασίζεται στη δημοσίευση του P. Henrici με τίτλο *Bounds for iterates, inverses, spectral variation and fields of values of non - normal matrices* στο περιοδικό *Numerische Mathematik*, 4, 24-40 (1962).

Το Κεφάλαιο 4 είναι βασισμένο στη δημοσίευση του J.H. Wilkinson με τίτλο *Note on Matrices with a Very Ill - Conditioned Eigenproblem* στο περιοδικό *Numer. Math.*, 19, 176-178 (1972) και στη δημοσίευση του A. Ruhe με τίτλο *Properties of a Matrix with a Very Ill - conditioned Eigenproblem* στο περιοδικό *Numer. Math.*, 15, 57-60 (1970). Το συγκεκριμένο κεφάλαιο παρουσιάζει ιδιαίτερο ενδιαφέρον, γιατί αναφέρεται σε πίνακες με ιδιοπρόβλημα κακής - κατάστασης. Επικεντρωνόμαστε στις ιδιότητές τους και στη συνέχεια διατυπώνουμε ένα πολύ σημαντικό θεώρημα που αναφέρεται στην απόσταση μεταξύ ενός δεδομένου πίνακα και ενός συνόλου \mathcal{M} των πινάκων που έχουν τουλάχιστον μια πολλαπλή ιδιοτιμή.

ΕΥΧΑΡΙΣΤΙΕΣ

Αναμφίβολα η εκπόνηση της διπλωματικής εργασίας διευρύνει τους πνευματικούς ορίζοντες του φοιτητή. Αυτό συμβαίνει, διότι ο φοιτητής καλείται να εμβαθύνει και να ξεφύγει από τα στενά πλαίσια της μελέτης ενός συγκεκριμένου αντικειμένου. Για να ξεπεράσει τις δυσκολίες, που προκύπτουν κατά την εκπόνηση της διπλωματικής εργασίας, αρωγοί στην προσπάθειά του είναι οι άνθρωποι που τον στηρίζουν.

Θα ήθελα, λοιπόν, να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή της εργασίας μου, κ. Παναγιώτη Ψαρράκο, Καθηγητή στον Τομέα Μαθηματικών της Σχολής Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών του Εθνικού Μετσόβιου Πολυτεχνείου, ο οποίος υπήρξε άριστος σύμβουλος με τις πολύτιμες παρατηρήσεις του και συμβουλές του και ουσιαστικός αρωγός στην προσπάθειά μου καθ' όλη τη διάρκεια της εκπόνησης της διπλωματικής μου εργασίας. Τον ευχαριστώ θερμά τόσο για την υπόδειξη του θέματος όσο και για την απεριόριστη βοήθεια που μου παρείχε. Η πολύτιμη καθοδήγησή του, η έμπρακτη στήριξή του και η άψογη συνεργασία μας συνετέλεσαν στην ολοκλήρωση αυτής της δημιουργικής διαδικασίας.

Είμαι ευγνώμων και θα ήθελα να εκφράσω τις ευχαριστίες μου προς την τριμελή συμβουλευτική επιτροπή μου. Συγκεκριμένα θα ήθελα να ευχαριστήσω τον κ. Κωνσταντίνο Χρυσάφινο (Αναπληρωτή Καθηγητή του Τομέα Μαθηματικών της Σχολής Ε.Μ.Φ.Ε.) καθώς και τον κ. Νικόλαο Γιαννακάκη (Επίκουρο Καθηγητή της ίδιας Σχολής) για την ενεργή συμμετοχή τους στην τριμελή συμβουλευτική επιτροπή.

Ευχαριστώ θερμά την οικογένειά μου που με στήριξαν σε όλη τη διάρκεια μελέτης και ανάπτυξης της εργασίας μου, παρέχοντάς μου στοργή, σθένος και διαρκή υποστήριξη. Ευχαριστώ τους γονείς μου, Ευάγγελο και Ελένη, οι οποίοι μου έδωσαν όλα τα απαραίτητα εφόδια στη ζωή, ηθικά και υλικά. Η παρούσα εργασία είναι αφιερωμένη σε αυτούς τους δύο τόσο σημαντικούς στη ζωή μου ανθρώπους.

Τέλος, θα ήθελα να ευχαριστήσω την Ελένη, το Σπύρο, το Χρήστο και την Κατερίνα για όσα μου χάρισαν αυτά τα πέντε χρόνια των σπουδών μου και για την πίστη τους στις δυνατότητές μου.

Συμβολισμοί

\mathbb{R}	Σύνολο πραγματικών αριθμών
\mathbb{C}	Σύνολο μιγαδικών αριθμών
\mathbb{R}^n	Σύνολο πραγματικών διανυσμάτων n διαστάσεων
\mathbb{C}^n	Σύνολο μιγαδικών διανυσμάτων n διαστάσεων
$\mathbb{R}^{m \times n}$	Σύνολο $m \times n$ πραγματικών πινάκων
$\mathbb{C}^{m \times n}$	Σύνολο $m \times n$ μιγαδικών πινάκων
$A = [a_{ij}], B = [b_{ij}]$ κ.ά.	Πίνακες μιγαδικοί ή πραγματικοί
O, O_n	Μηδενικός $n \times n$ πίνακας
I, I_n	Μοναδιαίος (Ταυτοτικός) $n \times n$ πίνακας
A^{-1}	Αντίστροφος του πίνακα A
\bar{A}	Συζυγής του πίνακα A
A^T	Ανάστροφος του πίνακα A
A^*, \bar{A}^T	Αναστροφosuζυγής του πίνακα A
x^T	Ανάστροφος του διανύσματος x
x^*	Αναστροφosuζυγής του διανύσματος x
$H(A)$	Το ερμιτιανό μέρος ενός τετραγωνικού πίνακα A
$S(A)$	Το αντιερμιτιανό μέρος ενός τετραγωνικού πίνακα A
$\text{Adj}A$	Ο συμπληρωματικός (προσαρτημένος) του πίνακα A
A^r	Δύναμη πίνακα
$\text{tr}A$ ή $\text{trace}A$	Ίχνος ενός τετραγωνικού πίνακα A
$\det(A)$	Ορίζουσα ενός τετραγωνικού πίνακα A
$\text{rank}(A)$	Ο βαθμός ενός πίνακα A
$\text{Re}(\cdot)$	Πραγματικό μέρος
$\text{Im}(\cdot)$	Φανταστικό μέρος
$\sigma(A)$	Το φάσμα ενός τετραγωνικού πίνακα A
$\rho(A)$	Η φασματική ακτίνα ενός τετραγωνικού πίνακα A
$F(A)$	Αριθμητικό πεδίο πίνακα A
$r(A)$	Αριθμητική ακτίνα πίνακα A
$\text{span}\{x_1, x_2, \dots, x_k\}$	Η γραμμική θήκη των διανυσμάτων x_1, x_2, \dots, x_k
$\text{Co}(\Omega)$	Κυρτή θήκη ενός συνόλου Ω
$\ \cdot\ $	Διανυσματική νόρμα
$\ \cdot\ _p$	Η p -νόρμα διανυσμάτων ή l_p διανυσματική νόρμα
$\ \cdot\ _2$	Η 2-νόρμα διανυσμάτων ή l_2 διανυσματική νόρμα ή Ευκλείδεια νόρμα
$\ \cdot\ _\infty$	Η ∞ -νόρμα διανυσμάτων ή max-νόρμα ή μέγιστη νόρμα
$\ \cdot\ $	Νόρμα πινάκων
$\ \cdot\ _F$	Η νόρμα Frobenius
$\ \cdot\ _{l_1}$	Η l_1 - νόρμα πίνακα
$\ \cdot\ _1$	Νόρμα μεγίστου αθροίσματος στηλών πίνακα
$\ \cdot\ _2$	Η φασματική (τελεστική) νόρμα πινάκων
$\ \cdot\ _\infty$	Νόρμα μεγίστου αθροίσματος γραμμών πίνακα

$\binom{n}{k}$	Ανάπτυγμα $\frac{n!}{k!(n-k)!}$
$\langle u, v \rangle$ ή (u, v)	Εσωτερικό γινόμενο διανυσμάτων u, v
$\kappa(A) = \kappa_{\ \cdot\ }(A) = \kappa_\alpha(A)$	Δείκτης κατάστασης του πίνακα A τυχαίας νόρμας $\ \cdot\ _\alpha$
$\kappa_2(A)$	Δείκτης κατάστασης του πίνακα A Ευκλείδειας νόρμας
$\kappa_F(A)$	Δείκτης κατάστασης του πίνακα A Frobenious νόρμας
$\kappa_A(\lambda) = \frac{\ x\ _2 \ y\ _2}{ y^*x }$	Δείκτης κατάστασης απλής ιδιοτιμής λ του πίνακα A ως προς τη 2-νόρμα (όπου x και y το δεξιό και αριστερό ιδιοδιάνυσμα, αντίστοιχα)
s_j	Ο αριθμός Wilkinson, ώστε $\kappa_A(\lambda_j) = s_j ^{-1}$ για $A \in \mathbb{C}^{n \times n}$, $\lambda_j \in \sigma(A)$ και $j = 1, 2, \dots$ (συνημίτονο γωνίας μεταξύ αριστερού και δεξιού ιδιόχωρου)
$\underline{\sigma}(A), \sigma_{\min}(A)$	Η μικρότερη ιδιάζουσα τιμή του A
$\Lambda_\varepsilon(A)$	Το ε -ψευδοφάσμα ενός πίνακα $A \in \mathbb{C}^{n \times n}$ για δεδομένο $\varepsilon > 0$
$h(S, T)$	Η Hausdorff απόσταση μεταξύ ενός συνόλου T και ενός άλλου συνόλου S
$\delta_{\ \cdot\ }(A)$	Απομάκρυνση κατά Henrici του A από την κανονικότητα ως προς μια νόρμα πινάκων $\ \cdot\ $
$s = s_A(B) = \max_{1 \leq i \leq n} \{ \min_{1 \leq j \leq n} \mu_i - \lambda_j \}$	Φασματική μεταβολή του $B = (b_{ij})$ με ιδιοτιμές μ_i σε σχέση με τον $A = (a_{ij})$ με ιδιοτιμές λ_i
$v = v(A, B) = \min_p \{ \max_{1 \leq i \leq n} \lambda_i - \mu_{p(i)} \}$	Μεταβολή ιδιοτιμών του $B = (b_{ij})$ με ιδιοτιμές μ_i σε σχέση με τον $A = (a_{ij})$ με ιδιοτιμές λ_i , όπου p : μεταθέσεις
$E(\lambda_i, \lambda_j, \frac{1}{2}a)$	Έλλειψη με εστίες λ_i, λ_j και μικρό άξονα a

ΚΕΦΑΛΑΙΟ 1: Εισαγωγή

1.1 Γενικοί Ορισμοί

Ο σκοπός του κεφαλαίου αυτού είναι να ορίσουμε και να υπενθυμίσουμε κάποιες εισαγωγικές έννοιες που αφορούν τη Γραμμική Άλγεβρα και την Ανάλυση Πινάκων και θα μας χρειαστούν παρακάτω.

Πριν ξεκινήσουμε αξίζει να σημειώσουμε ότι θα συμβολίζουμε το σύνολο των μιγαδικών $m \times n$ πινάκων με το σύνολο $\mathbb{C}^{m \times n}$ και το σύνολο των πραγματικών $m \times n$ πινάκων θα συμβολίζεται με $\mathbb{R}^{m \times n}$.

Ορισμός 1. Ένας πίνακας $A \in \mathbb{C}^{m \times n}$ καλείται *τετραγωνικός*, αν ο αριθμός των στηλών του πίνακα είναι ο ίδιος με τον αριθμό των γραμμών του, δηλαδή ισχύει $m = n$.

Ορισμός 2. Ο πίνακας $I_n = (\delta_{ij})_{n \times n} = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & 1 \end{bmatrix}$, όπου $\delta_{ij} = \begin{cases} 1, & \text{αν } i = j \\ 0, & \text{αν } i \neq j \end{cases}$ είναι η

συνάρτηση δέλτα του Kronecker, ονομάζεται *μοναδιαίος πίνακας* και έχει τις ιδιότητες:

$$AI_n = I_n A = A, \text{ για κάθε πίνακα } A \text{ τύπου } n \times n,$$

$$AI_n = A, \text{ για κάθε πίνακα } A \text{ τύπου } m \times n,$$

$$I_n A = A, \text{ για κάθε πίνακα } A \text{ τύπου } n \times l.$$

Ορισμός 3. Ένας τετραγωνικός πίνακας $A = (a_{ij})$ θα καλείται *κάτω τριγωνικός*, αν όλα τα στοιχεία που βρίσκονται πάνω από την κύρια διαγώνιο είναι μηδέν, δηλαδή αν $a_{ij} = 0$ για $i < j$.

Αντίστοιχα, ένας τετραγωνικός πίνακας $A = (a_{ij})$ θα καλείται *άνω τριγωνικός*, αν όλα τα στοιχεία που βρίσκονται κάτω από την κύρια διαγώνιο είναι μηδέν, δηλαδή αν $a_{ij} = 0$ για $i > j$.

Ορισμός 4. Ένας τετραγωνικός πίνακας $A = (a_{ij})$ θα καλείται *διαγώνιος*, αν τα μόνα μη-μηδενικά στοιχεία του, είναι αυτά που βρίσκονται στην κύρια διαγώνιο, δηλαδή αν $a_{ij} = 0$ για $i \neq j$.

Ορισμός 5. *Τρχος* ενός πίνακα $A = (a_{ij})$, που συμβολίζουμε με $\text{tr}(A)$ ή $\text{trace}(A)$, ονομάζουμε το άθροισμα των διαγώνιων στοιχείων του πίνακα, δηλαδή $\text{tr}(A) = \sum_i a_{ii}$.

Ορισμός 6. Αν ο πίνακας A είναι τετραγωνικός τύπου $n \times n$ και υπάρχει πίνακας X τέτοιος ώστε $AX = XA = I_n$, τότε λέμε ότι ο πίνακας A είναι αντιστρέψιμος και ο πίνακας X είναι ο αντίστροφος πίνακας του A . Γράφουμε τότε $X = A^{-1}$.

Σημείωση: Αν A, B δύο $n \times n$ αντιστρέψιμοι πίνακες, τότε ισχύει ότι $(AB)^{-1} = B^{-1}A^{-1}$.

Ορισμός 7. Έστω A, B δύο $n \times n$ πίνακες. Τότε οι A, B λέγονται όμοιοι, αν υπάρχει αντιστρέψιμος πίνακας P , τέτοιος ώστε $B = P^{-1}AP$.

Ορισμός 8. Αν $A = (a_{ij}) \in \mathbb{C}^{m \times n}$, τότε ο πίνακας που προκύπτει από τον A με εναλλαγή μεταξύ γραμμών και στηλών του λέγεται ανάστροφος πίνακας του A και συμβολίζεται με A^T , έχουμε δηλαδή $A^T = (a_{ji}) \in \mathbb{C}^{n \times m}$.

Σημείωση: Αν A, B δύο $n \times n$ πίνακες, τότε ισχύει ότι $(AB)^T = B^T A^T$.

Ορισμός 9. Αν $A = (a_{ij}) \in \mathbb{C}^{m \times n}$, τότε ο πίνακας με τα συζυγή στοιχεία του λέγεται συζυγής πίνακας $\bar{A} = [\bar{a}_{ij}]$. Ο ανάστροφος συζυγής του πίνακα A συμβολίζεται με $A^* = \bar{A}^T$. Εάν ισχύει $A = A^*$, τότε ο πίνακας ονομάζεται ερμιτιανός. Ακόμη εάν $A^* = -A$, τότε ο πίνακας ονομάζεται αντιερμιτιανός.

Σημείωση: Αν A, B δύο $n \times n$ πίνακες, τότε ισχύει ότι $(AB)^* = B^* A^*$.

Ορισμός 10. Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ λέγεται ορθομοναδιαίος αν ισχύει

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

Αν ο πίνακας $A \in \mathbb{R}^{n \times n}$, τότε ο πίνακας A λέγεται ορθογώνιος αν ισχύει

$$A^T A = A A^T = I \Leftrightarrow A^T = A^{-1}.$$

Ορισμός 11. Ένας τετραγωνικός πίνακας $A \in \mathbb{C}^{n \times n}$ καλείται κανονικός όταν αντιμετατίθεται με τον αναστροφοσυζυγή του, δηλαδή όταν ικανοποιεί τη σχέση

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

Ορισμός 12. Ένας πίνακας A λέγεται συμμετρικός, αν ο πίνακας είναι ίσος με τον ανάστροφό του, δηλαδή ισχύει $A = A^T$ και αντισυμμετρικός, αν ο ανάστροφος του πίνακα είναι ο αντίθετος πίνακας. Τα διαγώνια στοιχεία αντισυμμετρικών πινάκων είναι μηδενικά.

Σημείωση: Κάθε πίνακας μπορεί να γραφτεί ως άθροισμα ενός συμμετρικού και ενός αντισυμμετρικού πίνακα.

Ορισμός 13. Έστω V διανυσματικός χώρος πάνω στο σώμα $K = \mathbb{R}$ ή \mathbb{C} και $f: V \times V \rightarrow K$ μία συμμετρική διγραμμική μορφή $f(x, y) = x^T A y$ επί του V . Τότε η απεικόνιση

$$Q: V \rightarrow V, x \rightarrow Q(x) := f(x, x) = x^T A x$$

λέγεται τετραγωνική μορφή επί του V αντίστοιχη της f .

Ορισμός 14. Η τετραγωνική μορφή $x^T Ax$ θα ονομάζεται *θετικά ορισμένη*, εάν $x^T Ax > 0$ για κάθε $x \neq 0$, και ο συμμετρικός πίνακας A θα ονομάζεται *θετικά ορισμένος* πίνακας, εάν $x^T Ax$ είναι θετικά ορισμένη τετραγωνική μορφή.

Σημείωση: Ένας συμμετρικός πίνακας A θα είναι θετικά ορισμένος, εάν όλες οι ιδιοτιμές του είναι θετικές.

Ορισμός 15. Έστω A ένας $n \times n$ πίνακας. Ονομάζουμε *υποπίνακα* του A , τον πίνακα που προκύπτει από τον A , αν διαγράψουμε έναν αριθμό στηλών ή έναν αριθμό γραμμών ή έναν αριθμό γραμμών και στηλών.

Ορισμός 16. Έστω A είναι ένας $n \times n$ (τετραγωνικός) πίνακας. Ένα μη-μηδενικό διάνυσμα $x \in \mathbb{C}^n$ ονομάζεται *ιδιοδιάνυσμα* του A , αν

$$Ax = \lambda x$$

για κάποιο $\lambda \in \mathbb{C}$. Το $\lambda \in \mathbb{C}$ ονομάζεται *ιδιοτιμή* του A και λέμε ότι το x είναι ένα ιδιοδιάνυσμα του A που αντιστοιχεί στη λ .

Η λέξη “ιδιοδιάνυσμα” σχεδόν πάντα εκφράζει ένα *δεξί ιδιοδιάνυσμα*, δηλαδή ένα διάνυσμα στήλη που πρέπει να τοποθετηθεί δεξιά από τον πίνακα A στη χαρακτηριστική εξίσωση

$$Ax = \lambda x.$$

Μπορεί επίσης να υπάρχει διάνυσμα στήλη, το οποίο αν αναστραφεί, δεν μεταβάλλεται, όταν τοποθετηθεί στα αριστερά ενός γινομένου με έναν τετραγωνικό πίνακα A . Δηλαδή ικανοποιεί την εξίσωση

$$y^* A = \lambda y^*.$$

Κάθε διάνυσμα y με αυτήν την ιδιότητα ονομάζεται *αριστερό ιδιοδιάνυσμα* του A .

Ορισμός 17. Μια ιδιοτιμή ενός πίνακα A είναι *απλή* (*simple*), αν αυτή έχει αλγεβρική πολλαπλότητα 1 και *πολλαπλή* (*multile*) διαφορετικά. Μια ιδιοτιμή είναι *μη-μειωτική* (*nonderogatory*), εάν η γεωμετρική πολλαπλότητα είναι 1 και *μειωτική* (*derogatory*) διαφορετικά. Μια ιδιοτιμή είναι *ημισπλή* (*semisimple*), εάν οι γεωμετρικές και οι αλγεβρικές πολλαπλότητες συμπίπτουν.

Σ' αυτό το σημείο που ορίσαμε και την έννοια του ιδιοδιανύσματος αξίζει να αναφερθούμε στον τρόπο κατασκευής της *κανονικής μορφής Jordan* ενός τετραγωνικού πίνακα.

Ορισμός 18. Έστω $A \in \mathbb{C}^{n \times n}$ και λ_0 μία ιδιοτιμή του. Ένα μη μηδενικό διάνυσμα $x \in \mathbb{C}^n$ 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 $(A - \lambda_0 I_n)^k = 0$, ενώ $(A - \lambda_0 I_n)^{k-1} \neq 0$.

Από ένα γενικευμένο ιδιοδιάνυσμα x τάξης k , δημιουργείται μία αλυσίδα διανυσμάτων

$$x_k = x, x_{k-1} = (A - \lambda_0 I_n) x_k, x_{k-2} = (A - \lambda_0 I_n)^2 x_k, \dots, x_1 = (A - \lambda_0 I_n)^{k-1} x_k.$$

Έτσι κατασκευάσαμε 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

$$Ax_r = x_{r-1} + \lambda_0 x_r, r = 1, 2, \dots, k.$$

Επομένως,

$$\begin{aligned} A[x_1 \ x_2 \ \dots \ x_{k-1} \ x_k] &= [Ax_1 \ Ax_2 \ \dots \ Ax_{k-1} \ Ax_k] = [\lambda_0 x_1 \ \lambda_0 x_2 + x_1 \ \dots \ \lambda_0 x_k + x_{k-1}] \\ &= [x_1 \ x_2 \ \dots \ x_{k-1} \ x_k] \begin{bmatrix} \lambda_0 & 1 & \dots & 0 \\ 0 & \lambda_0 & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \dots & 0 & \lambda_0 \end{bmatrix}, \end{aligned}$$

όπου ο υποπίνακας $J_k \in \mathbb{C}^{k \times k}$ είναι το Jordan block της ιδιοτιμής λ_0 .

Σημείωση: Ο πίνακας Jordan ονομάζεται κάθε block-διαγώνιος πίνακας $J \in \mathbb{C}^{n \times n}$ που έχει διαγώνιους (τετραγωνικούς) υποπίνακες τους υποπίνακες Jordan (Jordan blocks) που αντιστοιχούν στις ιδιοτιμές του A .

Για κάθε $A \in \mathbb{C}^{n \times n}$, λοιπόν, υπάρχει αντιστρέψιμος $P \in \mathbb{C}^{n \times n}$, τέτοιος ώστε $P^{-1}AP = J$, όπου $J \in \mathbb{C}^{n \times n}$ είναι πίνακας Jordan που αντιστοιχεί στον A .

1.2 Στοιχεία Ανάλυσης Πινάκων

Για βαθύτερη ανάλυση θα ορίσουμε ακόμη κάποιες έννοιες που αφορούν τις νόρμες πινάκων στο $\mathbb{C}^{n \times n}$.

Ορισμός 19. Μια 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 μια γενίκευση της έννοιας της νόρμας, η οποία επιτρέπει σε μη μηδενικά διανύσματα να έχουν μηδενικό μέτρο.

Χαρακτηριστικά παραδείγματα νορμών διανυσμάτων στο \mathbb{C}^n είναι τα εξής:

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

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

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

$$\|x\|_2 = \|[x_1 \ x_2 \ \dots \ x_n]^T\|_2 = (|x_1|^2 + |x_2|^2 + \dots + |x_n|^2)^{1/2}.$$

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

$$\langle x, y \rangle = y^* x = [\bar{y}_1 \ \bar{y}_2 \ \dots \ \bar{y}_n] \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \bar{y}_1 x_1 + \bar{y}_2 x_2 + \dots + \bar{y}_n x_n.$$

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

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

- Η *μέγιστη νόρμα* (ή \max -νόρμα, ή ∞ -νόρμα) ορίζεται ως

$$\|x\|_\infty = \|[x_1 \ x_2 \ \dots \ x_n]^T\|_\infty = \max\{|x_1|, |x_2|, \dots, |x_n|\}.$$

Είναι γνωστό ότι το σύνολο $\mathbb{C}^{n \times n}$ των $n \times n$ μιγαδικών πινάκων είναι ένας διανυσματικός χώρος διάστασης n^2 , ο οποίος είναι ισόμορφος με το διανυσματικό χώρο \mathbb{C}^{n^2} .

Έτσι μπορούμε, επομένως, να ορίσουμε νόρμες πινάκων με τρόπο ανάλογο με αυτόν που ορίσαμε τις νόρμες διανυσμάτων.

Ορισμός 20. Μια συνάρτηση $\|\cdot\|: \mathbb{C}^{n \times n} \rightarrow \mathbb{R}$ ονομάζεται *νόρμα πινάκων* αν για κάθε $A, B \in \mathbb{C}^{n \times n}$, ικανοποιεί τα ακόλουθα:

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

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

$$\|I_n\| = \|I_n^2\| \leq \|I_n\|^2 \Rightarrow \|I_n\| \geq 1.$$

Επομένως, για κάθε $A \in \mathbb{C}^{n \times n}$,

$$1 \leq \|I_n\| = \|AA^{-1}\| \leq \|A\| \|A^{-1}\| \Rightarrow \|A^{-1}\| \geq \frac{1}{\|A\|}.$$

Κάποιες από τις νόρμες διανυσμάτων αποτελούν νόρμες πινάκων, όταν εφαρμόζονται στο διανυσματικό χώρο $\mathbb{C}^{n \times n}$, ενώ κάποιες άλλες όχι.

Τα πιο γνωστά παραδείγματα είναι οι l_p -νόρμες, για $p = 1, 2, \infty$.

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

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

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

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

όπου με $\text{trace}(\cdot)$ συμβολίζουμε το ίχνος του πίνακα και με A^* είναι ο αναστροφosuζυγής του πίνακα A .

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

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

Για παράδειγμα, έστω $A = \begin{bmatrix} 1 & 1 \\ -2 & 3 \end{bmatrix}$. Τότε

- $\|A\|_{l_1} = \left\| \begin{bmatrix} 1 & 1 \\ -2 & 3 \end{bmatrix} \right\|_{l_1} = \sum_{i,j=1}^2 |a_{ij}| = |1| + |1| + |-2| + |3| = 7.$
- $\|A\|_F = \left\| \begin{bmatrix} 1 & 1 \\ -2 & 3 \end{bmatrix} \right\|_F = \left(\sum_{i,j=1}^2 |a_{ij}|^2 \right)^{1/2} = (1^2 + 1^2 + (-2)^2 + 3^2)^{1/2} = \sqrt{15}.$
- $\|A\|_{l_\infty} = \left\| \begin{bmatrix} 1 & 1 \\ -2 & 3 \end{bmatrix} \right\|_{l_\infty} = \max\{|a_{ij}|: i, j = 1, 2, 3, 4\} = \max\{|1|, |1|, |-2|, |3|\} = 3.$

Με απλές πράξεις, μπορεί κανείς να επαληθεύσει ότι οι συναρτήσεις $\|\cdot\|_{l_1}$ και $\|\cdot\|_{l_2}$ αποτελούν πράγματι νόρμες πινάκων. Αντίθετα, η $\|\cdot\|_{l_\infty}$ αν και ικανοποιεί τα (i) - (iv) του Ορισμού 20, 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 $A \in \mathbb{C}^{n \times n}$, συμβολίζουμε με $a_1, a_2, \dots, a_n \in \mathbb{C}^n$ τις στήλες του A , τότε ο A γράφεται ως $A = [a_1 \ a_2 \ \dots \ a_n]$.

Έτσι μπορεί κανείς να παρατηρήσει ότι

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

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

Ορισμός 21. Έστω $\|\cdot\|$ μια νόρμα διανυσμάτων στο \mathbb{C}^n . Η επαγόμενη από την $\|\cdot\|$ νόρμα στο $\mathbb{C}^{n \times n}$ (*lub norm*) ορίζεται ως

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

Σημείωση: Αν $\|\cdot\|$ μια νόρμα διανυσμάτων στο \mathbb{C}^n και $\|\cdot\|$ η επαγόμενη από την $\|\cdot\|$ νόρμα στο $\mathbb{C}^{n \times n}$, τότε αποδεικνύεται ότι η $\|\cdot\|$ αποτελεί νόρμα πινάκων. Επιπλέον, ισχύει ότι $\|Ax\| \leq \|A\| \|x\|$ για κάθε $x \in \mathbb{C}^n$.

Ακόμη αναφέρουμε τα σημαντικότερα παραδείγματα νορμών πινάκων που επάγονται από τις γνωστές l_p -νόρμες. Θεωρούμε έναν τετραγωνικό πίνακα $A = [a_{ij}] \in \mathbb{C}^{n \times n}$.

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

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

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

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

Επομένως,

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

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

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

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

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

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

Η νόρμα $\|\cdot\|_{\infty}$ 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. Προφανώς και η νόρμα $\|\cdot\|_{\infty}$ (ως επαγόμενη) αποτελεί νόρμα πινάκων.

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

$$\|A\|_2 = \max_{x \neq 0} \left[\frac{x^* A^* A x}{x^* x} \right]^{\frac{1}{2}}.$$

Επίσης, ορίζεται και ως την μέγιστη ιδιάζουσα τιμή του πίνακα A , δηλαδή

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

όπου με $\sigma(A)$ συμβολίζουμε το φάσμα ενός πίνακα $A \in \mathbb{C}^{n \times n}$, δηλαδή το σύνολο των ιδιοτιμών του A και με A^* συμβολίζουμε τον αναστροφοσυζυγή του πίνακα A . Η νόρμα αυτή είναι νόρμα πινάκων, η οποία επάγεται από τη διανυσματική νόρμα l_2 , δηλαδή

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

Από την άλλη πλευρά, μερικές νόρμες πινάκων, όπως η νόρμα Frobenius και η l_1 -νόρμα παραπάνω, δεν επάγονται.

Ορισμός 22. Έστω ένας πίνακας $A \in \mathbb{C}^{n \times n}$ με φάσμα $\sigma(A) = \{ \lambda \in \mathbb{C} : \det(\lambda I_n - A) = 0 \}$.

Η φασματική ακτίνα του A ορίζεται ως

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

Σημείωση: Η φασματική ακτίνα δεν αποτελεί νόρμα πινάκων.

1.3 Παραγοντοποίηση QR

Η παραγοντοποίηση QR έχει δύο μορφές: την απλή παραγοντοποίηση QR και την πλήρη παραγοντοποίηση QR .

Θα εστιάσουμε την προσοχή μας στην απλή παραγοντοποίηση QR .

1.3.1 Ορθοκανονικοποίηση Gram-Schmidt. Έστω ένας πίνακας $A \in \mathbb{C}^{n \times m}$ ($n \geq m$), του οποίου οι στήλες a_1, a_2, \dots, a_m είναι γραμμικά ανεξάρτητες, δηλαδή ο βαθμός του A (= ο αριθμός των γραμμικά ανεξάρτητων στηλών του A) είναι $\text{rank}(A) = m$. Ο σκοπός μας είναι να κατασκευάσουμε ένα ορθοκανονικό σύστημα διανυσμάτων q_1, q_2, \dots, q_m για το οποίο θα ισχύει

$$\text{span}\{a_1, a_2, \dots, a_m\} = \text{span}\{q_1, q_2, \dots, q_m\},$$

όπου $\text{span}\{a_1, a_2, \dots, a_m\}$, $\text{span}\{q_1, q_2, \dots, q_m\}$ είναι οι γραμμικές θήκες των a_1, a_2, \dots, a_m και των q_1, q_2, \dots, q_m διανυσμάτων, αντίστοιχα.

Η μεθοδολογία που θα περιγραφεί παρακάτω είναι γνωστή ως *ορθοκανονικοποίηση Gram-Schmidt* και είναι βασισμένη στην παρακάτω παρατήρηση. Αν $x, y \in \mathbb{C}^n$ είναι δύο γραμμικά ανεξάρτητα διανύσματα και συμβολίζουμε 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 $q_1 = a_1 / \|a_1\|_2$. Στη συνέχεια κατασκευάζουμε ένα ορθογώνιο στο q_1 διάνυσμα $w_2 = a_2 - (q_1^* a_2)q_1$ και το κανονικοποιούμε και αυτό 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 αυτόν τον τρόπο, λαμβάνουμε τα ακόλουθα διανύσματα

$$q_1 = \frac{a_1}{r_{11}}, \quad q_2 = \frac{a_2 - r_{12} q_1}{r_{22}}, \quad q_3 = \frac{a_3 - r_{13} q_1 - r_{23} q_2}{r_{33}}, \dots, \quad q_m = \frac{a_m - \sum_{i=1}^{m-1} r_{im} q_i}{r_{mm}},$$

όπου

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

Ορίζουμε τον $n \times m$ πίνακα $Q = [q_1 \ q_2 \ \dots \ q_m]$ με ορθοκανονικό σύστημα στηλών και τον $m \times m$ άνω τριγωνικό πίνακα

$$\begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ 0 & r_{22} & \dots & r_{2m} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & r_{mm} \end{bmatrix}$$

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

$$A = QR \Leftrightarrow [a_1 \ a_2 \ \dots \ a_m] = [q_1 \ q_2 \ \dots \ q_m] \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1m} \\ 0 & r_{22} & \dots & r_{2m} \\ 0 & 0 & \ddots & \vdots \\ 0 & 0 & \dots & r_{mm} \end{bmatrix}$$

είναι άμεση και καλείται απλή παραγοντοποίηση QR του πίνακα A .

Παρατήρηση: Κάθε πίνακας $A \in \mathbb{C}^{n \times m}$ ($n \geq m$) βαθμού $\text{rank}(A) = m$ έχει μοναδική παραγοντοποίηση QR .

1.4 Παραγοντοποίηση SVD

Αξίζει να αναφέρουμε ότι μια από τις σημαντικότερες παραγοντοποιήσεις πινάκων είναι η γνωστή *παραγοντοποίηση ιδιάζουσών τιμών (singular value decomposition)*, την οποία και θα καλούμε *παραγοντοποίηση SVD*.

Θεώρημα: Αν A είναι ένας $n \times m$ μιγαδικός πίνακας, τότε υπάρχουν ορθομοναδιαίοι πίνακες $U \in \mathbb{C}^{n \times n}$ και $V \in \mathbb{C}^{m \times m}$ τέτοιοι ώστε $A = U \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_{\min\{n,m\}}\} V^*$, όπου $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min\{n,m\}} \geq 0$.

Ο διαγώνιος πίνακας $\text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_{\min\{n,m\}}\}$ συμβολίζεται με Σ και οι τιμές $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_{\min\{n,m\}} \geq 0$ καλούνται *ιδιάζουσες τιμές* του A . Αν γράψουμε $U = [u_1 \ u_2 \ \dots \ u_n]$ και $V = [v_1 \ v_2 \ \dots \ v_m]$, τότε τα διανύσματα στήλης u_i και v_i (με $1 \leq i \leq \min\{n,m\}$) λέγονται *αριστερά και δεξιά ιδιάζοντα διανύσματα* που αντιστοιχούν στην ιδιάζουσα τιμή σ_i . Από τις σχέσεις $AV = U\Sigma$ και $A^*U = V\Sigma^T$ προκύπτει ότι $A v_i = \sigma_i u_i$ και $A^* u_i = \sigma_i v_i$ ($1 \leq i \leq \min\{n,m\}$).

Ένα απλό παράδειγμα παραγοντοποίησης SVD είναι

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

Γενικά ισχύει $\sigma_1 = \sigma_1(A) = \|A\|_2$, ενώ μπορεί κανείς να δει ότι η (ορθομοναδιαία αναλλοίωτη) νόρμα Frobenius του A , $\|A\|_F = (\sum_{i,j} |a_{ij}|^2)^{1/2}$, γράφεται

$$\|A\|_F = \sqrt{\sigma_1^2 + \sigma_2^2 + \dots + \sigma_{\min\{n,m\}}^2}.$$

(Υπενθύμιση: Μια 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 $\|Ux\| = \|x\|$.)

Επιπλέον για $n \geq m$, ισχύει

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

ενώ για $n = m$,

$$|\det(A)| = \sigma_1 \sigma_2 \dots \sigma_n.$$

1.5 Ευστάθεια Τετραγωνικών Γραμμικών Συστημάτων

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

$$A = U \Sigma V^* = U \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_n\} V^* = \sum_{i=1}^n \sigma_i u_i v_i^*.$$

Τότε ισχύει

$$(U \text{diag}\{\sigma_1, \sigma_2, \dots, \sigma_n\} V^*) (V \text{diag}\{\sigma_1^{-1}, \sigma_2^{-1}, \dots, \sigma_n^{-1}\} U^*) = I_n,$$

δηλαδή

$$A^{-1} = V \text{diag}\left\{\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_n}\right\} U^*,$$

άρα και $\|A^{-1}\|_2 = 1/\sigma_n$. Επομένως το σύστημα $Ax = \beta$ έχει ως λύση τη

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

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

Για να μελετήσουμε την ευστάθεια του συστήματος

$$Ax = \beta,$$

θα θεωρήσουμε το διαταραγμένο και παραμετρικοποιημένο σύστημα

$$(A + \varepsilon F) x(\varepsilon) = \beta + \varepsilon f, \quad x(0) = x, \quad (1.5.1)$$

όπου $F \in \mathbb{C}^{n \times n}$ και $f \in \mathbb{C}^n$ δεδομένα και ε μία πραγματική μεταβλητή.

Επειδή ο πίνακας A είναι αντιστρέψιμος, υπάρχει περιοχή του 0, όπου ορίζεται η συνεχής διανυσματική συνάρτηση

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

Ακόμη, παραγωγίζοντας τη σχέση (1.5.1) κατά μέλη, καταλήγουμε στη διανυσματική παράγωγο

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

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

$$x(\varepsilon) = x + \varepsilon \dot{x}(0) + O(\varepsilon^2) = x + \varepsilon A^{-1} (f - F x) + O(\varepsilon^2).$$

Συνεπώς,

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

Σε αυτό το σημείο πρέπει να ορίσουμε το βαθμό (δείκτη) κατάστασης του πίνακα A .

Ορισμός 23. Έστω ένας $n \times n$ τετραγωνικός πίνακας $A \in \mathbb{C}^{n \times n}$ και $\|\cdot\|_\alpha$ μια τυχαία νόρμα πινάκων στο $\mathbb{C}^{n \times n}$. Η ποσότητα

$$\kappa(A) = \kappa_\alpha(A) = \begin{cases} \|A^{-1}\|_\alpha \|A\|_\alpha, & \text{αν ο } A \text{ είναι αντιστρέψιμος,} \\ \infty, & \text{διαφορετικά} \end{cases},$$

ονομάζεται *δείκτης κατάστασης (condition number)* του πίνακα A .

Άρα για τη 2-νόρμα ο δείκτης κατάστασης είναι

$$\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\sigma_1}{\sigma_n}$$

δεχόμενοι κατά σύμβαση ότι $\kappa_2(A) = \infty$, όταν ο A δεν είναι αντιστρέψιμος.

Με τη χρήση της ανισότητας $\|\beta\|_2 \leq \|A\|_2 \|x\|_2$, εύκολα 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 , αντίστοιχα. Δηλαδή το σχετικό σφάλμα της λύσης x μπορεί να γίνει ίσο με $\kappa_2(A)$ φορές το συνολικό σχετικό σφάλμα των A και β . Αυτό, ουσιαστικά, σημαίνει ότι ο βαθμός κατάστασης $\kappa_2(A)$ “ποσοτικοποιεί” την ευαισθησία του συστήματος $Ax = \beta$.

Σημείωση: Ο δείκτης κατάστασης ενός τετραγωνικού $n \times n$ πίνακα $A \in \mathbb{C}^{n \times n}$ παίρνει τιμές μεγαλύτερες ή ίσες του ένα. Δηλαδή

$$\kappa_\alpha(A) = \|A^{-1}\|_\alpha \|A\|_\alpha \geq \|A^{-1}A\|_\alpha = \|I_n\|_\alpha \geq 1.$$

Παρατήρηση: Οι πίνακες με σχετικά μικρό βαθμό κατάστασης, δηλαδή παραπλήσιο της μονάδας, ονομάζονται *πίνακες καλής κατάστασης (well-conditioned)*. Εάν ο δείκτης κατάστασης είναι μεγάλος, τότε λέμε ότι ο πίνακας είναι *κακής κατάστασης (ill-conditioned)*. Ο δείκτης κατάστασης είναι ∞ ή πολύ μεγάλος αν το δεξιό και το αριστερό ιδιοδιάνυσμα είναι ορθογώνια (δηλαδή κάθετα μεταξύ τους με το εσωτερικό τους γινόμενο να είναι μηδέν) ή σχεδόν ορθογώνια. Τέλος, εάν ο πίνακας έχει δείκτη κατάστασης ίσο με τη μονάδα, τότε λέμε πως έχει *ιδανική κατάσταση*. Πίνακες με ιδανική κατάσταση είναι ορθομοναδιαίοι πίνακες.

Πρόταση για το δείκτη κατάστασης: Έστω οι $n \times n$ τετραγωνικοί αντιστρέψιμοι πίνακες $A, B \in \mathbb{C}^{n \times n}$ και $\|\cdot\|_\alpha$ μια νόρμα πινάκων στο $\mathbb{C}^{n \times n}$. Τότε ισχύουν τα ακόλουθα:

(i) $\kappa_\alpha(A) = \kappa_\alpha(A^{-1})$.

(ii) $\kappa_\alpha(AB) \leq \kappa_\alpha(A) \kappa_\alpha(B)$.

Απόδειξη: (i) Έστω ότι ο αντιστρέψιμος πίνακας A έχει δείκτη κατάστασης

$$\kappa_\alpha(A) = \|A^{-1}\|_\alpha \|A\|_\alpha$$

και ο αντίστροφός του έχει δείκτη κατάστασης

$$\kappa_\alpha(A^{-1}) = \|A\|_\alpha \|A^{-1}\|_\alpha.$$

Άρα $\kappa_\alpha(A) = \kappa_\alpha(A^{-1})$.

(ii) Έστω ότι οι αντιστρέψιμοι πίνακες A, B έχουν δείκτη κατάστασης

$$\begin{aligned} \kappa_\alpha(AB) &= \|AB\|_\alpha \|(AB)^{-1}\|_\alpha = \|AB\|_\alpha \|B^{-1}A^{-1}\|_\alpha \\ &\leq \|A\|_\alpha \|B\|_\alpha \|A^{-1}\|_\alpha \|B^{-1}\|_\alpha \leq \kappa_\alpha(A) \kappa_\alpha(B). \quad \square \end{aligned}$$

Ορισμός 24. Έστω ένας πίνακας $A \in \mathbb{C}^{n \times n}$, λ μια απλή ιδιοτιμή του και $x, y \in \mathbb{C}^n$ τα αντίστοιχα δεξιά και αριστερά ιδιοδιανύσματα της ιδιοτιμής. Ονομάζουμε *δείκτη κατάστασης της ιδιοτιμής λ ως προς την Ευκλείδεια νόρμα*, την ακόλουθη ποσότητα

$$\kappa_A(\lambda) = \frac{\|x\|_2 \|y\|_2}{|y^*x|}.$$

Ο δείκτης κατάστασης απλής ιδιοτιμής αποτελεί ένα μέτρο ευαισθησίας της ιδιοτιμής αυτής.

1.6 Κανονικοί Πίνακες

Οι κανονικοί πίνακες εμφανίζονται σε ένα πολύ μεγάλο πλήθος σύγχρονων προβλημάτων. Η έντονη χρήση τους αποτυπώνεται στην ύπαρξη πάρα πολλών ισοδύναμων ορισμών στη βιβλιογραφία. Για την καλύτερη κατανόηση της έννοιας του κανονικού πίνακα αναφέρουμε παρακάτω κάποιες συνθήκες κανονικότητας χωρίς τις αποδείξεις τους. Υπενθυμίζουμε ότι ένας τετραγωνικός πίνακας $A \in \mathbb{C}^{n \times n}$ καλείται *κανονικός* όταν αντιμετατίθεται με τον αναστροφοσυζυγή του, δηλαδή όταν ικανοποιεί τη σχέση

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

Η έννοια, λοιπόν, του κανονικού πίνακα θα συνδυαστεί με τις ακόλουθες συνθήκες.

Πρόταση 1.6.1: (Συνθήκη 1) Ένας αντιστρέψιμος πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν ο A^{-1} είναι κανονικός.

Πρόταση 1.6.2: (Συνθήκη 2) Ένας αντιστρέψιμος πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν ο $A^{-1} A^*$ είναι ορθομοναδιαίος.

Πρόταση 1.6.3: (Συνθήκη 3) Ένας αντιστρέψιμος πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν ο A αντιμετατίθεται με τον πίνακα $A^{-1} A^*$.

Πρόταση 1.6.4: (Συνθήκη 4) Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν ο πίνακας U^*AU είναι κανονικός για οποιοδήποτε ορθομοναδιαίο πίνακα $U \in \mathbb{C}^{n \times n}$.

Πρόταση 1.6.5: (Συνθήκη 5) Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν υπάρχει ένας ορθομοναδιαίος πίνακας $U \in \mathbb{C}^{n \times n}$ και ένας διαγώνιος πίνακας $\Lambda \in \mathbb{C}^{n \times n}$ τέτοιοι ώστε $U^*AU = \Lambda$.

Πρόταση 1.6.6: (Συνθήκη 6) Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν είναι διαγωνοποιήσιμος και υπάρχει ορθοκανονική βάση του $\mathbb{C}^{n \times n}$ που αποτελείται από τα ιδιοδιανύσματα του A .

Πρόταση 1.6.7: (Συνθήκη 7) Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν ο A είναι διαγωνοποιήσιμος και οποιαδήποτε ιδιοδιανύσματά του, τα οποία προέρχονται από διαφορετικές ιδιοτιμές, είναι ορθογώνια.

Πρόταση 1.6.8: (Συνθήκη 8) Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν ο A είναι διαγωνοποιήσιμος και κάθε ιδιοδιάνυσμα $x \in \mathbb{C}^n$ του A είναι ιδιοδιάνυσμα και του A^* .

Πρόταση 1.6.9: (Συνθήκη 9) Ένας πίνακας $A \in \mathbb{C}^{n \times n}$, με (όχι απαραίτητα διακεκριμένες) ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_n$, είναι κανονικός αν και μόνο αν οι ιδιάζουσες τιμές του είναι $|\lambda_1|, |\lambda_2|, \dots, |\lambda_n|$.

1.7 Αριθμητικό Πεδίο-Αριθμητική Ακτίνα

Ορισμός 25. Το *αριθμητικό πεδίο* (*numerical range, field of values*) ενός τετραγωνικού πίνακα $A \in \mathbb{C}^{n \times n}$ συμβολίζεται με $F(A)$ και ορίζεται ως το σύνολο

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

Το φάσμα και το αριθμητικό πεδίο είναι σύνολα από τα οποία αντλούμε πληροφορίες για τον πίνακα. Το αριθμητικό πεδίο ενός πίνακα δίνει πληροφορίες που το φάσμα δε δύναται να δώσει.

Το φάσμα είναι ένα διακριτό σημειοσύνολο. Αντίθετα, το αριθμητικό πεδίο είναι πάντα ένα συμπαγές, δηλαδή κλειστό και φραγμένο, υποσύνολο του μιγαδικού επιπέδου. Επίσης, με βάση το Θεώρημα Toeplitz-Hausdorff, το αριθμητικό πεδίο είναι και κυρτό υποσύνολο του μιγαδικού επιπέδου.

Το $F(A)$ μπορεί να θεωρηθεί και ως η εικόνα της Ευκλείδειας μοναδιαίας σφαίρας $\{x \in \mathbb{C}^n, x^*x = 1\}$ στο \mathbb{C} , μέσω της απεικόνισης $x \rightarrow x^*Ax$. Η σφαίρα είναι συμπαγές σύνολο και η εν λόγω απεικόνιση είναι συνεχής. Από το θεώρημα του Weierstrass προκύπτει ότι το $F(A)$ είναι συμπαγές σύνολο στο \mathbb{C} ως συνεχής εικόνα συμπαγούς συνόλου.

Τα αριθμητικά πεδία κάποιων απλών πινάκων είναι τα ακόλουθα:

1.

$$\begin{aligned} F(\alpha I_n) &= \{x^*(\alpha I_n)x : x \in \mathbb{C}^n, x^*x = 1\} \\ &= \{\alpha(x^*x) : x \in \mathbb{C}^n, x^*x = 1\} \\ &= \{\alpha\}. \end{aligned}$$

2.

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

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

3.

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

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

Ορισμός 26. Η *αριθμητική ακτίνα* (*numerical radius*) ενός τετραγωνικού πίνακα A ορίζεται ως

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

Σημείωση: Η $r(A)$ στο διανυσματικό χώρο των τετραγωνικών πινάκων είναι διανυσματική νόρμα, αλλά δεν αποτελεί νόρμα πινάκων.

Ιδιότητες: Για κάθε $A \in \mathbb{C}^{n \times n}$ και $\alpha \in \mathbb{C}$, ισχύουν

$$(i) F(A + \alpha I_n) = F(A) + \alpha.$$

$$(ii) F(\alpha A) = \alpha F(A).$$

Παρατήρηση: Έστω ένας τετραγωνικός πίνακας A . Συμβολίζουμε με

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

το *ερμιτιανό μέρος* και το *αντιερμιτιανό μέρος* του A , αντίστοιχα. Κάθε τετραγωνικός πίνακας A μπορεί να γραφτεί στη μορφή $A = H(A) + S(A)$. Επιπλέον, οι πίνακες $H(A)$ και $-i S(A)$ είναι ερμιτιανοί. Για ένα μιγαδικό αριθμό, το φανταστικό και το πραγματικό του μέρος είναι οι προβολές του στο φανταστικό και στον πραγματικό άξονα, αντίστοιχα. Με ανάλογο τρόπο, το αριθμητικό πεδίο του ερμιτιανού και του αντιερμιτιανού μέρους του A είναι οι προβολές του $F(A)$ στον πραγματικό και το φανταστικό άξονα, αντίστοιχα.

Στη συνέχεια παραθέτουμε κάποιες προτάσεις που είναι χρήσιμες για την περαιτέρω κατανόηση του αριθμητικού πεδίου και της αριθμητικής ακτίνας και ορισμένες από αυτές θα μας χρειαστούν στη συνέχεια.

Πρόταση 1.7.1: Για κάθε πίνακα $A \in \mathbb{C}^{n \times n}$, ισχύουν

$$F(H(A)) = F\left(\frac{A+A^*}{2}\right) = \text{Re}(F(A)) \quad \text{και} \quad F(S(A)) = F\left(\frac{A-A^*}{2}\right) = i \text{Im}(F(A)),$$

όπου με $\text{Re}(\cdot)$ και $\text{Im}(\cdot)$ συμβολίζουμε το πραγματικό και το φανταστικό μέρος, αντίστοιχα.

Πρόταση 1.7.2: Για κάθε πίνακα $A \in \mathbb{C}^{n \times n}$, το φάσμα του, $\sigma(A)$, περιέχεται στο αριθμητικό του πεδίο του $F(A)$. Δηλαδή

$$\sigma(A) \subseteq F(A).$$

Πρόταση 1.7.3 (Υπο-προσθετικότητα): Για κάθε $A, B \in \mathbb{C}^{n \times n}$, ισχύει

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

Πρόταση 1.7.4: Για κάθε πίνακα $A, U \in \mathbb{C}^{n \times n}$, με U ορθομοναδιαίο, ισχύει

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

Απόδειξη: Εύκολα βλέπουμε ότι

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

Έτσι, έχει ολοκληρωθεί η απόδειξή μας. □

Πρόταση 1.7.5 (Κανονικότητα): Για κάθε κανονικό πίνακα $A \in \mathbb{C}^{n \times n}$, ισχύει

$$F(A) = \text{Co}(\sigma(A)),$$

δηλαδή το αριθμητικό πεδίο κανονικού πίνακα είναι η κυρτή θήκη των ιδιοτιμών του.

Απόδειξη: Αν ο πίνακας A είναι κανονικός, τότε $A = U^*\Lambda U$, όπου ο $\Lambda = \text{diag}\{\lambda_1, \lambda_2, \dots, \lambda_n\}$ είναι διαγώνιος πίνακας με $\lambda_1, \lambda_2, \dots, \lambda_n$ να είναι οι ιδιοτιμές του A (όχι κατ' ανάγκη διακεκριμένες) και U είναι ορθομοναδιαίος πίνακας. Από την ιδιότητα της ορθομοναδιαίας αναλλοιωσιμότητας (Πρόταση 1.7.4) ισχύει

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

Το $F(A)$ είναι απλά το σύνολο όλων των κυρτών συνδυασμών των διαγώνιων στοιχείων του A . Τα διαγώνια στοιχεία του A , όμως, είναι οι ιδιοτιμές του A και αυτό, επομένως, σημαίνει ότι $F(A) = \text{Co}(\sigma(A))$. Η απόδειξή μας έχει ολοκληρωθεί. \square

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 $\sqrt{\text{trace}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2}$.

Πρόταση 1.7.6: Για κάθε πίνακα $A \in \mathbb{C}^{n \times n}$, ισχύει $\frac{1}{2} \|A\|_2 \leq r(A) \leq \|A\|_2$.

Πρόταση 1.7.7: Έστω ένας πίνακας $A \in \mathbb{C}^{n \times n}$ με φασματική ακτίνα

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

Τότε ισχύει

$$\rho(A) \leq r(A).$$

Πρόταση 1.7.8 (Ανισότητα δυνάμεων): Έστω ένας πίνακας $A \in \mathbb{C}^{n \times n}$. Τότε για κάθε $n \in \mathbb{N}$, ισχύει

$$r(A^n) \leq r(A)^n.$$

1.8 Ψευδοφάσμα Πίνακα και Διαταραχές Ιδιοτιμών

Ιδιαίτερο ενδιαφέρον παρουσιάζει το ψευδοφάσμα ενός πίνακα $A \in \mathbb{C}^{n \times n}$, $\Lambda_\varepsilon(A)$, για δεδομένο $\varepsilon > 0$. Το σύνολο αυτό “διαστελλεται”, καθώς η παράμετρος ε αυξάνει. Αξίζει να σημειωθεί ότι το ψευδοφάσμα ορίζεται μόνο για νόρμες πινάκων που επάγονται από νόρμες διανυσμάτων.

Για δεδομένο $\varepsilon > 0$, το ψευδοφάσμα ενός $n \times n$ 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 . (Τα ψευδοφάσματα του A είναι εγκλιβωτισμένα υποσύνολα του μιγαδικού επιπέδου \mathbb{C} που επεκτείνονται για να συμπληρώσουν το επίπεδο καθώς $\varepsilon \rightarrow \infty$.)

Θεωρούμε έναν σταθερό $n \times n$ μιγαδικό πίνακα A . Δηλώνουμε τη μικρότερη ιδιάζουσα τιμή του A με $\underline{\sigma}(A)$ ($= \sigma_{\min}(A)$).

Ορισμός 27. Για κάθε $\varepsilon \geq 0$, το ε -ψευδοφάσμα $\Lambda_\varepsilon(A)$ του A είναι το σύνολο των αριθμών $w \in \mathbb{C}$ που ικανοποιούν οποιαδήποτε από τις παρακάτω ισοδύναμες συνθήκες:

- 1) $\Lambda_\varepsilon(A) = \{w \in \mathbb{C} : \|(w I_n - A)^{-1}\| \geq \varepsilon^{-1}\}$.
- 2) $\Lambda_\varepsilon(A) = \{w \in \mathbb{C} : w \in \Lambda(A+E) \text{ για κάποιο } E \text{ με } \|E\| \leq \varepsilon\}$.
- 3) $\Lambda_\varepsilon(A) = \{w \in \mathbb{C} : \|(w I_n - A) u\| \leq \varepsilon \text{ για κάποιο } u \in \mathbb{C}^n \text{ με } \|u\| = 1\}$.
- 4) Όταν αναφερόμαστε στη νόρμα $\|\circ\|_2$, συμβολίζοντας με $\underline{\sigma}(A)$ τη μικρότερη ιδιάζουσα τιμή του A , ο ακόλουθος ορισμός είναι επίσης ισοδύναμος:

$$\Lambda_\varepsilon(A) = \{w \in \mathbb{C} : \underline{\sigma}(w I_n - A) \leq \varepsilon\}. \quad (1.8.1)$$

Στην περίπτωση που $w \in \Lambda(A)$, θεωρούμε ότι $\|(w I_n - A)^{-1}\| = \infty$.

Σαφώς $\Lambda_0(A)$ είναι ακριβώς το φάσμα του A (εμείς το δηλώνουμε απλά $\Lambda(A)$).

Απόδειξη της ισοδυναμίας των ορισμών: Αρχικά θα αποδείξουμε τις συνεπαγωγές $(1) \Rightarrow (3) \Rightarrow (2) \Rightarrow (1)$ και έπειτα την ισοδυναμία $(1) \Leftrightarrow (4)$.

$[(1) \Rightarrow (3)]$ Θα υποθέσουμε ότι $\|(w I_n - A)^{-1}\| \geq \varepsilon^{-1}$. Τότε υπάρχει μοναδιαίο διάνυσμα $v \in \mathbb{C}^n$, έτσι ώστε

$$\|(w I_n - A)^{-1}\| = \|(w I_n - A)^{-1} v\|.$$

Αν ορίσουμε $z = (w I_n - A)^{-1} v$, τότε

$$\frac{1}{\varepsilon} \leq \|(w I_n - A)^{-1}\| = \frac{\|(w I_n - A)^{-1} v\|}{\|v\|} = \frac{\|z\|}{\|(w I_n - A) z\|}.$$

Επομένως, το διάνυσμα $u = z / \|z\|$ είναι μοναδιαίο και ικανοποιεί τη σχέση

$$\|(w I_n - A) u\| \leq \varepsilon.$$

$[(3) \Rightarrow (2)]$ Υποθέτουμε ότι υπάρχει ένα μοναδιαίο διάνυσμα $u \in \mathbb{C}^n$, τέτοιο ώστε $\|(w I_n - A) u\| \leq \varepsilon$. Τότε θα υπάρχει ακόμη ένα μοναδιαίο διάνυσμα $v \in \mathbb{C}^n$ που ικανοποιεί τη σχέση $(w I_n - A) u = \hat{\varepsilon} v$, για κάποιο θετικό αριθμό $\hat{\varepsilon} \leq \varepsilon$. Από το Θεώρημα Αναπαράστασης του Riesz και το Θεώρημα Hahn-Banach, υπάρχει ένα διάνυσμα $z \in \mathbb{C}^n$, τέτοιο ώστε $z^* u = 1$ και $\|z\| = 1$. Κατά συνέπεια, $w u = A u - \hat{\varepsilon} v z^* u = (A - \hat{\varepsilon} v z^*) u$. Δηλαδή $w \in \Lambda(A+E)$, όπου $E = -\hat{\varepsilon} v z^*$ και $\|E\| \leq \varepsilon$.

$[(2) \Rightarrow (1)]$ Υποθέτουμε ότι $w \in \Lambda(A+E)$ για κάποιον πίνακα E με $\|E\| \leq \varepsilon$. Τότε υπάρχει ακόμη ένα μοναδιαίο διάνυσμα $u \in \mathbb{C}^n$ για το οποίο ισχύει $(A+E) u = w u$. Με απλές πράξεις επαληθεύουμε εύκολα ότι $u = \|(w I_n - A)^{-1}\| E u$ και

$$1 = \|u\| = \|(w I_n - A)^{-1}\| \|E u\| \leq \|(w I_n - A)^{-1}\| \|E\| \leq \|(w I_n - A)^{-1}\| \varepsilon.$$

Έτσι θα ισχύει ότι $\|(w I_n - A)^{-1}\| \geq \varepsilon^{-1}$.

[(1) \Rightarrow (4)] Ας υποθέσουμε τώρα ότι $\|\circ\| = \|\circ\|_2$. Αν

$$\sigma_1(A) (= \bar{\sigma}(A) = \sigma_{\max}(A)) \geq \sigma_2(A) \geq \sigma_3(A) \geq \dots \geq \sigma_{n-1}(A) \geq \sigma_n(A) (= \underline{\sigma}(A) = \sigma_{\min}(A))$$

είναι οι θετικές ιδιάζουσες τιμές ενός $n \times n$ αντιστρέψιμου πίνακα A , τότε οι ιδιάζουσες τιμές του A^{-1} είναι οι

$$\frac{1}{\sigma_1(A)} \leq \frac{1}{\sigma_2(A)} \leq \frac{1}{\sigma_3(A)} \leq \dots \leq \frac{1}{\sigma_{n-1}(A)} \leq \frac{1}{\sigma_n(A)}.$$

Ακόμη, $\|A\|_2 = \sigma_1(A)$ και $\|A^{-1}\|_2 = 1/\sigma_n(A)$. Εύκολα, λοιπόν, βλέπει κανείς ότι

$$\|A^{-1}\|_2 = \sigma_{\max}(A^{-1}) = \frac{1}{\sigma_{\min}(A)}.$$

Συνεπώς, προκύπτει ότι

$$\|(w I_n - A)^{-1}\| \geq \frac{1}{\varepsilon}$$

αν και μόνο αν

$$\sigma_{\max}((w I_n - A)^{-1}) = \frac{1}{\sigma_{\min}(w I_n - A)} \geq \frac{1}{\varepsilon},$$

ή ισοδύναμα αν και μόνο αν

$$\sigma_{\min}(w I_n - A) \leq \varepsilon. \quad \square$$

Το ψευδοφάσμα είναι απλά σύνολα όλων των ιδιοτιμών των πινάκων εντός μιας δεδομένης απόστασης

$$\Lambda_\varepsilon(A) = \cup_{\|X-A\| \leq \varepsilon} \Lambda(X). \quad (1.8.2)$$

Πολλά θεωρήματα που αφορούν ιδιοτιμές είναι ειδικές περιπτώσεις θεωρημάτων ε -ψευδοφάσματος για $\varepsilon = 0$.

Αυτοί οι ορισμοί είναι για τη φασματική νόρμα στην οποία περιορίζουμε την προσοχή μας.

Μια απλή εφαρμογή της αρχής μεγίστου μέτρου δείχνει ότι κάθε συνεκτική συνιστώσα του $\Lambda_\varepsilon(A)$ περιέχει μια ιδιοτιμή του A . Θεωρούμε μια συγκεκριμένη ιδιοτιμή z του A και δείχνουμε ότι η ψευδοφασματική συνιστώσα περιέχει το z από την $\Lambda_\varepsilon^z(A)$. Για μικρό ε , διακεκριμένες ιδιοτιμές αντιστοιχούν σε διακεκριμένες συνιστώσες. Στην ορολογία του Wilkinson, κάθε συνιστώσα, τότε, αποτελείται από όλες τις πιθανές “διαδόχους” της αντίστοιχης ιδιοτιμής που συνδέονται με διαταραχές του A , μεγέθους όχι μεγαλύτερο από ε . Για μεγαλύτερα ε , οι συνιστώσες συγχωνεύονται και η έννοια της “διαδοχής” είναι πιο λεπτή.

Κάθε απλή ιδιοτιμή του z είναι μια αναλυτική συνάρτηση, τοπικά, του πίνακα A . Η κλίση αυτής της συνάρτησης είναι xy^* για αντίστοιχα μοναδιαία αριστερά και δεξιά ιδιοδιανύσματα y και x . Συνεπώς, για μικρό $\varepsilon \geq 0$, η ψευδοφασματική συνιστώσα $\Lambda_\varepsilon^z(A)$ είναι περίπου μια μπάλα με κέντρο το z και ακτίνα $|y^*x|^{-1}\varepsilon$ και ως εκ τούτου αυξάνεται σχεδόν γραμμικά στο ε με ρυθμό $|y^*x|^{-1}$, που αποτελεί τον κλασικό “δείκτη κατάστασης” της ιδιοτιμής του z . Επιπλέον, για μικρό ε , η συνιστώσα $\Lambda_\varepsilon^z(A)$ είναι μια κυρτή γειτονιά του z . Αυτή η ιδιότητα στην κυρτότητα ισχύει γενικότερα για μη-μειωτικές ιδιοτιμές, αλλά μπορεί και να “αποτύχει”.

Θεώρημα 1.1 (δείκτης κατάστασης): Το αντίστροφο του δείκτη κατάστασης μιας ιδιοτιμής είναι η διάταξή του ως μία ρίζα του ελάχιστου πολυωνύμου.

Υπενθύμιση: Με τον όρο *ελάχιστο πολυώνυμο* ενός πίνακα ορίζουμε το μη-μηδενικό πολυώνυμο με τον ελάχιστο βαθμό, το οποίο μηδενίζεται από τον πίνακα.

Αξίωμα - Υπόθεση: Ο πίνακας $A \in \mathbb{C}^{n \times n}$ έχει μια ιδιοτιμή $z \in \mathbb{C}$ και μετασχηματίζει σε Jordan σχηματισμό μέσω ενός πίνακα $P \in \mathbb{C}^{n \times n}$

$$P^{-1}AP = \begin{bmatrix} J & \dots & \dots & \dots & \dots \\ \dots & J & \dots & \dots & \dots \\ \vdots & \ddots & \ddots & \ddots & \ddots \\ \vdots & \ddots & \ddots & J & \vdots \\ \dots & \dots & \dots & \dots & \bar{A} \end{bmatrix}, \quad \text{όπου } J = \begin{bmatrix} z & 1 & 0 & \dots & 0 \\ 0 & z & 1 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & z \end{bmatrix}.$$

Επιπλέον, το Jordan block J εμφανίζεται k φορές και κάθε Jordan block στον πίνακα \bar{A} είναι είτε μικρότερο από J είτε αντιστοιχεί σε μια διαφορετική ιδιοτιμή. Οι στήλες του P είναι p_1, p_2, \dots, p_n και οι γραμμές του P^{-1} είναι q_1, q_2, \dots, q_n και ορίζουμε πίνακες

$$\tilde{Q} = \begin{bmatrix} q_m \\ q_{2m} \\ \vdots \\ q_{km} \end{bmatrix} \quad \text{και} \quad \tilde{P} = [p_1, p_{m+1}, \dots, p_{(k-1)m+1}]$$

έχοντας γραμμές και στήλες, αντίστοιχα στα αριστερά και δεξιά ιδιοδιανύσματα, που αντιστοιχούν στα k $m \times m$ Jordan blocks για την ιδιοτιμή z και για μια ποσότητα

$$\tilde{\alpha} = \|\tilde{P}\tilde{Q}\|.$$

Με την παραδοχή αυτή, m είναι η τάξη της ιδιοτιμής z ως τη ρίζα του ελαχίστου πολυωνύμου του A .

Θεώρημα 1.2 (Διαταραχή ιδιοτιμής): Υποθέτουμε ότι το αξίωμα ισχύει και θεωρούμε οποιαδήποτε ακολουθία αριθμών $\varepsilon_r \downarrow 0$ και πίνακες $A_r \in \mathbb{C}^{n \times n}$ που ικανοποιούν τη σχέση

$$A_r = A + \varepsilon_r B + o(\varepsilon_r) \quad \text{καθώς } r \rightarrow \infty$$

για κάποιο πίνακα $B \in \mathbb{C}^{n \times n}$. Θεωρούμε τις ιδιοτιμές του πίνακα $\tilde{Q}B\tilde{P} \in \mathbb{C}^{k \times k}$ με w_1, w_2, \dots, w_k (απαριθμημένα με αλγεβρική πολλαπλότητα). Τότε km από τις ιδιοτιμές του A_r μπορούν να εκφραστούν στη μορφή

$$z + \gamma^s w_j^{1/m} \varepsilon_r^{1/m} + o(\varepsilon_r^{1/m}),$$

για $s = 0, 1, \dots, m-1$ και $j = 1, 2, \dots, k$, όπου $w_j^{1/m}$ χαρακτηρίζει οποιαδήποτε $m^{\text{η}}$ ρίζα του w_j και γ είναι μια αρχική $m^{\text{η}}$ ρίζα της μονάδας, καθώς οι εναπομείναντες $n - km$ ιδιοτιμές είναι είτε εντός μιας απόστασης $o(\varepsilon_r^{1/m})$ της ιδιοτιμής z είτε παραμένουν οριοθετημένες ομοίμορφα μακριά από το z , καθώς $r \rightarrow \infty$.

Για την καλύτερη κατανόηση του παραπάνω θεωρήματος μπορούμε να το επαναδιατυπώσουμε με απλούστερες έννοιες:

Θεώρημα 1.3: Έστω λ μία απλή ιδιοτιμή του πίνακα A , με $x, y \in \mathbb{C}^n$ το δεξιό και αριστερό ιδιοδιάνυσμα, αντίστοιχα. Έστω $\tilde{A} = A + E$ μία “αρκετά μικρή” διαταραχή του A . Τότε υπάρχει μοναδική ιδιοτιμή $\tilde{\lambda}$ του \tilde{A} , τέτοια ώστε

$$\tilde{\lambda} = \lambda + \frac{y^* E x}{y^* x} + O(\|E\|_2^2).$$

Το λ ουσιαστικά αποτελεί το κέντρο του κύκλου και το $y^* E x$ “παίζει” το ρόλο της παραγώγου.

Από την παραπάνω σχέση 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. (Γενικά, στο δεύτερο μέλος κρατήσαμε τον πρώτο όρο από το ανάπτυγμα του Taylor.)

Όπως αναφέρθηκε, ο συντελεστής $\kappa_A(\lambda) = \frac{\|x\|_2 \|y\|_2}{|y^* x|}$ είναι ο γνωστός μας δείκτης κατάστασης της ιδιοτιμής λ , ως προς τη 2-νόρμα και προφανώς αποτελεί ένα μέτρο της ευαισθησίας της ιδιοτιμής αυτής.

Θεώρημα 1.4 (Ψευδοφασματική ανάπτυξη): Υποθέτουμε ότι το αξίωμα ισχύει, έτσι ειδικότερα, η ιδιοτιμή z είναι πολλαπλότητας m για το ελάχιστο πολυώνυμο του πίνακα A , δηλαδή το μεγαλύτερο Jordan block είναι τάξεως m . Τότε η Hausdorff απόσταση μεταξύ της ψευδοφασματικής συνιστώσας $A_\varepsilon^z(A)$ και του δίσκου $z + (\tilde{\alpha}\varepsilon)^{1/m} D$ είναι $o(\varepsilon^{1/m})$ καθώς $\varepsilon \downarrow 0$.

Σημείωση: Μικρές ψευδοφασματικές συνιστώσες είναι περίπου δίσκοι με κέντρο τις ιδιοτιμές. Ορίζουμε τον κλειστό μοναδιαίο δίσκο στο \mathbb{C} με D . Η απόσταση από ένα σημείο $s \in \mathbb{C}$ σε ένα σύνολο $T \subset \mathbb{C}$ είναι

$$d_T(s) = \inf_{t \in T} |s-t|$$

και η Hausdorff απόσταση μεταξύ του T και ενός άλλου συνόλου $S \subset \mathbb{C}$ είναι

$$h(S,T) = \max \{ \sup_{s \in S} d_T(s), \sup_{t \in T} d_S(t) \}.$$

Το παραπάνω θεώρημα μας υποδηλώνει ότι για μικρό ε το ψευδοφάσμα που προκύπτει μοιάζει με “μικρά κυκλάκια” με κέντρο την αντίστοιχη ιδιοτιμή κάθε φορά.

Απόδειξη του Θεωρήματος 1.4: Η απόδειξη θα γίνει για $m = 1$. Ανάλογα αποδεικνύεται και για τη γενικότερη περίπτωση (με χρήση του κατάλληλου ορισμού δείκτη κατάστασης πολλαπλής ιδιοτιμής και με κατάλληλη χρήση των Jordan block). Χωρίς βλάβη της γενικότητας, μπορούμε να υποθέσουμε ότι η ιδιοτιμή z είναι μηδέν. Εάν δεχθούμε ότι το αποτέλεσμα δεν ισχύει, τότε υπάρχει μια ακολουθία αριθμών $\varepsilon_r \downarrow 0$ και μια σταθερά $k > 0$, τέτοια ώστε

$$h((\tilde{\alpha} \varepsilon_r)^{1/m} D, \Lambda_{\varepsilon_r}^0(A)) > k \varepsilon_r^{1/m} \text{ για } r = 1, 2, \dots$$

Για $m = 1$, έχουμε

$$h(\tilde{\alpha} \varepsilon_r D, \Lambda_{\varepsilon_r}^0(A)) > k \varepsilon_r \text{ για } r = 1, 2, \dots,$$

όπου $\tilde{\alpha} = \kappa_A(z)$.

Επομένως, μπορούμε να υποθέσουμε την ύπαρξη μιας ακολουθίας σημείων $z_r \in \mathbb{C}$, τέτοια ώστε

$$z_r \in \Lambda_{\varepsilon_r}^0(A) \text{ και } |z_r| > (\tilde{\alpha} + k) \varepsilon_r \quad (1.8.3)$$

ή

$$|z_r| \leq \tilde{\alpha} \varepsilon_r \text{ και } d_{\Lambda_{\varepsilon_r}^0(A)}(z_r) > k \varepsilon_r. \quad (1.8.4)$$

Παίρνοντας μια υπακολουθία, μπορούμε να υποθέσουμε στην πραγματικότητα ότι μια από αυτές τις καταστάσεις ισχύει για όλα τα r .

Ας ξεκινήσουμε υποθέτοντας ότι η κατάσταση (1.8.3) ισχύει για όλα τα r . Ξεκάθαρα, $z_r \rightarrow 0$. Επιπλέον, υπάρχει μια ακολουθία των πινάκων $A_r \in \mathbb{C}^{n \times n}$ που ικανοποιούν την ανίσωση $\|A_r - A\| \leq \varepsilon_r$ και $z_r \in \Lambda(A_r)$ για όλα τα r . Προφανώς, ισχύει $A_r \neq A$ για όλα τα μεγάλα r . Έτσι, μπορούμε να υποθέσουμε ότι η ακολουθία των αριθμών $\delta_r = \|A_r - A\|$ δεν είναι ποτέ μηδέν. Παίρνοντας μια υπακολουθία, μπορούμε να υποθέσουμε ότι το $\delta_r^{-1}(A_r - A)$ συγκλίνει σε έναν πίνακα B με νόρμα ένα, έτσι ώστε

$$A_r = A + \delta_r B + o(\delta_r), \text{ καθώς } r \rightarrow \infty.$$

Μπορεί τώρα να εφαρμοστεί το Θεώρημα 1.2 (διαταραχή ιδιοτιμής). Στη συνέχεια, παίρνοντας μια επιπλέον υπακολουθία, υπάρχει μια ιδιοτιμή w του πίνακα $\tilde{Q}B\tilde{P}$, τέτοια ώστε

$$z_r = \delta_r w + o(\delta_r) \quad (\text{για } m = 1).$$

Επειδή η w είναι μια μη-μηδενική ιδιοτιμή του $\tilde{Q}B\tilde{P}$ (από τη δεύτερη Συνθήκη στη (1.8.3)), είναι επίσης και μια ιδιοτιμή του πίνακα $\tilde{P}\tilde{Q}B$, η οποία έχει νόρμα όχι μεγαλύτερη από $\tilde{\alpha}$. Ως εκ τούτου, παίρνουμε

$$|z_r| \leq \varepsilon_r \tilde{\alpha} + o(\varepsilon_r),$$

σε αντίθεση με την υπόθεσή μας ότι η Συνθήκη (1.8.3) ισχύει.

Υποθέτουμε από την άλλη πλευρά ότι η Συνθήκη (1.8.4) ισχύει για όλα τα r . Επειδή $0 \in \Lambda_{\varepsilon_r}^0(A)$, ξέρουμε ότι κάθε σημείο z_r είναι μη-μηδενικό. Έτσι μετά τη λήψη μιας υπακολουθίας μπορούμε να υποθέσουμε ότι η ποσότητα $|z_r|^{-1}z_r$ προσεγγίζει ένα σημείο w πάνω στο μοναδιαίο κύκλο.

Καθορίζουμε ένα ζευγάρι από μοναδιαία αριστερά και δεξιά μοναδικά διανύσματα $y, x \in \mathbb{C}^n$ που αντιστοιχούν στη μεγαλύτερη μοναδική τιμή $\tilde{\alpha}$ του πίνακα $\tilde{P}\tilde{Q}$. Επειδή $y^* \tilde{P} \tilde{Q} x = \tilde{\alpha}$, συμπεραίνουμε ότι $\tilde{\alpha} \in \Lambda(\tilde{Q} x y^* \tilde{P})$ και έτσι $\tilde{\alpha} w \in \Lambda(\tilde{Q}(w x y^*)\tilde{P})$. Το Θεώρημα 1.2 (διαταραχή ιδιοτιμής) τώρα δείχνει ότι

$$\tilde{\alpha} \delta w \in \Lambda(A + \delta w x y^*) + o(\delta) \subset \Lambda_\delta(A) + o(\delta), \quad (\text{για } m = 1) \quad (1.8.5)$$

καθώς $\delta \downarrow 0$. Στην πραγματικότητα μπορούμε να αντικαταστήσουμε το ψευδοφάσμα $\Lambda_\delta(A)$ με τη συνιστώσα του $\Lambda_\delta^0(A)$ πάνω στη δεξιά πλευρά. Διαφορετικά, θα υπήρχε μια ακολουθία $\delta_r \downarrow 0$ και μια μη-μηδενική ιδιοτιμή z' του πίνακα A , τέτοια ώστε

$$\tilde{\alpha} \delta_r w \in \Lambda_{\delta_r}^0(A) + o(\delta_r) \quad \text{καθώς } r \rightarrow \infty.$$

Αυτό είναι αδύνατο δεδομένου ότι η δεξιά πλευρά συγκλίνει σε μεμονωμένα $\{z'\}$, καθώς $r \rightarrow \infty$.

Έστω $\delta = \tilde{\alpha}^{-1} |z_r|$. Επειδή $\delta \leq \varepsilon_r$ (από τη Συνθήκη (1.8.4)), συμπεραίνουμε ότι

$$|z_r| w \in \Lambda_{\varepsilon_r}^0(A) + o(z_r),$$

καθώς $r \rightarrow \infty$, έτσι

$$d_{\Lambda_{\varepsilon_r}^0(A)}(z_r) \leq |z_r| w + o(z_r) - z_r = o(z_r) = o(\varepsilon_r).$$

Αλλά αυτό έρχεται σε αντίθεση με τη Συνθήκη (1.8.4).

Άρα η Hausdorff απόσταση μεταξύ της ψευδοφασματικής συνιστώσας $\Lambda_\varepsilon^Z(A)$ και στο δίσκο $z + (\tilde{\alpha}\varepsilon)D$ είναι $o(\varepsilon)$ καθώς $\varepsilon \downarrow 0$. \square

Το παραπάνω θεώρημα επαληθεύει το γεγονός ότι για αρκετά μικρές τιμές του ε , κάθε συνεκτική συνιστώσα του ψευδοφάσματος είναι σχεδόν κυκλικός δίσκος με κέντρο μια ιδιοτιμή και ακτίνα “ανάλογη” του δείκτη κατάστασης της ιδιοτιμής, όπως είναι ξεκάθαρο στα παραδείγματα της Παραγράφου 3.2.1 παρακάτω.

ΚΕΦΑΛΑΙΟ 2: Οι Δείκτες Κατάστασης του Προβλήματος Ιδιοτιμών Πίνακα

Για έναν $n \times n$ πίνακα A με διακριτές ιδιοτιμές μπορούμε να λάβουμε ορισμένες εκφράσεις για κάποιους δείκτες κατάστασης που σχετίζονται με την απλοποίηση του A στη Jordan κανονική μορφή του. Θα μελετήσουμε τους δείκτες κατάστασης και τη συμπεριφορά τους ως προς τη διακρίνουσα των ιδιοτιμών του A και ως προς τη Gram ορίζουσα των ιδιοδιανυσμάτων του A .

2.1 Οι Jordan Δείκτες Κατάστασης

Όπως αναφέραμε και προηγουμένως, αν $A = (a_{ij})$ ένας $n \times n$ πίνακας ορίζουμε ως $A^* = (\bar{a}_{ji})$ τον ανάστροφοσυζυγή του και

$$\|A\|_F = (\sum_{i,j=1}^n |a_{ij}|^2)^{1/2}, \quad \|A\|_2 = \max_{x \neq 0} \frac{\|Ax\|_2}{\|x\|_2},$$

όπου $\|x\|_2$ είναι το διάνυσμα της Ευκλείδειας νόρμας.

Επιπλέον, όπως αναφέρθηκε, αν ο J είναι ο Jordan κανονικής μορφής του A , τότε υπάρχουν όχι-μοναδικοί τετραγωνικοί πίνακες P , τέτοιοι ώστε

$$A = PJP^{-1}. \quad (2.1.1)$$

Σ' αυτό το κεφάλαιο, ορίζουμε τους Jordan δείκτες κατάστασης κ_F, κ_2 του A ως εξής

$$\kappa_F = \inf\{\|P^{-1}\|_F \cdot \|P\|_F\}, \quad \kappa_2 = \inf\{\|P^{-1}\|_2 \cdot \|P\|_2\}$$

για όλους τους P που ικανοποιούν την (2.1.1). Πάντα θα ισχύει

$$\kappa_2 \geq \|P^{-1}P\|_2 = \|I\|_2 \geq 1,$$

όπου $\kappa_2 = 1$, όταν ο A είναι ένας κανονικός πίνακας, γιατί η (2.1.1) τότε ικανοποιείται από ορθομοναδιαίους P .

Οι ανισότητες που αφορούν τους Jordan δείκτες κατάστασης έχουν δοθεί από τους Gautchi, Marcus, Bauer, Fike και πολλούς άλλους συγγραφείς.

Το ακόλουθο θεώρημα οφείλεται κυρίως στον Bauer του οποίου οι ιδέες χρησιμοποιήθηκαν στην απόδειξη που δίνεται παρακάτω.

Θεώρημα 2.1: Ισχύει η ακόλουθη ανισότητα

$$n - 2 + \kappa_2 + \kappa_2^{-1} \leq \kappa_F \leq \frac{1}{2} n (\kappa_2 + \kappa_2^{-1}).$$

Απόδειξη: Έστω $q_1 \geq q_2 \geq \dots \geq q_n$ είναι οι ιδιοτιμές του P^*P .

Είναι εύκολο να δούμε ότι η νόρμα $\|\cdot\|_F$ είναι ορθομοναδιαία αναλλοίωτη.

(Απόδειξη: Έστω πίνακας $A = (a_{ij}) \in \mathbb{C}^{n \times n}$. Δεδομένου ότι η l_2 -νόρμα διανυσμάτων είναι ορθομοναδιαία αναλλοίωτη, για κάθε ορθομοναδιαίο πίνακα $U \in \mathbb{C}^{n \times n}$, ισχύει

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

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

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

Άρα η νόρμα Frobenius στο $\mathbb{C}^{n \times n}$ είναι ορθομοναδιαία αναλλοίωτη. \square)

Με τη χρήση της μεθόδου SVD (singular value decomposition) και αξιοποιώντας το γεγονός ότι η νόρμα $\|\cdot\|_F$ είναι ορθομοναδιαία αναλλοίωτη προκύπτει ότι

$$\|P\|_F = \sqrt{q_1 + q_2 + \dots + q_n} \quad (= \sqrt{\text{trace}(P^*P)}),$$

όπου q_1, q_2, \dots, q_n είναι οι ιδιοτιμές του πίνακα P^*P .

Έχουμε, λοιπόν,

$$\|P\|_F^2 \|P^{-1}\|_F^2 = (q_1 + q_2 + \dots + q_n) (q_1^{-1} + q_2^{-1} + \dots + q_n^{-1}) = e^* Q e \cdot e^* Q^{-1} e,$$

όπου $Q = \text{diag}(q_i)$ και e^* είναι το διάνυσμα $(1, 1, \dots, 1)$.

Για να συνεχίσουμε την απόδειξη του θεωρήματος θα χρησιμοποιήσουμε την ανισότητα Kantorovich.

Ανισότητα Kantorovich: Έστω ο πίνακας A ένας $n \times n$ τετραγωνικός και θετικά ορισμένος συμμετρικός πίνακας με ιδιοτιμές $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n > 0$ και x είναι ένα μη-μηδενικό διάνυσμα τάξης n . Τότε ισχύει

$$1 \leq \frac{(x^T A x) \cdot (x^T A^{-1} x)}{(x^T x)^2} \leq \left\{ \frac{1}{2} \left[\left(\frac{\lambda_1}{\lambda_n} \right)^{\frac{1}{2}} + \left(\frac{\lambda_1}{\lambda_n} \right)^{-\frac{1}{2}} \right] \right\}^2.$$

Η ανισότητα του Kantorovich, λοιπόν, δείχνει ότι

$$(e^* Q e) \cdot (e^* Q^{-1} e) \leq \left\{ \frac{1}{2} (p + p^{-1}) \right\}^2 (e^* e)^2 = \left\{ \frac{1}{2} n (p + p^{-1}) \right\}^2,$$

όπου $p = (q_1/q_n)^{\frac{1}{2}} = \|P\|_2 \cdot \|P^{-1}\|_2$.

Ως εκ τούτου

$$\kappa_F \leq \|P\|_F \cdot \|P^{-1}\|_F \leq \frac{1}{2} n (p + p^{-1}).$$

Η δεξιά πλευρά του Θεωρήματος 2.1 προκύπτει αμέσως, αφού ισχύει $\kappa_2 = \inf p$ για όλα τα P που ικανοποιούν την (2.1.1).

Εάν

$$u = (q_1 q_n)^{-\frac{1}{2}} (q_2 + \dots + q_{n-1}), \quad v = (q_1 q_n)^{\frac{1}{2}} (q_2^{-1} + \dots + q_{n-1}^{-1}),$$

τότε $uv \geq (n-2)^2$ από την ανισότητα του Cauchy και

$$\begin{aligned} \|P\|_F^2 \|P^{-1}\|_F^2 &= (q_1 q_n)^{-\frac{1}{2}} (q_1 + q_2 + \dots + q_n) (q_1 q_n)^{\frac{1}{2}} (q_1^{-1} + q_2^{-1} + \dots + q_n^{-1}) \\ &= (p + u + p^{-1}) (p^{-1} + v + p) \geq [p + p^{-1} + (uv)^{\frac{1}{2}}]^2. \end{aligned}$$

Στο σημείο αυτό υπενθυμίζουμε το Θεώρημα Cauchy-Schwarz.

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

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

ή

$$\|x_1 \bar{y}_1 + x_2 \bar{y}_2 + \dots + x_n \bar{y}_n\|^2 \leq (|x_1|^2 + \dots + |x_n|^2) (|y_1|^2 + \dots + |y_n|^2).$$

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

Από $(uv)^{\frac{1}{2}} \geq (n-2)$ και $p \geq \kappa_2 \geq 1$, έχουμε

$$\|P\|_F \cdot \|P^{-1}\|_F \geq p + p^{-1} + n - 2 \geq \kappa_2 + \kappa_2^{-1} + n - 2,$$

από την οποία προκύπτει η αριστερή πλευρά του Θεωρήματος 2.1. Και οι δύο πλευρές του Θεωρήματος 2.1 τείνουν στην ισότητα, όταν ο πίνακας A είναι κανονικός, γιατί τότε $\kappa_2 = 1$. \square

2.2 Οι Δείκτες Κατάστασης Ιδιοτιμών Πίνακα

Όπως έχουμε αναφέρει και στο πρώτο Κεφάλαιο, ο δείκτης κατάστασης αντιστρέψιμου πίνακα είναι $\kappa(A) = \kappa_\alpha(A) = \|A^{-1}\|_\alpha \|A\|_\alpha$, όπου $A \in \mathbb{C}^{n \times n}$ και $\|\cdot\|_\alpha$ μια τυχαία νόρμα πινάκων στο $\mathbb{C}^{n \times n}$. Επιπλέον, η ποσότητα $\kappa_A(\lambda_j) = \frac{\|x\|_2 \|y\|_2}{|y^* x|}$ είναι γνωστή ως δείκτης κατάστασης της απλής ιδιοτιμής λ_j ως προς τη 2-νόρμα, όπου $\lambda_j \in \sigma(A)$ και x, y είναι τα δεξιά και αριστερά ιδιοδιανύσματα του A που αντιστοιχούν στις ιδιοτιμές λ_j . Δηλαδή,

$$A x = \lambda_j x, \quad y^* A = \lambda_j y^*. \quad (2.2.1)$$

Ακόμη θέτουμε

$$s_j = y^* x / \|x\|_2 \|y\|_2 \quad (2.2.2)$$

να είναι ο αριθμός Wilkinson, ώστε $\kappa_A(\lambda_j) = |s_j|^{-1}$ για $A \in \mathbb{C}^{n \times n}$, $\lambda_j \in \sigma(A)$ και $j = 1, 2, \dots$.

Εάν λ_j είναι μια απλή ιδιοτιμή, τότε τα x, y είναι μοναδικά, εκτός από ένα βαθμωτό πολλαπλάσιο και το $\kappa_A^{-1}(\lambda_j)$ προσδιορίζεται μονοσήμαντα. Ο Wilkinson έδειξε ότι ο $\kappa_A(\lambda_j)$ δίνει κάποιο μέτρο της ευαισθησίας της ιδιοτιμής λ_j για μικρές διαταραχές των στοιχείων του A . Ο Bauer και ο Fike έδειξαν ότι

$$1 \leq \kappa_A(\lambda_j) \leq \frac{1}{2} (\kappa_2 + \kappa_2^{-1}). \quad (2.2.3)$$

Στη συνέχεια, υπενθυμίζουμε την έννοια ενός συμπληρωματικού (adjoint) πίνακα που θα μας χρειαστεί σε παρακάτω θεωρήματα.

Ορισμός 28. Έστω $A = (a_{ij})$ ένας $n \times n$ πίνακας. Αν πάρουμε το στοιχείο a_{ij} και “κόψουμε” την i γραμμή και τη j στήλη, τότε προκύπτει ένας $(n-1) \times (n-1)$ πίνακας M_{ij} του οποίου η ορίζουσα $|M_{ij}|$ λέγεται *ελάσσονα ορίζουσα (minor)* του στοιχείου a_{ij} . Αν πάρουμε την τιμή της ελάσσονας ορίζουσας με πρόσημο ανάλογο με τη θέση του στοιχείου a_{ij} , δηλαδή $A_{ij} = (-1)^{i+j} |M_{ij}|$, τότε έχουμε *αλγεβρικό συμπλήρωμα* A_{ij} του στοιχείου a_{ij} . *Συμπληρωματικός (adjoint) πίνακας* του A λέγεται ο πίνακας

$$\text{adj}A = (A_{ij})^T = (A_{ji}),$$

όπου A_{ij} είναι το αλγεβρικό συμπλήρωμα του στοιχείου a_{ij} του πίνακα A με $i, j = 1, 2, \dots, n$.

Στο ακόλουθο θεώρημα ο $\text{adj}C$ φανερώνει το πίνακα (C_{ij}) , όπου C_{ij} είναι το αλγεβρικό συμπλήρωμα του στοιχείου στην $j^{\text{η}}$ γραμμή και στην $i^{\text{η}}$ στήλη του πίνακα C .

Θεώρημα 2.2: Εάν λ_j είναι μια απλή ιδιοτιμή του A , τότε

$$\kappa_A(\lambda_j) = \|\text{adj}(\lambda_j I - A)\|_F \cdot \prod_j |\lambda_j - \lambda_i|^{-1}, \quad (2.2.4)$$

όπου το \prod_j δηλώνει το γινόμενο για όλα $i \neq j$ με $1 \leq i \leq n$.

Συγκεκριμένα συμβολίζουμε $\prod_j |\lambda_j - \lambda_i|^{-1} = \prod_{\substack{i=1 \\ i \neq j}}^n |\lambda_j - \lambda_i|^{-1}$.

Απόδειξη: Για μη-μοναδικά C, Q, R οι σχέσεις

$$C \operatorname{adj} C = (\operatorname{adj} C) C = I \det C, \quad (2.2.5)$$

$$(\operatorname{adj} C) (\operatorname{adj} C)^* = (\operatorname{adj} C^* C), \quad (2.2.6)$$

$$R [\operatorname{adj}(QCR)] Q = (\operatorname{adj} C) \det(RQ) \quad (2.2.7)$$

δίνουν την ιδιότητα $\operatorname{adj} C = C^{-1} \det C$, όπου $\det(\cdot)$ συμβολίζουμε την ορίζουσα πίνακα.

(Υπενθύμιση: Αν $A = [\alpha_{11}]$ είναι ένας 1×1 πίνακας, τότε η ορίζουσά του ορίζεται ως $\det(A) = \alpha_{11}$. Η ορίζουσα ενός $n \times n$ πίνακα ορίζεται αναδρομικά με όρους των ορίζουσών των $(n-1) \times (n-1)$ πινάκων, όπως παρακάτω. Για κάθε j

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

όπου A_{ij} είναι ο $(n-1) \times (n-1)$ πίνακας που σχηματίζεται, αν σβήσουμε την $i^{\text{η}}$ γραμμή και τη $j^{\text{η}}$ στήλη του A .)

{ Ενδεικτικά, αποδεικνύουμε τη (2.2.7). Ομοίως αποδεικνύονται και οι υπόλοιπες.

Έχουμε

$$\begin{aligned} R [\operatorname{adj}(QCR)] Q &= R [R^{-1} C^{-1} Q^{-1} \det(QCR)] Q = C^{-1} Q^{-1} \det(QR) \det(C) Q = \\ &= C^{-1} \det(C) \det(RQ) Q^{-1} Q = C^{-1} \det(C) \det(RQ) = \\ &= (\operatorname{adj} C) \det(RQ). \end{aligned}$$

Οι παραπάνω σχέσεις ισχύουν επίσης από τη συνέχεια, όταν οι πίνακες είναι μοναδικοί.

Από το θεώρημα του Schur, το οποίο θα συζητηθεί και θα αποδειχθεί αναλυτικά παρακάτω στο Κεφάλαιο 3, θα ισχύει $A = U^* T U$, όπου U είναι ένας ορθομοναδιαίος πίνακας και T είναι ένας άνω τριγωνικός πίνακας. Τα διαγώνια στοιχεία του T είναι οι ιδιοτιμές του πίνακα A . Επιπλέον, ο πίνακας U μπορεί να επιλεγεί, έτσι ώστε η συγκεκριμένη ιδιοτιμή λ_j να είναι το οδηγό διαγώνιο στοιχείο του T . Με $Q = U$, $R = U^*$ και $C = \lambda_j I - A$ η σχέση (2.2.7) μας δίνει

$$\|\operatorname{adj}(\lambda_j I - A)\|_F = \|\operatorname{adj}(\lambda_j I - T)\|_F,$$

αφού ο U είναι ορθομοναδιαίος. Συγκεκριμένα

$$\begin{aligned} \|\operatorname{adj}(\lambda_j I - A)\|_F &= \|U^* [\operatorname{adj}(U(\lambda_j I - A)U^*)] U\|_F = \|\operatorname{adj}(U(\lambda_j I - A)U^*)\|_F = \\ &= \|\operatorname{adj}(U\lambda_j I U^* - UAU^*)\|_F = \|\operatorname{adj}(\lambda_j I - UAU^*)\|_F = \|\operatorname{adj}(\lambda_j I - T)\|_F. \end{aligned}$$

Δεδομένου ότι ο Wilkinson έδειξε ότι οι $\kappa_A^{-1}(\lambda_j)$ είναι αμετάβλητοι στο πλαίσιο ενός ορθομοναδιαίου μετασχηματισμού αρκεί να αποδείξουμε το Θεώρημα 2.2 στην ειδική περίπτωση, όπου $A = T$. Τότε ο $\lambda_j I - A$ και κατ' επέκταση ο $\text{adj}(\lambda_j I - A)$ είναι άνω τριγωνικοί πίνακες και το οδηγό διαγώνιο στοιχείο του $\text{adj}(\lambda_j I - A)$ είναι $\prod_j (\lambda_j - \lambda_i)$. Παρατηρούμε ότι όλα τα στοιχεία στην πρώτη στήλη του $\lambda_j I - A$ είναι μηδέν. Τότε τα μόνα μη-μηδενικά αλγεβρικά συμπληρώματα είναι εκείνα τα οποία σχηματίζουν την πρώτη γραμμή του $\text{adj}(\lambda_j I - A)$.

Για να το καταλάβουμε καλύτερα ας δούμε ένα παράδειγμα

Έστω $A = T = \begin{pmatrix} \lambda_j & \alpha & \beta \\ 0 & \lambda_1 & \gamma \\ 0 & 0 & \lambda_2 \end{pmatrix}$ με $\alpha, \beta, \gamma \in \mathbb{C}$ και λ_1, λ_2 είναι ιδιοτιμές του A .

Τότε $\lambda_j I - A = \begin{pmatrix} 0 & -\alpha & -\beta \\ 0 & \lambda_j - \lambda_1 & -\gamma \\ 0 & 0 & \lambda_j - \lambda_2 \end{pmatrix}$.

Παρατηρούμε πράγματι ότι όλα τα στοιχεία στην πρώτη στήλη του $\lambda_j I - A$ είναι ίσα με το μηδέν.

Το οδηγό διαγώνιο στοιχείο του $\text{adj}(\lambda_j I - A)$ είναι

$$A_{11} = (-1)^{1+1} \begin{vmatrix} \lambda_j - \lambda_1 & -\gamma \\ 0 & \lambda_j - \lambda_2 \end{vmatrix} = (\lambda_j - \lambda_1)(\lambda_j - \lambda_2).$$

Ακόμη

$$A_{21} = (-1)^{2+1} \begin{vmatrix} -\alpha & -\beta \\ 0 & \lambda_j - \lambda_2 \end{vmatrix} = \alpha(\lambda_j - \lambda_2), \quad A_{31} = (-1)^{3+1} \begin{vmatrix} -\alpha & -\beta \\ \lambda_j - \lambda_1 & -\gamma \end{vmatrix} = \alpha\gamma + \beta(\lambda_j - \lambda_1),$$

$$A_{12} = (-1)^{1+2} \begin{vmatrix} 0 & -\gamma \\ 0 & \lambda_j - \lambda_2 \end{vmatrix} = 0, \quad A_{22} = (-1)^{2+2} \begin{vmatrix} 0 & -\beta \\ 0 & \lambda_j - \lambda_2 \end{vmatrix} = 0, \quad A_{32} = (-1)^{3+2} \begin{vmatrix} 0 & -\beta \\ 0 & -\gamma \end{vmatrix} = 0,$$

$$A_{13} = (-1)^{1+3} \begin{vmatrix} 0 & \lambda_j - \lambda_1 \\ 0 & 0 \end{vmatrix} = 0, \quad A_{23} = (-1)^{2+3} \begin{vmatrix} 0 & -\alpha \\ 0 & 0 \end{vmatrix} = 0, \quad A_{33} = (-1)^{3+3} \begin{vmatrix} 0 & -\alpha \\ 0 & \lambda_j - \lambda_1 \end{vmatrix} = 0.$$

Άρα

$$\text{adj}(\lambda_j I - A) = \begin{pmatrix} (\lambda_j - \lambda_1)(\lambda_j - \lambda_2) & \alpha(\lambda_j - \lambda_2) & \alpha\gamma + \beta(\lambda_j - \lambda_1) \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}.$$

Πράγματι, τα μόνα μη-μηδενικά στοιχεία του $\text{adj}(\lambda_j I - A)$ είναι εκείνα, τα οποία σχηματίζουν την πρώτη γραμμή του πίνακα.

Ομοίως αποδεικνύεται και στη γενική περίπτωση.

Ως εκ τούτου,

$$\|adj(\lambda_j I - A)\|_F = \|y\|_2, \quad (2.2.8)$$

όπου το διάνυσμα y^* είναι η πρώτη γραμμή του $adj(\lambda_j I - A)$.

Εάν το διάνυσμα x είναι η πρώτη στήλη του $adj(\lambda_j I - A)$, τότε το x έχει μόνο ένα μη-μηδενικό στοιχείο και συγκεκριμένα το $\prod_j (\lambda_j - \lambda_i)$. Έχουμε, επομένως,

$$\|x\|_2 = \prod_j |\lambda_j - \lambda_i| \text{ και } y^* x = \prod_j (\lambda_j - \lambda_i)^2. \quad (2.2.9)$$

Με $C = \lambda_j I - A$, η (2.2.5) γίνεται

$$(\lambda_j I - A) adj(\lambda_j I - A) = \{adj(\lambda_j I - A)\} (\lambda_j I - A) = 0,$$

από την οποία προκύπτει ότι τα x, y ικανοποιούν τη (2.2.1).

Τότε από τις σχέσεις (2.2.2), (2.2.8) και (2.2.9), έχουμε

$$\begin{aligned} \kappa_A^{-1}(\lambda_j) &= |y^* x| / \|x\|_2 \cdot \|y\|_2 = \\ &= \prod_j |\lambda_j - \lambda_i|^2 / \{ \prod_j |\lambda_j - \lambda_i| \cdot \|adj(\lambda_j I - A)\|_F \} = \\ &= \prod_j |\lambda_j - \lambda_i| / \|adj(\lambda_j I - A)\|_F. \end{aligned}$$

$$\text{Άρα } \kappa_A(\lambda_j) = \|adj(\lambda_j I - A)\|_F \cdot \prod_j |\lambda_j - \lambda_i|^{-1}.$$

Έτσι το θεώρημα αποδείχθηκε. □

Το επόμενο θεώρημα τροποποιεί ελαφρώς ένα αποτέλεσμα του Wilkinson. Παρόμοιες ιδέες χρησιμοποιήθηκαν νωρίτερα από τον Bauer σε ένα γενικότερο πλαίσιο.

Θεώρημα 2.3: Εάν όλες οι ιδιοτιμές του A είναι απλές, τότε

$$\kappa_F = \kappa_A(\lambda_1) + \kappa_A(\lambda_2) + \dots + \kappa_A(\lambda_n). \quad (2.2.10)$$

Απόδειξη: Έστω X ένας $n \times n$ πίνακας του οποίου οι στήλες είναι τα δεξιά ιδιοδιανύσματα u_1, u_2, \dots, u_n του A κανονικοποιημένα έτσι ώστε $\|u_j\|_2 = 1$ για όλα τα j . Εάν Y είναι ένας πίνακας του οποίου οι γραμμές είναι τα κανονικοποιημένα αριστερά ιδιοδιανύσματα του A , τότε

$$AX = XJ, \quad YA = JY, \quad YXJ = JYX, \quad (2.2.11)$$

όπου $J = \text{diag}(\lambda_j)$.

Δεδομένου ότι οι ιδιοτιμές είναι διακριτές, μόνο οι διαγώνιοι πίνακες μπορούν να αντιμετατίθονται με τον J και ως εκ τούτου ο YX είναι ένας διαγώνιος πίνακας. Από τη σχέση (2.2.2) συμπεραίνουμε ότι το j° διαγώνιο στοιχείο είναι το s_j . Έτσι, λοιπόν, έχουμε

$$YX = S = \text{diag}(s_1, s_2, \dots, s_n) \quad (2.2.12)$$

και $X^{-1} = S^{-1}Y$. Δεδομένου ότι τα ιδιοδιανύσματα του A είναι μοναδικά εκτός, ίσως, από ένα βαθμωτό πολλαπλάσιο, κάθε πίνακας που ικανοποιεί την (2.1.1) είναι της μορφής $P = XR$, όπου $R = \text{diag}(r_1, r_2, \dots, r_n)$. Τότε $\|P\|_F^2 = \sum |r_i|^2$, γιατί οι στήλες του X έχουν νόρμα ίση με τη μονάδα (τα διανύσματα είναι κανονικοποιημένα). Επίσης, $P^{-1} = R^{-1}X^{-1} = R^{-1}S^{-1}Y$ και $\|P^{-1}\|_F^2 = \sum |r_i s_i|^{-2}$, γιατί οι γραμμές του Y έχουν νόρμα ίση με τη μονάδα (τα διανύσματα είναι και αυτά κανονικοποιημένα). Η ανισότητα Cauchy θα μας δώσει

$$\|P\|_F^2 \|P^{-1}\|_F^2 = \sum |r_i|^2 \sum |r_i s_i|^{-2} \geq (\sum |s_i|^{-1})^2,$$

για όλα τα (r_i) με ισότητα, όταν $(r_i) = (|s_i|^{-1/2})$.

Ως εκ τούτου, $\sum |s_i|^{-1} = \sum \kappa_A(\lambda_i) = \inf\{\|P\|_F \cdot \|P^{-1}\|_F\} = \kappa_F$ με $i = 1, 2, \dots, n$, και η απόδειξη ολοκληρώθηκε. \square

Μια σαφής έκφραση για το κ_F δίνεται από τις (2.2.4) και (2.2.10), υπό την προϋπόθεση ότι οι ιδιοτιμές του A είναι διακριτές. Όταν οι ιδιοτιμές δεν είναι διακριτές, το πρόβλημα υπολογισμού του κ_F έχει σημαντικές δυσκολίες, μερικές από τις οποίες θα αναφερθούν παρακάτω.

2.3 Οπτικοποίηση του σημαντικού ρόλου του Δείκτη Κατάστασης της Ιδιοτιμής για μικρές Διαταραχές

2.3.1 Αριθμητικά Παραδείγματα και Συμπεράσματα. Αφού ορίσαμε τους δείκτες κατάστασης, εξετάζουμε τη συμπεριφορά τους. Ο δείκτης κατάστασης αποτελεί χρήσιμο εργαλείο για τη διερεύνηση της ευστάθειας (ή αστάθειας) ενός πίνακα. Ορίζεται μόνο για αντιστρέψιμους πίνακες και αν ο πίνακας τείνει να γίνει μη αντιστρέψιμος, τότε ο δείκτης κατάστασης αυτού τείνει στο άπειρο. Ακόμη, ο δείκτης κατάστασης καθορίζει και το πώς διαταραχές του πίνακα επηρεάζουν τη λύση.

Μπορούμε να δούμε κάποια παραδείγματα προκειμένου να κατανοήσουμε τον τρόπο υπολογισμού του δείκτη κατάστασης και φυσικά τη συμπεριφορά του.

Παράδειγμα 1.

Έστω ότι έχουμε τον 4×4 άνω τριγωνικό πίνακα

$$A = \begin{bmatrix} 1 & 0.1 & 0.02 & 0.05 \\ 0 & 2 & 0.1 & 0.03 \\ 0 & 0 & 3 & 0.04 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

Ο δείκτης κατάστασης είναι $\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2$.

Έχουμε, κατ' αρχάς,

$$\begin{aligned} A^*A &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0.1 & 2 & 0 & 0 \\ 0.02 & 0.1 & 3 & 0 \\ 0.05 & 0.03 & 0.04 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0.1 & 0.02 & 0.05 \\ 0 & 2 & 0.1 & 0.03 \\ 0 & 0 & 3 & 0.04 \\ 0 & 0 & 0 & 4 \end{bmatrix} = \\ &= \begin{bmatrix} 1 & 0.1 & 0.02 & 0.05 \\ 0.1 & 4.01 & 0.202 & 0.065 \\ 0.02 & 0.202 & 9.0104 & 0.124 \\ 0.05 & 0.065 & 0.124 & 16.005 \end{bmatrix}. \end{aligned}$$

Οι ιδιοτιμές του A^*A υπολογίζονται από τη σχέση

$$|\lambda_j - A^*A| = 0.$$

Προκύπτει, λοιπόν,

$$\lambda_1 = 16.008, \lambda_2 = 9.016, \lambda_3 = 4.005, \lambda_4 = 0.997.$$

Άρα $\|A\|_2 = \max\{\sqrt{\lambda}: \lambda \in \sigma(A^*A)\} = \sqrt{16.008} \cong 4.000999875 \cong 4$.

Ακόμη έχουμε ότι

$$A^{-1} = \begin{bmatrix} 1 & -0.05 & -0.005 & -0.012 \\ 0 & 0.5 & -0.017 & -0.004 \\ 0 & 0 & 0.333 & -0.003 \\ 0 & 0 & 0 & 0.25 \end{bmatrix}$$

και συνεπώς,

$$(A^{-1})^* A^{-1} = \begin{bmatrix} 1 & -0.05 & -0.005 & -0.012 \\ -0.05 & 0.2525 & -0.00825 & -0.0014 \\ -0.005 & -0.00825 & 0.111203 & -0.000871 \\ -0.012 & -0.0014 & -0.000871 & 0.062669 \end{bmatrix}.$$

Οι ιδιοτιμές του $(A^{-1})^* A^{-1}$ υπολογίζονται από τη σχέση

$$|\lambda_j - (A^{-1})^* A^{-1}| = 0.$$

Προκύπτει, λοιπόν,

$$\lambda_1 = 1.004, \lambda_2 = 0.250, \lambda_3 = 0.111, \lambda_4 = 0.062.$$

Άρα $\|A^{-1}\|_2 = \max\{\sqrt{\lambda}: \lambda \in \sigma((A^{-1})^* A^{-1})\} = \sqrt{1.004} \cong 1.001998 \cong 1$ και $\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2 = 4 \times 1 = 4$.

Παράδειγμα 2.

Έστω ότι έχουμε τον 4×4 άνω τριγωνικό πίνακα

$$A = \begin{bmatrix} 1 & 10 & 100 & 100 \\ 0 & 2 & 20 & 50 \\ 0 & 0 & 3 & 100 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

Ο δείκτης κατάστασης είναι $\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2$.

Έχουμε, κατ' αρχάς,

$$\begin{aligned} A^* A &= \begin{bmatrix} 1 & 0 & 0 & 0 \\ 10 & 2 & 0 & 0 \\ 100 & 20 & 3 & 0 \\ 100 & 50 & 100 & 4 \end{bmatrix} \begin{bmatrix} 1 & 10 & 100 & 100 \\ 0 & 2 & 20 & 50 \\ 0 & 0 & 3 & 100 \\ 0 & 0 & 0 & 4 \end{bmatrix} = \\ &= \begin{bmatrix} 1 & 10 & 100 & 100 \\ 10 & 104 & 1040 & 1100 \\ 100 & 1040 & 10409 & 11300 \\ 100 & 1100 & 11300 & 22516 \end{bmatrix}. \end{aligned}$$

Οι ιδιοτιμές του $A^* A$ υπολογίζονται από τη σχέση

$$|\lambda_j - A^* A| = 0.$$

Προκύπτει, λοιπόν,

$$\lambda_1 = 29357.149, \lambda_2 = 3672.811, \lambda_3 = 0.040, \lambda_4 = 0.000.$$

Άρα $\|A\|_2 = \max\{\sqrt{\lambda}: \lambda \in \sigma(A^* A)\} = \sqrt{29357.149} \cong 171,339280$.

Ακόμη

$$A^{-1} = \begin{bmatrix} 1 & -5 & 0 & 37.5 \\ 0 & 0.5 & -3.333 & 77.083 \\ 0 & 0 & 0.333 & -8.333 \\ 0 & 0 & 0 & 0.25 \end{bmatrix}.$$

Άρα

$$(A^{-1})^* A^{-1} = \begin{bmatrix} 1 & -5 & 0 & 37.5 \\ -5 & 25.25 & -1.6665 & -148.9585 \\ 0 & -1.6665 & 11.219778 & -259.692528 \\ 37.5 & -148.9585 & -259.692528 & 7417.540278 \end{bmatrix}.$$

Οι ιδιοτιμές του $(A^{-1})^* A^{-1}$ υπολογίζονται από τη σχέση

$$|\lambda_j - (A^{-1})^* A^{-1}| = 0.$$

Προκύπτει, λοιπόν,

$$\lambda_1 = 7429.816, \lambda_2 = 25.194, \lambda_3 = 0, \lambda_4 = 0.$$

Άρα $\|A^{-1}\|_2 = \max\{\sqrt{\lambda}: \lambda \in \sigma((A^{-1})^* A^{-1})\} = \sqrt{7429.816} \cong 86.196380 \cong 86.20$ και $\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2 = 171,339280 \times 86.20 = 14769,445936 \cong 14769,45$.

Παρατηρούμε, δηλαδή, ότι όσο μεγαλύτερα είναι τα στοιχεία του άνω τριγωνικού πίνακα, τόσο “εκτοξεύεται” ο δείκτης κατάστασης. Στην αρχή βάλαμε μικρές τιμές και ο δείκτης κατάστασης βγήκε μικρός. Μεγαλώνοντας κατά πολύ τα νούμερα παρατηρούμε ότι και ο δείκτης κατάστασης αυξήθηκε πάρα πολύ.

Έχοντας ορίσει το ψευδοφάσμα, παρατηρήσαμε τη συμπεριφορά του με την ύπαρξη διαταραχής.

Θα χρειαστούμε το μαθηματικό πακέτο Matlab για να “κατασκευάσουμε” το ψευδοφάσμα. Για το σχεδιασμό του ψευδοφάσματος ενός πίνακα, έστω A , θα ορίσουμε ένα κατάλληλο χωρίο που περιέχει ολόκληρο ή μέρος του $\Lambda_\varepsilon(A)$ και θα υπολογίσουμε την ποσότητα $\underline{\sigma}(wI - A)$ πάνω στα σημεία της διαμέρισης αυτού. Έτσι, θα σχεδιάσουμε την καμπύλη $g(w) = \underline{\sigma}(wI - A) = \varepsilon$ με την κατάλληλη εντολή στο Matlab για τις διάφορες τιμές του ε .

Ανοίγουμε ένα new-script στο Matlab και γράφουμε τις παρακάτω εντολές:

```
function pseudofasma = psegridA(A, epsilons, xrange, yrange, hx, hy)
% Στη συγκεκριμένη εντολή κατασκευάζουμε τη συνάρτηση pseudofasma
% προκειμένου να εκτιμήσουμε το φάσμα μας και για διάφορα ε θέλουμε
% να εκτιμήσουμε το όρια του ε-ψευδοφάσματος ενός τετραγωνικού πίνακα
% A.
%Ως epsilons θα τοποθετήσουμε τις διάφορες τιμές που θέλουμε να
%βάλουμε στο ε.
%Ως x-range ορίζουμε το αριστερό και το δεξί όριο του άξονα x.
%Ως y-range ορίζουμε το αριστερό και το δεξί όριο του άξονα y.
%Ως hx ορίζουμε το μήκος πλέγματος για τον x-άξονα.
%Ως hy ορίζουμε το μήκος πλέγματος για τον y-άξονα.
figure, hold on, tic
epsilons = detach(epsilons);
% Το tic αφορά το χρόνο.
% Διαχωρίσαμε τα epsilons (e1>=...>=eN).
N = length(epsilons);
eps = epsilons(N);
```

```

[t,d] = size(A);
% Παίρνουμε την ελάχιστη τιμή του ε.
n1 = round((xrange(2)-xrange(1))/hx);
n2 = round((yrange(2)-yrange(1))/hy);
Z = zeros(n1+1,n2+1);
Iterations = 0; expoints = 0; inpoints = 0;
% Κατασκευάσαμε ένα πλέγμα και θέσαμε κάποιες αρχικές συνθήκες.
% Στο παρακάτω θα προσπαθήσουμε να βρούμε την ελάχιστη ιδιάζουσα τιμή
% σε κάθε σημείο της διαμέρισης.
for x = xrange(1):hx:xrange(2)
I = round((x-xrange(1))/hx)+1);
for y = yrange(1):hy:yrange(2)
J = round((y-yrange(1))/hy)+1);
%Βρίσκουμε τη x-συνταταγμενη του σημείου.
%Βρίσκουμε τη y-συνταταγμενη του σημείου.
mu = x+(y*i);
Z(I,J) = min(svd(mu*eye(d,d)-A));
Iterations = Iterations+1;
Y(J) = yrange(1)+(J-1)*hy;
end
X(I) = xrange(1)+(I-1)*hx;
end
% Στο παραπάνω σημείο mu εφαρμόσαμε τη μέθοδο της SVD και, έτσι,
% βρίσκουμε την ελάχιστη ιδιάζουσα τιμή.
%
% Στη συνέχεια δίνουμε των αριθμό των επαναλήψεων και
% σχεδιάζουμε τη γραφική μας.
number_of_total_grid_points=(n1+1)*(n2+1)
number_of_iterations=Iterations
hold on
for j = 1:N
ZZ = Z-epsilons(j)*ones(size(Z));
contour(X,Y,ZZ',[0 0], 'g')
end
axis([xrange(1),xrange(2),yrange(1),yrange(2)])
xlabel('Ο άξονας των Πραγματικών Αριθμών')
ylabel('Ο άξονας των Φανταστικών Αριθμών')
% Υπολογίζουμε τώρα τις ιδιοτιμές μας.
eigen = eig(A);
plot(real(eigen), imag(eigen), 'r+')
hold off, toc
% Το toc αφορά το χρόνο.

```

Στη συνέχεια σε ένα άλλο νέο script γράφουμε τις εντολές:

```
function D = detach(V)
N = length(V);
for i = 1:N-1
for j = 1:N-1
if V(j)<V(j+1)
temp = V(j);
V(j) = V(j+1);
V(j+1) = temp;
end
end
end
D = V;
```

Αφού αποθηκεύσουμε τα αρχεία μας, “τρέχουμε” το πρόγραμμά μας.

Μπορούμε πλέον παρατηρώντας διάφορα παραδείγματα να εξετάσουμε το ψευδοφάσμα.

Παράδειγμα 3.

Έστω ότι έχουμε τον 3×3 άνω τριγωνικό πίνακα $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix}$. Ο πίνακας αυτός

είναι διαγωνοποιήσιμος και $A = \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} = S^{-1} \begin{bmatrix} 1 & 0 & 0 \\ 0 & 4 & 0 \\ 0 & 0 & 6 \end{bmatrix} S$, όπου ο πίνακας

μετασχηματισμού $S = \begin{bmatrix} 1 & 0.5547 & 0.5108 \\ 0 & 0.8321 & 0.7982 \\ 0 & 0 & 0.3193 \end{bmatrix}$ έχει δείκτη κατάστασης $\kappa_2(S) = \|S\|_2 \|S^{-1}\|_2 = 6.8066$.

Πηγαίνοντας, λοιπόν, στο παράθυρο εργασιών του Matlab καλούμε τη συνάρτηση psegridA. Επιπλέον, επιλέγουμε το εύρος τιμών στον άξονα x να είναι το $[-4,8]$ και στον άξονα y να είναι το $[-4,4]$. Ακόμη, επιλέγουμε για μήκη πλέγματος τις τιμές 0,01 και 0,01 για τους άξονες x, y , αντίστοιχα.

Αρχικά, θα βρούμε το ψευδοφάσμα του πίνακα A για τις τιμές $\varepsilon = 0.2, 1$.

Έχουμε, λοιπόν,

```
>> A=[1,2,3;0,4,5;0,0,6]
```

A =

```
1 2 3
0 4 5
0 0 6
```

```
>> c=psegridA(A,[0.2,1],[-4,8],[-4,4],0.01,0.01)
```

```
number_of_total_grid_points =
```

```
962001
```

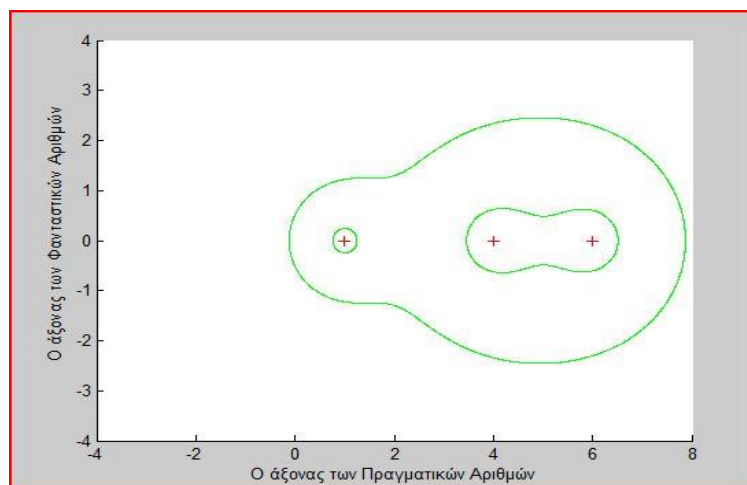
```
number_of_iterations =
```

```
962001
```

```
Elapsed time is 31.831681 seconds.
```

Η παραπάνω συνάρτηση μας απέδωσε σε 962001 επαναλήψεις 962001 σημεία. Ο χρόνος που χρειάστηκε ήταν 31.831681 δευτερόλεπτα.

Η γραφική παράσταση του ε -ψευδοφάσματος του πίνακα A για $\varepsilon = 0.2$ και $\varepsilon = 1$ είναι



Σχήμα 1: Τα ψευδοφάσματα του A για $\varepsilon = 0.2$ και $\varepsilon = 1$.

Παρατηρούμε εκ πρώτης όψεως ότι για $\varepsilon = 0.2$ προσεγγίζουμε πιο πολύ την ιδιοτιμή παρά για $\varepsilon = 1$.

Για να σιγουρέψουμε το αποτέλεσμά μας μπορούμε να εκτελέσουμε την ίδια εντολή χρησιμοποιώντας αυτή τη φορά περισσότερες τιμές για το ε . Με βάση το θεώρημά μας όσο μικρότερο είναι το ε , τόσο περισσότερο θα τείνουμε σε κύκλο με κέντρο την αντίστοιχη ιδιοτιμή. Εκτελούμε, επομένως, την ίδια εντολή για $\varepsilon = 0.01, 0.05, 0.1, 0.2, 0.5, 0.8, 1$. Έτσι, έχουμε

```
>> A=[1,2,3;0,4,5;0,0,6]
```

```
A =
```

```
1 2 3
0 4 5
0 0 6
```

```
>> c=psegridA(A,[0.01,0.05,0.1,0.2,0.5,0.8,1],[-4,8],[-4,4],0.01,0.01)
```

```
number_of_total_grid_points =
```

```
962001
```

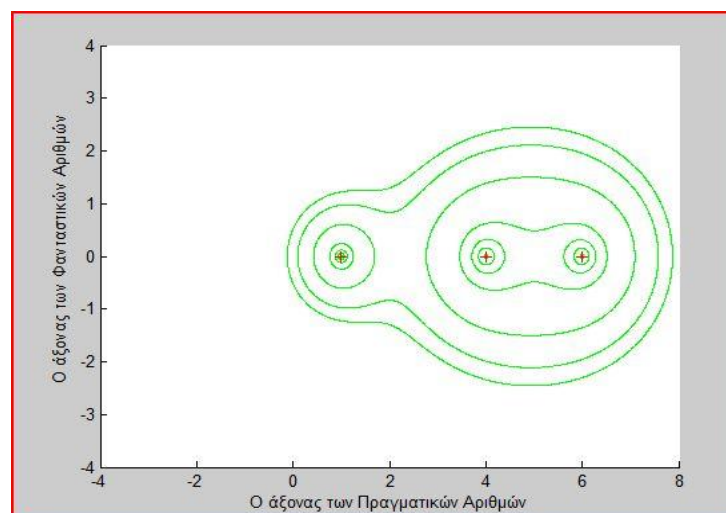
```
number_of_iterations =
```

```
962001
```

Elapsed time is 35.470292 seconds.

Ομοίως πάλι αποδίδουμε 962001 σημεία σε 962001 επαναλήψεις. Για να υλοποιηθεί αυτό χρειάστηκε χρόνος 35.470292 δευτερόλεπτα.

Το γράφημά μας είναι



Σχήμα 2: Τα ψευδοφάσματα του A για $\varepsilon = 0.01, 0.05, 0.1, 0.2, 0.5, 0.8, 1$.

Παρατηρούμε ότι πράγματι όσο μικρότερο γίνεται το ε , τόσο πιο κοντά προσεγγίζουμε την αντίστοιχη ιδιοτιμή. Έτσι, πράγματι για μικρές τιμές του ε προκύπτουν κύκλοι με κέντρο τις αντίστοιχες ιδιοτιμές. Το ψευδοφάσμα μας, δηλαδή, θα παριστάνει “μικρά κυκλάκια” γύρω από τις ιδιοτιμές για $\varepsilon \rightarrow 0$.

Μπορούμε στο συγκεκριμένο παράδειγμα να βάλουμε για ε μια πολύ μικρή τιμή ($\varepsilon = 0.045$). Έχουμε

```
>> c=psegridA(A,[0.045],[-4,8],[-4,4],0.01,0.01)
```

```
number_of_total_grid_points =
```

```
962001
```

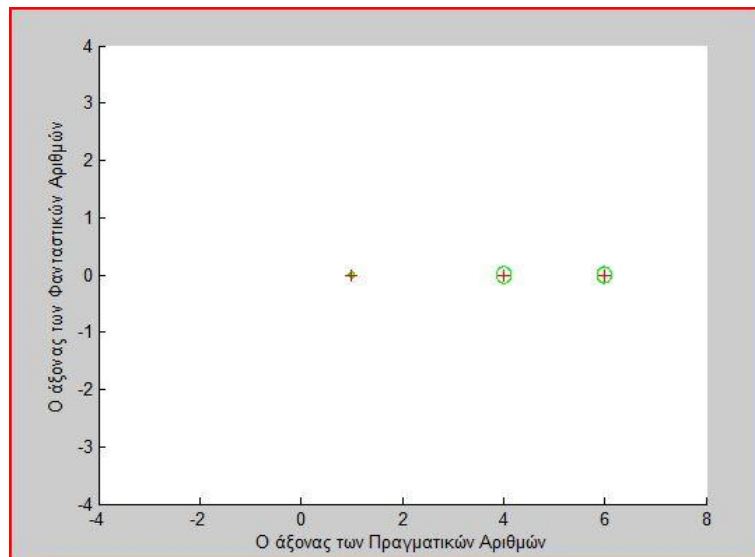
```
number_of_iterations =
```

```
962001
```

Elapsed time is 30.644394 seconds.

Ο χρόνος που απαιτείται για να ολοκληρωθεί η απεικόνιση 962001 σημείων είναι 30.644394 δευτερόλεπτα.

Το γράφημά μας είναι



Σχήμα 3: Τα ψευδοφάσματα του A για $\varepsilon = 0.045$.

Πράγματι, διαπιστώνουμε την ύπαρξη των κύκλων. Το ψευδοφάσμα, επομένως, αναπαριστά κύκλους με κέντρο τις ιδιοτιμές.

Παράδειγμα 4.

Έστω ότι έχουμε τον 4×4 άνω τριγωνικό πίνακα

$$A = \begin{bmatrix} 1 & 0.1 & 0.02 & 0.05 \\ 0 & 2 & 0.1 & 0.03 \\ 0 & 0 & 3 & 0.04 \\ 0 & 0 & 0 & 4 \end{bmatrix}.$$

Πηγαίνουμε, λοιπόν, πάλι στο παράθυρο εργασιών του Matlab καλούμε τη συνάρτηση `psegridA`. Επιπλέον, επιλέγουμε το εύρος τιμών στον άξονα x να είναι το $[-4,8]$ και στον άξονα y να είναι το $[-4,4]$. Ακόμη, επιλέγουμε για μήκη πλέγματος τις τιμές 0,01 και 0,01 για τους άξονες x , y , αντίστοιχα.

Αρχικά θα βρούμε το ψευδοφάσμα του πίνακα A για τις τιμές $\varepsilon = 0.2, 1$.

Έχουμε, λοιπόν,

```
>> A=[1,0.1,0.02,0.05;0,2,0.1,0.03;0,0,3,0.04;0,0,0,4]
```


A =

```
1.0000  0.1000  0.0200  0.0500
      0   2.0000  0.1000  0.0300
      0    0    3.0000  0.0400
      0    0     0    4.0000
```

```
>> c=psegridA(A,[0.2,1],[-4,8],[-4,4],0.01,0.01)
```

number_of_total_grid_points =

962001

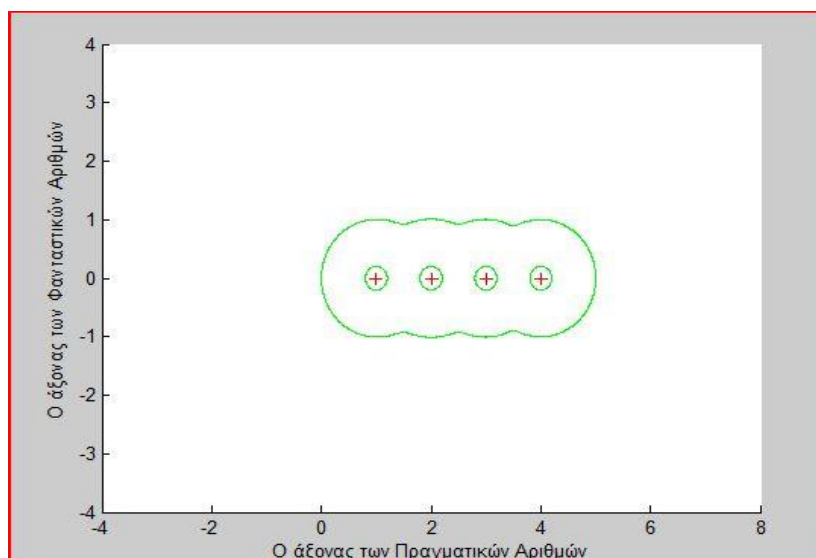
number_of_iterations =

962001

Elapsed time is 37.657316 seconds.

Η παραπάνω συνάρτηση μας απέδωσε σε 962001 επαναλήψεις 962001 σημεία. Ο χρόνος που χρειάστηκε ήταν 37.657316 δευτερόλεπτα.

Η γραφική παράσταση του ε -ψευδοφάσματος του πίνακα A για $\varepsilon = 0.2$ και $\varepsilon = 1$ είναι



Σχήμα 4: Τα ψευδοφάσματα του A για $\varepsilon = 0.2$ και $\varepsilon = 1$.

Παρατηρούμε εκ πρώτης όψεως ότι για $\varepsilon = 0.2$ προσεγγίζουμε πιο πολύ την ιδιοτιμή παρά για $\varepsilon = 1$.

Για να σιγουρέψουμε το αποτέλεσμά μας μπορούμε να εκτελέσουμε την ίδια εντολή χρησιμοποιώντας αυτή τη φορά περισσότερες τιμές για το ε . Με βάση το θεώρημά

μας όσο μικρότερο είναι το ϵ , τόσο περισσότερο θα τείνουμε σε κύκλο με κέντρο την αντίστοιχη ιδιοτιμή. Εκτελούμε, επομένως, την ίδια εντολή για $\epsilon = 0.01, 0.05, 0.1, 0.2, 0.5, 0.8, 1$. Έτσι, θα έχουμε

```
>> A=[1,0.1,0.02,0.05;0,2,0.1,0.03;0,0,3,0.04;0,0,0,4]
```

A =

```
1.0000 0.1000 0.0200 0.0500
0      2.0000 0.1000 0.0300
0      0      3.0000 0.0400
0      0      0      4.0000
```

```
>> c=psegridA(A,[0.01,0.05,0.1,0.2,0.5,0.8,1],[-4,8],[-4,4],0.01,0.01)
```

number_of_total_grid_points =

962001

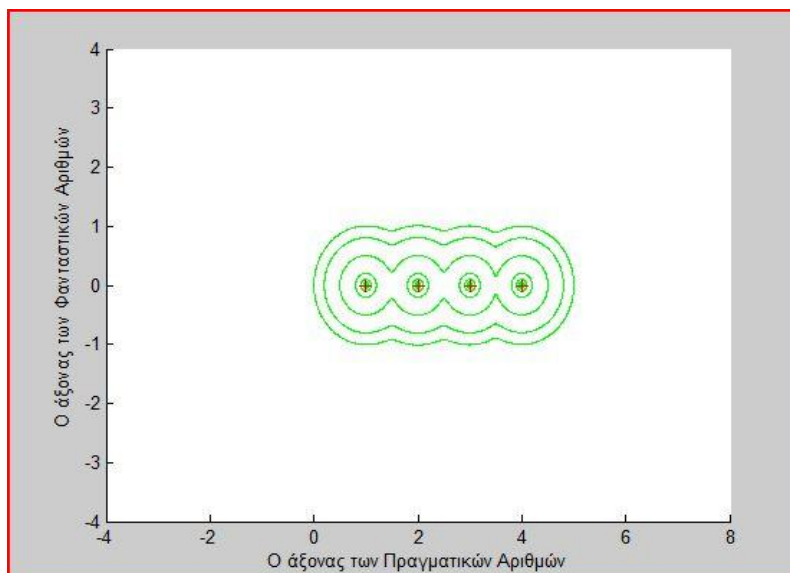
number_of_iterations =

962001

Elapsed time is 40.030690 seconds.

Ομοίως πάλι αποδίδουμε 962001 σημεία σε 962001 επαναλήψεις. Για να υλοποιηθεί αυτό χρειάστηκε χρόνος 40.03069 δευτερόλεπτα.

Το γράφημά μας είναι



Σχήμα 5: Τα ψευδοφάσματα του A για $\epsilon = 0.01, 0.05, 0.1, 0.2, 0.5, 0.8, 1$.

Διαπιστώνουμε ότι όσο μικρότερο γίνεται το ε , τόσο πιο κοντά προσεγγίζουμε την αντίστοιχη ιδιοτιμή.

Το ψευδοφάσμα μας θα παριστάνει “μικρά κυκλάκια” γύρω από τις ιδιοτιμές για $\varepsilon \rightarrow 0$.

Μπορούμε στο συγκεκριμένο παράδειγμα να βάλουμε για ε μια πολύ μικρή τιμή ($\varepsilon = 0.08$). Έχουμε

```
>> c=psegridA(A,[0.08],[-4,8],[-4,4],0.01,0.01)
```

```
number_of_total_grid_points =
```

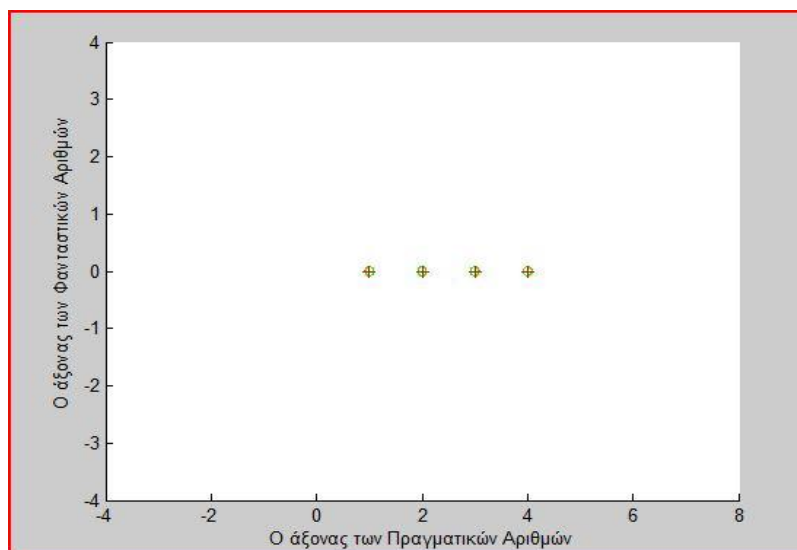
```
962001
```

```
number_of_iterations =
```

```
962001
```

```
Elapsed time is 38.187546 seconds.
```

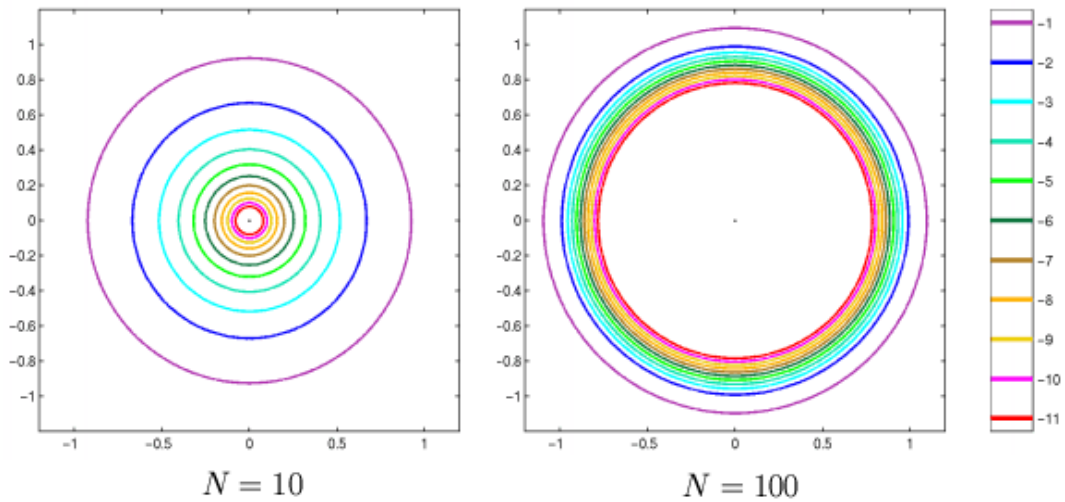
Ο χρόνος που απαιτείται για να ολοκληρωθεί η απεικόνιση 962001 σημείων είναι 38.187546 δευτερόλεπτα. Η γραφική μας είναι



Σχήμα 6: Τα ψευδοφάσματα του A για $\varepsilon = 0.08$.

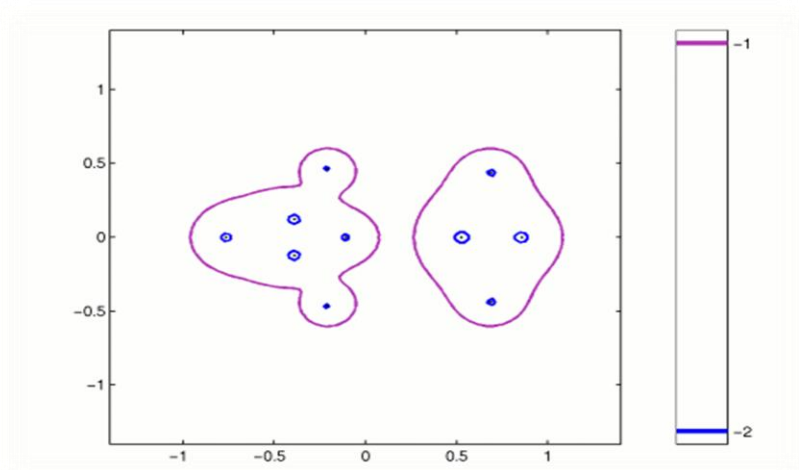
Πράγματι, διαπιστώνουμε την ύπαρξη των κύκλων. Το ψευδοφάσμα, επομένως, αναπαριστά κύκλους με κέντρο τις ιδιοτιμές.

2.3.2 Διάφορα γραφήματα ψευδοφάσματος τυχαίων πινάκων από την ιστοσελίδα Pseudospectra Gateway (<http://www.cs.ox.ac.uk/pseudospectra/>).

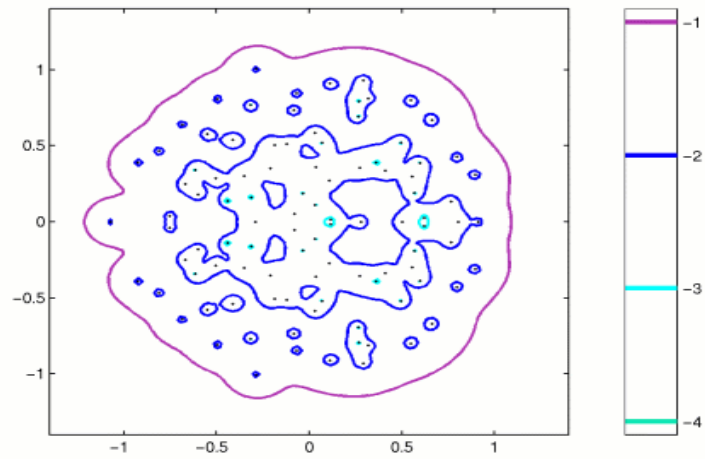


Σχήμα 7: Το φάσμα και το ε -ψευδοφάσμα για έναν πίνακα με διάσταση 10 (αριστερά) και διάσταση 100 (δεξιά). Να σημειώσουμε τον υψηλό βαθμό μη κανονικότητας στον μικρό πίνακα και πώς εκρήγνυται αυτή η μη κανονικότητα και η διάσταση αυξάνεται.

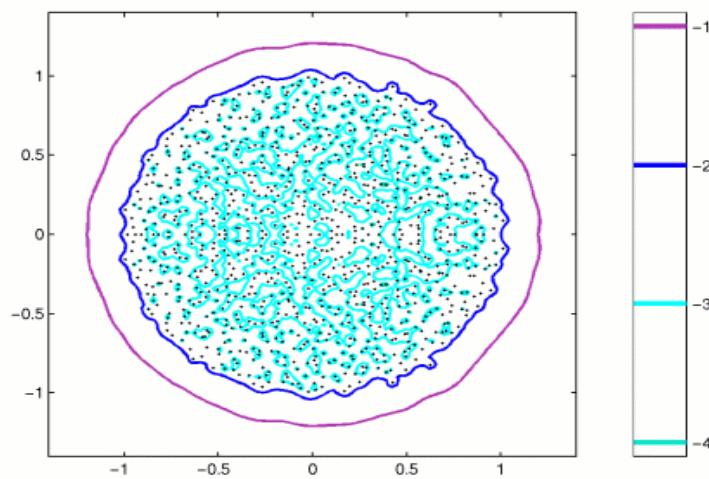
Οι φασματικές ιδιότητες των πυκνών μη συμμετρικών τυχαίων πινάκων έχουν μελετηθεί από τη δεκαετία του 1960. Εξετάζουμε τον πίνακα A , n -διαστάσεων, του οποίου τα στοιχεία αντλούνται από την κανονική κατανομή με διακύμανση n^{-1} .



Σχήμα 8: Πυκνοί τυχαίοι πίνακες με πραγματικές συμμετοχές, διάστασης $N = 10$.



Σχήμα 9: Πυκνός τυχαίος πίνακας με πραγματικές συμμετοχές, διάστασης $N = 100$.



Σχήμα 10: Πυκνός τυχαίος πίνακας με πραγματικές συμμετοχές, διάστασης $N = 1000$. Οι τιμές που αναγράφονται στην colorbar δείχνουν το $\log(\epsilon)$.

2.4 Άνω Φράγμα του Δείκτη Κατάστασης Ιδιοτιμών Πίνακα

Επειδή $\lambda_j I - A$ είναι ένας μοναδικός πίνακας, το $\|adj(\lambda_j I - A)\|_F$ στη (2.2.4) μπορεί να εκτιμηθεί με τη βοήθεια του ακόλουθου αποτελέσματος.

Θεώρημα 2.4: Αν C είναι κάθε $n \times n$ πίνακας, τότε

$$\|adjC\|_2 = (\rho_1 \rho_2 \dots \rho_{n-1})^{\frac{1}{2}}, \quad (2.4.1)$$

$$\|adjC\|_2 \leq \{(n-1)^{-1} \|C\|_F^2\}^{(n-1)/2}, \quad (2.4.2)$$

$$\|adjC\|_2 \leq \{1 + (n-1)^{-1}\}^{(n-1)/2} (d_1 d_2 \dots d_{n-1})^{\frac{1}{2}}, \quad (2.4.3)$$

όπου $\rho_1 \geq \rho_2 \geq \dots \geq \rho_n$ είναι οι ιδιοτιμές του C^*C και d_1, d_2, \dots, d_n είναι τα διαγώνια στοιχεία του C^*C αναδιατεταγμένα έτσι ώστε $d_1 \geq \dots \geq d_n$. Επιπλέον,

$$\|adjC\|_F = \|adjC\|_2, \quad \text{όταν } \det C = 0. \quad (2.4.4)$$

Απόδειξη: Οι ιδιοτιμές του $adj(C^*C)$ είναι $\prod_j \rho_i$, ($j = 1, 2, \dots, n$), όπου \prod_j δηλώνει το γινόμενο για όλα τα $i \neq j$ με $1 \leq i \leq n$. Δεδομένου ότι $\prod_n \rho_i$ είναι η μεγαλύτερη από αυτές τις ιδιοτιμές, η σχέση (2.2.6) θα μας δώσει $\|adjC\|_2^2 = \prod_n \rho_i$. Πράγματι, από τη σχέση (2.2.6) ισχύει

$$\|adjC\|_2^2 = \lambda_{\max}(adjC^* adjC) = \lambda_{\max}(adj(CC^*)) = \lambda_{\max}(adj(C^*C)) = \prod_n \rho_i.$$

Η σχέση $\|adjC\|_2^2 = \prod_n \rho_i$ αποδεικνύει τη σχέση (2.4.1).

Σημείωση: Αν είναι $\alpha_1, \alpha_2, \dots, \alpha_n$ θετικοί και

$$A = \frac{\alpha_1 + \alpha_2 + \dots + \alpha_n}{n} \quad (\text{αριθμητικός μέσος})$$

$$G = \sqrt[n]{\alpha_1 \alpha_2 \dots \alpha_n} \quad (\text{γεωμετρικός μέσος}),$$

τότε $G \leq A$ και $A \leq G^2$.

Συγκρίνοντας το γεωμετρικό μέσο $(\prod_n \rho_i)^{1/(n-1)}$ με τον αντίστοιχο αριθμητικό μέσο $(\rho_1 + \rho_2 + \dots + \rho_{n-1}) / (n-1)$ δείχνουμε ότι

$$(\prod_n \rho_i)^{1/(n-1)} \leq (\rho_1 + \rho_2 + \dots + \rho_{n-1}) / (n-1).$$

Άρα προκύπτει η (2.4.5)

$$(n-1) \|adjC\|_2^{2/(n-1)} \leq \rho_1 + \rho_2 + \dots + \rho_{n-1} \leq \rho_1 + \rho_2 + \dots + \rho_{n-1} + \rho_n = \text{trace}(C^*C).$$

Επειδή $\text{trace}(CC^*) = \text{trace}(C^*C) = \sum_{i,j=1}^n |c_{ij}|^2 = \|C\|_F^2$, έχουμε

$$(n-1) \|adjC\|_2^{2/(n-1)} \leq \|C\|_F^2.$$

Άρα

$$\|adjC\|_2 \leq \{(n-1)^{-1} \|C\|_F^2\}^{(n-1)/2}.$$

Με $Q = I$, η σχέση (2.2.7) θα μας δώσει

$$|\det R| \|adjC\|_2 = \|R adj(CR)\|_2 \leq \|R\|_2 \|adj(CR)\|_2. \quad (2.4.6)$$

Αντικαθιστώντας το C στην (2.4.5) με CR και χρησιμοποιώντας την (2.4.6) παίρνουμε

$$|\det R| \|adjC\|_2 \leq \|R\|_2 \{(n-1)^{-1} \text{trace}(R^* C^* CR)\}^{(n-1)/2}. \quad (2.4.7)$$

Εάν $d_1 \geq \dots \geq d_n > 0$, τότε όλα τα διαγώνια στοιχεία του πίνακα $R^* C^* CR$ μπορούν να γίνουν ίσα με 1 επιλέγοντας R να είναι ο διαγώνιος πίνακας με $d_1^{-1/2}, d_2^{-1/2}, \dots, d_n^{-1/2}$ να είναι οι ιδιοτιμές του.

Τότε $\text{trace}(R^* C^* CR) = n$, $\|R\|_2 = d_n^{-1/2}$, $\det R = (d_1 d_2 \dots d_n)^{-1/2}$ και έτσι η σχέση (2.4.7) ανάγεται στην (2.4.3).

Στις περιπτώσεις που μερικά από τα στοιχεία d_1, \dots, d_n είναι μηδέν η (2.4.3) ισχύει ακόμη, λόγω συνέχειας. Εάν $\det C = 0$, τότε $\rho_n = 0$ και $\prod_j \rho_i = 0$, για $j = 1, 2, \dots, n-1$. Πράγματι αν $\det C = 0$, τότε $\det(CC^*) = (\det C)(\det C^*) = 0$. Άρα $\rho_n = 0$ και $\prod_j \rho_i = 0$ για κάθε $j \neq n$, αφού το γινόμενο αυτό θα περιέχει το ρ_n που ισούται με 0. Όλες οι ιδιοτιμές του $adj(C^*C)$ είναι επομένως μηδέν, εκτός ίσως από την $\prod_n \rho_i$. Δηλαδή ή θα ισχύει $\prod_n \rho_i = 0$ (αν $\rho_{n-1} = 0$) ή θα ισχύει $\prod_n \rho_i \neq 0$ (αν $\rho_{n-1} \neq 0$). Ως εκ τούτου η σχέση (2.2.6) δίνει

$$\|adjC\|_F^2 = \text{trace } adj(C^*C) = \prod_n \rho_i = \|adjC\|_2^2.$$

Έτσι αποδείχθηκε η (2.4.4) και ολοκληρώθηκε η απόδειξη του Θεωρήματος 2.4. \square

Στο Θεώρημα 2.4, οι $d_1^2, d_2^2, \dots, d_n^2$ είναι οι νόρμες των στηλών του πίνακα C . Αλλά θα μπορούσαν επίσης να ληφθούν για να είναι οι νόρμες των γραμμών του C , γιατί η (2.4.1) δείχνει ότι $\|adjC\|_2 = \|adjC^*\|_2 = (\rho_1 \rho_2 \dots \rho_{n-1})^{1/2}$. (Οι στήλες πλέον γίνονται γραμμές). Στην ειδική περίπτωση που $\det C = 0$, η (2.4.2) και η (2.4.4) εντείνουν ένα αποτέλεσμα του Richter, ο οποίος δείχνει ότι η $\|adjC\|_2$ ποτέ δεν υπερβαίνει το $n^{(2-n)/2} \|C\|_F^{n-1}$.

2.5 Gram Ορίζουσα και Διακρίνουσα

Εάν όλες οι ιδιοτιμές του A είναι απλές, ορίζουμε $G = X^*X$, όπου X είναι ένας $n \times n$ πίνακας του οποίου οι στήλες είναι τα δεξιά ιδιοδιανύσματα u_1, u_2, \dots, u_n του A κανονικοποιημένα έτσι ώστε $\|u_j\|_2 = 1$ για όλα τα j . Τότε η *Gram ορίζουσα* $\det G$ προσδιορίζεται μονοσήμαντα.

Θεώρημα 2.5: Εάν όλες οι ιδιοτιμές του A είναι απλές, τότε

$$\left\{ \frac{1}{2} (\kappa_2 + \kappa_2^{-1}) \right\}^{-2} \geq \det G \geq \left\{ \frac{1}{2} (\kappa_2 + \kappa_2^{-1}) \right\}^{2-2n} \quad (2.5.1)$$

$$|s_j|^2 \geq \det G \geq |s_j|^{-2} |s_1 s_2 \dots s_n|^2 \quad (2.5.2)$$

όπου $|s_j|$ είναι το ελάχιστο των αριθμών $|s_1|, |s_2|, \dots, |s_n|$.

Απόδειξη: Επειδή οι στήλες του X έχουν μοναδιαία νόρμα, όλα τα διαγώνια στοιχεία του X^*X είναι ίσα με το 1 και επομένως $\text{trace} G = n$. Η πρώτη εξίσωση της (2.2.11) δείχνει ότι η (2.1.1) ικανοποιείται από την $P = X$. Ως εκ τούτου,

$$\kappa_2 \leq \|X\|_2 \|X^{-1}\|_2 = \|G\|_2^{\frac{1}{2}} \|G^{-1}\|_2^{\frac{1}{2}}. \quad (2.5.3)$$

Οι Davis, Haynsworth και Marcus έδειξαν ότι

$$\|G\|_2 \|G^{-1}\|_2 \leq [1 + (1 - D_1)^{\frac{1}{2}}] / [1 - (1 - D_1)^{\frac{1}{2}}],$$

όπου $D_1 = n^n (\text{trace} G)^{-n} \det G$.

Αντικαθιστώντας το $\|G\|_2 \|G^{-1}\|_2$ με κ_2^2 (από την (2.5.3)) και μετά λύνοντας ως προς D_1 , παίρνουμε

$$\kappa_2^2 \leq [1 + (1 - D_1)^{\frac{1}{2}}] / [1 - (1 - D_1)^{\frac{1}{2}}] \Rightarrow \kappa_2^2 [1 - (1 - D_1)^{\frac{1}{2}}] \leq 1 + (1 - D_1)^{\frac{1}{2}} \Rightarrow$$

$$\Rightarrow \kappa_2^2 - \kappa_2^2 (1 - D_1)^{\frac{1}{2}} \leq 1 + (1 - D_1)^{\frac{1}{2}} \Rightarrow \kappa_2^2 - 1 \leq \kappa_2^2 (1 - D_1)^{\frac{1}{2}} + (1 - D_1)^{\frac{1}{2}} \Rightarrow$$

$$\Rightarrow \kappa_2^2 - 1 \leq (\kappa_2^2 + 1) (1 - D_1)^{\frac{1}{2}} \Rightarrow \frac{\kappa_2^2 - 1}{\kappa_2^2 + 1} \leq (1 - D_1)^{\frac{1}{2}} \Rightarrow \left(\frac{\kappa_2^2 - 1}{\kappa_2^2 + 1} \right)^2 \leq (1 - D_1) \Rightarrow$$

$$\Rightarrow D_1 \leq 1 - \left(\frac{\kappa_2^2 - 1}{\kappa_2^2 + 1} \right)^2 \Rightarrow D_1 \leq \frac{(\kappa_2^2 + 1)^2 - (\kappa_2^2 - 1)^2}{(\kappa_2^2 + 1)^2} \Rightarrow D_1 \leq \frac{4 \kappa_2^2}{(\kappa_2^2 + 1)^2} \Rightarrow$$

$$\Rightarrow \frac{1}{\left(\frac{\kappa_2^2 + 1}{2\kappa_2} \right)^2} \geq D_1 \Rightarrow \frac{1}{\left(\frac{\kappa_2}{2} + \frac{1}{2\kappa_2} \right)^2} \geq D_1 \Rightarrow \left\{ \frac{1}{2} (\kappa_2 + \kappa_2^{-1}) \right\}^{-2} \geq D_1 \Rightarrow$$

$$\Rightarrow \left\{ \frac{1}{2} (\kappa_2 + \kappa_2^{-1}) \right\}^{-2} \geq D_1 = n^n (\text{trace} G)^{-n} \det G.$$

Αυτό αποδεικνύει την αριστερή πλευρά της (2.5.1), γιατί $\text{trace}G = n$. Από αυτό, η αριστερή πλευρά της (2.5.2) προκύπτει μέσω της (2.2.3). Εάν ο Y είναι ένας πίνακας στη δεύτερη εξίσωση του (2.2.11) και $H = (Y^T)^* Y^T$, τότε $\det H$ είναι η Gram ορίζουσα που αντιστοιχεί στον A^T . Επειδή οι A και A^T έχουν τους ίδιους Jordan αριθμούς κατάστασης, η αριστερή πλευρά της (2.5.1) δίνει

$$\left\{ \frac{1}{2} (\kappa_2 + \kappa_2^{-1}) \right\}^{-2} \geq \det H. \quad (2.5.4)$$

Τότε η (2.2.3) δίνει $|s_j|^2 \geq \det H$. Επίσης η (2.2.12) μας δίνει

$$(\det G) \det H = |\det X|^2 |\det Y|^2 = |\det S|^2 = |s_1 s_2 \dots s_n|^2.$$

Άρα $(\det G) |s_j|^2 \geq (\det G) \det H = |s_1 s_2 \dots s_n|^2$. Αυτό αποδεικνύει τη δεξιά πλευρά της (2.5.2) και από αυτό η δεξιά πλευρά της (2.5.1) προκύπτει μέσω της (2.2.3). Συγκεκριμένα,

$$\det G \geq |s_j|^2 |s_1 s_2 \dots s_n|^2 = |s_1 s_2 \dots s_{j-1} s_{j+1} \dots s_n|^2 \geq \left\{ \frac{1}{2} (\kappa_2 + \kappa_2^{-1}) \right\}^{2-2n}.$$

Αυτό ολοκληρώνει την απόδειξη του Θεωρήματος 2.5. Η δεξιά πλευρά της (2.5.1) είναι στενά συνδεδεμένη με το αποτέλεσμα του Marcus από το οποίο θα μπορούσε επίσης να εξαχθεί. \square

Όλες οι ιδιοτιμές του A είναι απλές αν $\Delta \neq 0$, όπου

$$\Delta = \prod_{i \neq j} (\lambda_i - \lambda_j)^2. \quad (2.5.5)$$

Αυτό το Δ είναι η διακρίνουσα του χαρακτηριστικού πολυωνύμου

$$\det(\lambda I - A) = \lambda^n + \alpha_1 \lambda^{n-1} + \alpha_2 \lambda^{n-2} + \dots + \alpha_n.$$

Επειδή το Δ μπορεί να εκφραστεί σαν ένα πολυώνυμο σε $\alpha_1, \alpha_2, \dots, \alpha_n$, η συνθήκη $\Delta \neq 0$ είναι εύκολο να ελεγχθεί όταν το n είναι μικρό. Για παράδειγμα,

$$\Delta = \alpha_1^2 \alpha_2^2 - 4\alpha_1^3 \alpha_3 + 18\alpha_1 \alpha_2 \alpha_3 - 4\alpha_2^3 - 27\alpha_3^2, \text{ όταν } n = 3.$$

Εάν $n = 4$, τότε το 27Δ είναι ίσο με

$$4(12\alpha_4 - 3\alpha_1 \alpha_3 + \alpha_2^2)^2 - (72\alpha_2 \alpha_4 + 9\alpha_1 \alpha_2 \alpha_3 - 27\alpha_3^2 - 27\alpha_1^2 \alpha_4 - 2\alpha_2^3)^2.$$

Το ακόλουθο αποτέλεσμα δίνει ένα άνω φράγμα για το κ_2 το οποίο, αν και όχι ολοκληρωμένο, δεν απαιτεί τη γνώση των ιδιοτιμών του A .

Θεώρημα 2.6: Εάν $\Delta \neq 0$ τότε

$$(\kappa_2 + \kappa_2^{-1})^4 \leq 2^4 |\Delta|^{-2} \{ (n-1)^{-1} 2 (\|A\|_F^2 - n^{-1} |\text{trace}A|^2) \}^{n(n-1)}. \quad (2.5.6)$$

Απόδειξη: Όπως προκύπτει από την (2.5.1) και την (2.5.4) έχουμε

$$\left\{\frac{1}{2}(\kappa_2 + \kappa_2^{-1})\right\}^{-4} \geq (\det G) \det H = |s_1 s_2 \dots s_n|^2.$$

Αντικαθιστώντας το $|s_j|$ από τη (2.2.4) και το Δ από τη (2.5.5), παίρνουμε

$$\left\{\frac{1}{2}(\kappa_2 + \kappa_2^{-1})\right\}^4 \leq |\Delta|^{-2} \prod_{j=1}^n \|\text{adj}(\lambda_j I - A)\|_F^2.$$

Αντικαθιστώντας το $\|\text{adj}(\lambda_j I - A)\|_F$ από (2.4.4) και (2.4.2), παίρνουμε

$$\left\{\frac{1}{2}(\kappa_2 + \kappa_2^{-1})\right\}^4 \leq (n-1)^{n(n-1)} |\Delta|^{-2} \prod_{j=1}^n \|\lambda_j I - A\|_F^{2(n-1)}. \quad (2.5.7)$$

Η σύγκριση των γεωμετρικών και αριθμητικών μέσων μας δίνει

$$\begin{aligned} \prod_{j=1}^n \|\lambda_j I - A\|_F^{2/n} &\leq \frac{1}{n} \sum_{j=1}^n \|\lambda_j I - A\|_F^2 = \frac{1}{n} \sum_{j=1}^n \text{trace}[(\lambda_j I - A)^* (\lambda_j I - A)] \\ &= \frac{1}{n} \sum_{j=1}^n \text{trace}[\lambda_j^2 I - \lambda_j I A - \lambda_j I A^* + A^* A] \\ &= \frac{1}{n} \sum_{j=1}^n \{n |\lambda_j|^2 - \text{trace}[\lambda_j I A + \lambda_j I A^*] + \text{trace}[A^* A]\} \\ &= \frac{1}{n} \sum_{j=1}^n \{n |\lambda_j|^2 + \|A\|_F^2 - 2\text{Re}(\bar{\lambda}_j \text{trace} A)\} \\ &\leq 2 (\|A\|_F^2 - n^{-1} |\text{trace} A|^2). \end{aligned}$$

Αντικαθιστώντας αυτό στην (2.5.7), παίρνουμε την (2.5.6). □

2.6 Κάτω Φράγμα του Δείκτη κ_2

Από τη σχέση (2.1.1) έχουμε $\|A\|_2 \leq \|P^{-1}\|_2 \|J\|_2 \|P\|_2$. Ως εκ τούτου,

$$\kappa_2 \geq \|A\|_2 / \|J\|_2. \quad (2.6.1)$$

Επιπλέον, η (2.1.1) δίνει $f(A) = P f(J) P^{-1}$, όπου $f(z)$ είναι κάθε ρητή συνάρτηση, η οποία είναι ομαλή στις ιδιοτιμές του A . Περαιτέρω τα κάτω φράγματα των κ_2 μπορούν ως εκ τούτου να ληφθούν με την αντικατάσταση των A και J στην (2.6.1) με $f(A)$ και $f(J)$, αντίστοιχα. Το ακόλουθο αποτέλεσμα μας δίνει ένα κάτω φράγμα για το κ_2 , το οποίο δεν απαιτεί γνώση των ιδιοτιμών του A .

Θεώρημα 2.7: Εάν ο J στην (2.1.1) είναι ένας διαγώνιος πίνακας, τότε

$$\kappa_2^4 \geq 1 + \frac{1}{2} \{\|A^* A - A A^*\|_F / \|A^2\|_F\}^2. \quad (2.6.2)$$

Η απόδειξη του Θεωρήματος 2.7 χρησιμοποιεί το ακόλουθο λήμμα το οποίο μπορεί να εξαχθεί από τα πιο γενικά συμπεράσματα του Bauer, του Stoer και του Witzgall. Αλλά το αποδεικνύουμε για λόγους πληρότητας.

Λήμμα: Ισχύουν

$$\|BC\|_F \leq \|B\|_2 \|C\|_F \text{ και } \|BC\|_F \leq \|C\|_2 \|B\|_F.$$

Απόδειξη του Λήμματος: Εάν το διάνυσμα c_i είναι η $i^{\text{η}}$ στήλη του C , τότε το Bc_i είναι η $i^{\text{η}}$ στήλη του BC και

$$\|BC\|_F^2 = \sum_{i=1}^n \|Bc_i\|_F^2 \leq \sum_{i=1}^n \|B\|_2^2 \|c_i\|_F^2 = \|B\|_2^2 \|C\|_F^2.$$

(Η απόδειξη μπορεί να γίνει και με τη χρήση της μεθόδου *SVD*.)

Από αυτό η δεύτερη ανισότητα εξάγεται ως εξής

$$\|BC\|_F = \|C^*B^*\|_F \leq \|C^*\|_2 \|B^*\|_F = \|C\|_2 \|B\|_F. \quad \square$$

Απόδειξη του Θεωρήματος 2.7: Από τη σχέση (2.1.1) και από το Λήμμα παίρνουμε

$$\|A^*A\|_F = \|(P^*)^{-1} (J^*P^*P J) P^{-1}\|_F \leq \|P^{-1}\|_2^2 \|J^*P^*P J\|_F.$$

Βάζουμε $C = PJ$ στη $\|C^*C\|_F = \|CC^*\|_F$. Τότε το λήμμα μας δίνει

$$\|J^*P^*P J\|_F = \|PJ J^*P^*\|_F \leq \|P\|_2^2 \|J^*J\|_F.$$

Επειδή ο J είναι διαγώνιος, έχουμε

$$\|JJ^*\|_F = \|J^2\|_F \leq \|A^2\|_F.$$

Ως εκ τούτου,

$$\|P^{-1}\|_2^2 \|P\|_2^2 \geq \|JJ^*\|_F^{-1} \|A^*A\|_F \geq \|A^2\|_F^{-1} \|A^*A\|_F,$$

για όλα τα P που ικανοποιούν την (2.1.1). Έτσι λοιπόν προκύπτει

$$\kappa_2^2 \geq \|A^2\|_F^{-1} \|A^*A\|_F,$$

από το οποίο το Θεώρημα 2.7 ακολουθεί με τη χρήση της ταυτότητας

$$\|A^*A - AA^*\|_F^2 = 2\|A^*A\|_F^2 - 2\|A^2\|_F^2.$$

Το παραπάνω είναι εύκολο να δειχθεί από την επέκταση $\text{trace}(A^*A - AA^*)^2$. (Γνωρίζουμε ότι $\|A\|_F^2 = \sum |\alpha_{ij}|^2$.)

Έχουμε, επομένως,

$$\begin{aligned} \|A^*A - AA^*\|_F^2 &= \text{trace}[(A^*A - AA^*)^* (A^*A - AA^*)] = \text{trace}[(A^*A - AA^*)^2] \\ &= \text{trace}[(A^*A)^2 - 2(A^*A)(AA^*) + (AA^*)^2] \\ &= \text{trace}[(A^*A)^2 + (AA^*)^2] - 2\text{trace}[(A^*A)(AA^*)] \\ &= 2\|A^*A\|_F^2 - 2\|A^2\|_F^2. \end{aligned} \quad \square$$

2.7 Παρατηρήσεις σχετικά με τα Αποτελέσματα

Παράδειγμα 1.

Για να δοκιμάσουμε κάποιες από τις παρατηρήσεις που λήφθηκαν παραπάνω στο Κεφάλαιο 2 θεωρούμε τον πίνακα

$$A = \begin{pmatrix} u & 0 & v \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix}$$

στον οποίο η διακρίνουσα $\Delta = u^2(1+u)^2$ και η απομάκρυνση από την κανονικότητα μπορούν να μεταβάλλονται ανεξάρτητα. Οι δείκτες κατάστασης εύκολα υπολογίζονται.

Εάν $\lambda_1 = u$, $\lambda_2 = 0$, $\lambda_3 = -1$ και $\Delta \neq 0$, τότε

$$\kappa_A(\lambda_1) = \kappa_A(\lambda_3) = [1+(1+u)^{-2}v^2], \quad \kappa_A(\lambda_2) = 1,$$

$$\kappa_2 = \{|v| + [v^2 + (1+u)^2]^{1/2}\} |1+u|^{-1} \quad \text{και} \quad \det G = \{1+(1+u)^{-2}v^2\}^{-1}.$$

Και οι δύο πλευρές της σχέσης (2.5.2) “μειώνουν” την ισότητα όπως κάνει επίσης και η αριστερή πλευρά του (2.5.1). Οι εκτιμήσεις (2.5.6), (2.6.1), (2.6.2) δίνουν

$$(\kappa_2 + \kappa_2^{-1})^2 \leq 4u^{-2} (1+u)^2 \{1+u^2+v^2-3^{-1}(1-u)^2\}^3, \quad (2.5.6)$$

$$\kappa_2^2 \geq \{1+u^2+v^2+[(1+u^2+v^2)^2-4u^2]^{1/2}\} / 2 \max(1, u^2), \quad (2.6.1)$$

$$\kappa_2^4 \geq [1+2v^2+(u^2+v^2)^2] / [1+u^4+v^2(1-u)^2]. \quad (2.6.2)$$

Καθώς $u \rightarrow -1$ και κρατώντας το v σταθερό, οι δεξιές πλευρές του (2.6.1) και (2.6.2) παραμένουν πεπερασμένες καθώς $\kappa_2 \rightarrow +\infty$. Επίσης, το κ_2 παραμένει οριοθετημένο, όταν $u \rightarrow 0$, καθώς η δεξιά πλευρά του (2.5.6) τείνει στο ∞ . Φαίνεται ότι όλες οι εκτιμήσεις αυτές ενδέχεται να μην έχουν “υποστεί” επεξεργασία, όταν η διακρίνουσα Δ είναι μικρή ή η απομάκρυνση από την κανονικότητα $|v|$ είναι μεγάλη, όπως θα δούμε παρακάτω.

Για να τονίσουμε τις δυσκολίες που εμπλέκονται στην εκτίμηση του κ_2 , όταν ο A έχει επαναλαμβανόμενες ιδιοτιμές, εξετάζουμε την υπόθεση

$$A = \begin{pmatrix} \lambda & \alpha \\ 0 & \lambda \end{pmatrix}, \quad J = \begin{pmatrix} \lambda & 1 \\ 0 & \lambda \end{pmatrix}, \quad P = \xi \begin{pmatrix} 1 & \eta \\ 0 & \alpha \end{pmatrix}, \quad \text{όπου } \alpha \neq 0.$$

Για αυθαίρετα $\xi \neq 0$, η , ο πίνακας P είναι η γενική λύση του (2.1.1). Επειδή

$$\|P\|_F^2 \|P^{-1}\|_F^2 = (1+\eta^2+\alpha^2) (1+\alpha^{-2}+\alpha^{-2}\eta^2),$$

συνεπάγεται ότι $\kappa_F = (1 + a^2) / |a|$ και επομένως $\kappa_F \rightarrow \infty$, καθώς $a \rightarrow 0$. Αυτό συμβαίνει, γιατί ο κ_F είναι εξαιρετικά ευαίσθητος σε μικρές διακυμάνσεις της παραμέτρου a . Έτσι, λοιπόν, συνεπάγεται ότι στην περίπτωση των πολλαπλών ιδιοτιμών, δεν μπορούμε να “παράγουμε” ένα γενικό άνω φράγμα του κ_F εκτός αν έχουμε κάποιες πολύπλοκες ιδιότητες του πίνακα και των ιδιοτιμών. Ωστόσο, ακόμη και αν κάτι τέτοιο είχε “παραχθεί” μπορεί να μην είναι πολύ χρήσιμο στην αριθμητική ανάλυση εξαιτίας της πρακτικής δυσκολίας διάκρισης μιας πολλαπλής ρίζας από ένα ζεύγος διακριτών ριζών, οι οποίες βρίσκονται κοντά μεταξύ τους.

ΚΕΦΑΛΑΙΟ 3: Φράγματα για Πίνακες Υψωμένους σε Δύναμη και για Αντίστροφους Πίνακες, Φασματική Μεταβολή και Αριθμητικά Πεδία για Μη-Κανονικούς Πίνακες.

Θα θεωρούμε από εδώ και στο εξής ότι ένας πίνακας θα σημαίνει ένας τετραγωνικός πίνακας τάξης n με μιγαδικά στοιχεία και ένα διάνυσμα θα υποδηλώνει ένα διάνυσμα στήλη με n μιγαδικές συνιστώσες. Θα εξετάσουμε τα ακόλουθα προβλήματα της γραμμικής άλγεβρας που αφορούν

- (i) Την εκτίμηση των νορμών πινάκων A_n , $n = 1, 2, 3, \dots$, σε σχέση με τη $\rho(A)$.
- (ii) Την εκτίμηση του σφάλματος $\tilde{x} - A^{-1}b$ από μια προσεγγιστική λύση \tilde{x} της εξίσωσης $Ax = b$ σε σχέση με το υπόλοιπο $r = A\tilde{x} - b$ και με τη φασματική ακτίνα $\rho(A^{-1})$.
- (iii) Την εκτίμηση της απόστασης του φάσματος ενός πίνακα B από το φάσμα ενός πίνακα A σε σχέση με τη νόρμα του $B - A$.
- (iv) Την εύρεση φράγματος για το αριθμητικό πεδίο ενός πίνακα A σε σχέση με το φάσμα του A . Εδώ το αριθμητικό πεδίο, $F(A)$, ορίζεται ως το σύνολο των μιγαδικών αριθμών

$$\xi = \frac{x^*Ax}{x^*x},$$

όπου το x “διατρέχει” όλα τα μη-μηδενικά διανύσματα. (Το $*$ υποδηλώνει τον ανάστροφο συζυγή ενός διανύσματος ή ενός πίνακα.)

Οι λύσεις στα παραπάνω προβλήματα είναι κλασικές (και μάλιστα τετριμμένες σε μερικές περιπτώσεις), εάν ο A είναι κανονικός, δηλαδή αν ισχύει $AA^* = A^*A$. Επίσης, έχουν κατασκευαστεί και λύσεις για μη-κανονικό πίνακα A , αλλά με λιγότερο ικανοποιητικά αποτελέσματα. Μερικά από τα φράγματα που δίνονται εξαρτώνται από τη γνώση ενός πίνακα S στην αναπαράσταση $A = SJS^{-1}$, όπου J είναι η Jordan κανονική μορφή. Τα φράγματα που δίνονται εδώ εξαρτώνται, στο μέγιστο, από τις ιδιοτιμές του A . Η γνώση της Jordan κανονικής μορφής δεν απαιτείται. Οι εκτιμήσεις μας προσεγγίζουν τις κλασικές εκτιμήσεις για A κανονικό πίνακα. Η επιμονή μας στο να μη χρησιμοποιήσουμε τη Jordan μορφή βασίζεται εν μέρει σε λόγους υπολογιστικής ευκολίας, αλλά και από το γεγονός ότι η Jordan μορφή είναι μια ασυνεχής συνάρτηση στο χώρο των πινάκων. Επομένως είναι δύσκολα κατάλληλη για τους σκοπούς του υπολογισμού.

Το κύριο εργαλείο στην έρευνά μας είναι μια αριθμητικά-εκτιμώμενη συνάρτηση στο χώρο των πινάκων που χρησιμεύει ως ένα μέτρο απομάκρυνσης από την κανονικότητα του πίνακα. Αυτή η συνάρτηση θα οριστεί στη συνέχεια. Θα πρέπει, λοιπόν, να “παράγουμε” ένα φράγμα για αυτή τη συνάρτηση με βάση μια ρητή συνάρτηση των στοιχείων του πίνακα. Στα επόμενα τμήματα το μέρος της μη-

κανονικότητας θα εφαρμοστεί για τη λύση των τεσσάρων προβλημάτων που αναφέρθηκαν παραπάνω.

Σημείωση: Ο Wielandt είχε ορίσει ένα μέτρο για τη μη-κανονικότητα ενός πίνακα. Ωστόσο, το μέτρο του μπορεί να εφαρμοστεί μόνο σε πίνακες που είναι όμοιοι με έναν διαγώνιο πίνακα. Για να βρει, όμως, ένα άμεσο φράγμα γι' αυτό απαιτείται η γνώση ενός πίνακα S , ο οποίος πραγματοποιεί τη διαγωνοποίηση.

3.1 Ένα Μέτρο για τη Μη-Κανονικότητα κατά Henrici

3.1.1 *Μερικά εισαγωγικά πάνω στις νόρμες.* Οι ορισμοί που ακολουθούν βασίζονται στην επιλογή της νόρμας που χρησιμοποιούμε.

- Μια νόρμα πίνακα $\|\cdot\|$ καλείται *συμβατή (compatible)* με μία νόρμα διάνυσμα $\|\cdot\|$, εάν $\|Ax\| \leq \|A\| \|x\|$ για όλους τους πίνακες A και τα διανύσματα x . Μια επαγόμενη νόρμα είναι πάντα συμβατή με τη νόρμα διάνυσμα που το ορίζει.
- Μια νόρμα πίνακα $\|\cdot\|$ θα καλείται *ορθομοναδιαία αναλλοίωτη (unitarily invariant)*, εάν $\|UAU^*\| = \|A\|$ για όλους τους A και τους ορθομοναδιαίους U . Η νόρμα Frobenius και η φασματική νόρμα είναι ορθομοναδιαία αναλλοίωτες.
- Μια επαγόμενη νόρμα $\|\cdot\|$ καλείται *προσανατολισμένη κατά άξονα νόρμα (axis-oriented norm)*, εάν $\|D\| = \max_{1 \leq i \leq n} d_{ii} = \rho(D)$ για κάθε διαγώνιο πίνακα $D = (d_{ij})$. Είναι προφανές ότι οι επαγόμενες νόρμες $\|\cdot\|_1$, $\|\cdot\|_2$ και $\|\cdot\|_\infty$ είναι προσανατολισμένες κατά άξονα.
- Μια νόρμα $\|\cdot\|_a$ λέμε ότι είναι *μεγαλύτερη* μιας άλλης νόρμας $\|\cdot\|_b$ εάν $\|A\|_a \geq \|A\|_b$ για όλους τους πίνακες A . Ξεκάθαρα, η $\|A\|_1$ είναι η μεγαλύτερη από: τη $\|A\|_\infty$, τη $\|A\|_1$, τη νόρμα Frobenius και τη φασματική νόρμα. Επίσης, η νόρμα Frobenius είναι μεγαλύτερη από τη φασματική νόρμα.

Θα χρειαστούμε τις ακόλουθες συνέπειες από τις οριζόμενες ιδιότητες (i)-(iv) μιας νόρμας πινάκων.

(I) Έστω $\|\cdot\|$ μια νόρμα πινάκων στο $C^{n \times n}$. Τότε για κάθε $A \in C^{n \times n}$, ισχύει

$$\rho(A) \leq \|A\|, \quad (3.1.1)$$

όπου $\rho(A)$ υποδηλώνουμε τη φασματική ακτίνα του 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 $x_0 \in C^n$ του A που αντιστοιχεί στην ιδιοτιμή λ_0 και τον $n \times n$ πίνακα $X_0 = [x_0 \ x_0 \ \dots \ x_0]$ (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, $\rho(A) = |\lambda_0| \leq \|A\|$. □

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

(II) Έστω $\|\cdot\|_a$ και $\|\cdot\|_b$ δύο οποιεσδήποτε νόρμες πίνακα. Τότε υπάρχει μια σταθερά M_{ab} που εξαρτάται μόνο από αυτές τις δύο νόρμες, τέτοια ώστε

$$\|A\|_a \leq M_{ab} \|A\|_b \quad (3.1.2)$$

για όλους τους πίνακες A .

Υπόδειξη: Για την απόδειξή του θα βασιστούμε στην απόδειξη του παρακάτω θεωρήματος.

Υπάρχουν θετικοί (πεπερασμένοι και σταθεροί) αριθμοί $\mu, M \in \mathbb{R}$ τέτοιοι ώστε $\mu \|x\|_a \leq \|x\|_b \leq M \|x\|_a, \forall x \in \mathbb{C}^n$, όπου $\|\cdot\|_a, \|\cdot\|_b$ δύο τυχαίες νόρμες. \square

Για παράδειγμα, για κάθε πίνακα $A \in \mathbb{C}^{n \times n}$, ισχύουν οι ανισότητες:

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

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

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

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

και

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

Η ιδιότητα (II) είναι χρήσιμη, γιατί μας επιτρέπει συχνά να “μειώσουμε” την έρευνα της γενικής ιδιότητας των νορμών για τη μελέτη μιας ειδικής νόρμας.

3.1.2 Η απομάκρυνση από την κανονικότητα ως προς τη $\|\cdot\|$.

Θεώρημα του Schur: Έστω ένας τυχαίος πίνακας $A = (a_{ij}) \in \mathbb{C}^{n \times n}$. Τότε υπάρχουν ένας ορθομοναδιαίος πίνακας $U \in \mathbb{C}^{n \times n}$ και ένας άνω τριγωνικός πίνακας $T = (t_{ij}) \in \mathbb{C}^{n \times n}$ με διαγώνια στοιχεία τις ιδιοτιμές του A (λαμβάνοντας υπόψη και τις πολλαπλότητες), τέτοιοι ώστε

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

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

$$W_1 = [x_1 \ w_2 \ w_3 \ \dots \ w_n] = [x_1 \ \widehat{W}_1] \in \mathbb{C}^{n \times n},$$

όπου $\widehat{W}_1 = [w_2 \ w_3 \ \dots \ w_n] \in \mathbb{C}^{n \times (n-1)}$.

(Υπενθύμιση: Μια βάση $\{v_1, \dots, v_n\}$ ενός διανυσματικού χώρου V ονομάζεται ορθοκανονική, όταν τα διανύσματα είναι ανά δύο κάθετα (ορθογώνια) μεταξύ τους, δηλαδή $v_i \cdot v_j = 0$ με $i \neq j$ και καθένα απ' αυτά έχει μέτρο ίσο με 1.)

Αφού $w_2^* x_1 = w_3^* x_1 = \dots = w_n^* x_1 = 0$, έπεται ότι $\widehat{W}_1^* x_1 = 0$. Κατά συνέπεια, με απλές πράξεις παρατηρούμε ότι

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

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

$$W_2 = [x_2 \ \widehat{W}_2] \in \mathbb{C}^{(n-1) \times (n-1)} \quad (\text{όπου } \widehat{W}_2 \in \mathbb{C}^{(n-1) \times (n-2)}),$$

τέτοιον ώστε

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

Συνεπώς,

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

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

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

τέτοιους ώστε ο πίνακας

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

να ικανοποιεί τη σχέση

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

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

Ο πίνακας T καλείται μια *Schur* τριγωνική μορφή του A . Γενικά, ο T δεν προσδιορίζεται μοναδικά για ένα δεδομένο A . Θέτουμε

$$T = D + M,$$

όπου ο D υποδηλώνει το διαγώνιο πίνακα του οποίου η κύρια διαγώνιος συμπίπτει με αυτή της T (και έτσι αποτελείται από τις ιδιοτιμές του A). Προκύπτει ότι ο $M = (m_{ij})$ έχει μη-μηδενικά στοιχεία μόνο στα δεξιά της κύριας διαγωνίου (είναι, δηλαδή, γνήσια άνω τριγωνικός) και έτσι τότε

$$M^r = 0, \quad r \geq n. \quad (3.1.3)$$

Ορίζουμε την απομάκρυνση κατά *Henrici* του A από την κανονικότητα ως προς μια νόρμα πινάκων $\|\circ\|$ ως εξής

$$\delta_{\|\circ\|}(A) = \inf \{\|M\| : U^*AU = D + M, U^*U = UU^* = I_n\}, \quad (3.1.4)$$

όπου το \inf λαμβάνεται σε σχέση με όλα τα M τα οποία μπορούν να εμφανιστούν στη *Schur* τριγωνική μορφή.

Εύκολα βλέπουμε ότι η $\|\circ\|_F$ είναι ορθομοναδιαία αναλλοίωτη.

Τότε θα ισχύει $\|A\|_F^2 = \|UTU^*\|_F^2 = \|T\|_F^2 = \|D + M\|_F^2 = \|D\|_F^2 + \|M\|_F^2$. Άρα προκύπτει ότι η σχέση

$$\|M\|_F = \{\|A\|_F^2 - \sum_{i=1}^n |\lambda_i|^2\}^{\frac{1}{2}} = \{\sum_{i,j=1}^n |a_{ij}|^2 - \sum_{i=1}^n |\lambda_i|^2\}^{\frac{1}{2}}, \quad (3.1.5)$$

όπου λ_i , $i = 1, 2, \dots, n$ ιδιοτιμές του A (όχι κατ' ανάγκη διακεκριμένες), είναι ανεξάρτητη από την ειδική επιλογή της *Schur* τριγωνικής μορφής.

Χρησιμοποιώντας τις σχέσεις (3.1.4) και (3.1.5) παρατηρούμε ότι

$$\delta_{\|\circ\|_F}(A) = \sqrt{\|A\|_F^2 - \sum_{i=1}^n |\lambda_i|^2}. \quad (3.1.6)$$

Πρόταση: Ο παραπάνω πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν $\|M\|_F = 0$.

Απόδειξη: Θα δείξουμε πρώτα το ευθύ. Έστω ότι ο πίνακας A είναι κανονικός, δηλαδή ισχύει $A^*A = AA^*$. Επιπλέον, με βάση το Θεώρημα *Schur* ο A γράφεται στη

μορφή $A = U T U^*$. Επίσης ο U είναι ορθομοναδιαίος πίνακας. Άρα $U^* U = U U^* = I_n$. Έτσι προκύπτει

$$\begin{aligned} A^* A = A A^* &\Leftrightarrow (U T U^*)^* U T U^* = U T U^* (U T U^*)^* \Leftrightarrow \\ &\Leftrightarrow U T^* U^* U T U^* = U T U^* U T^* U^* \Leftrightarrow U T^* I_n T U^* = U T I_n T^* U^* \Leftrightarrow \\ &\Leftrightarrow U T^* T U^* = U T T^* U^* \Leftrightarrow U^1 U T^* T U^* (U^*)^{-1} = U^1 U T T^* U^* (U^*)^{-1} \Leftrightarrow \\ &\Leftrightarrow I_n T^* T I_n = I_n T T^* I_n \Leftrightarrow T^* T = T T^*. \end{aligned}$$

Άρα ο T είναι και αυτός κανονικός πίνακας. Επιπλέον ο T είναι άνω τριγωνικός και γράφεται στη μορφή

$$T = D + M,$$

όπου ο D υποδηλώνει το διαγώνιο πίνακα και ο M έχει μη-μηδενικά στοιχεία μόνο στα δεξιά της κύριας διαγωνίου.

Η μόνη, όμως, περίπτωση στην οποία ισχύει $T^* T = T T^*$ είναι τα στοιχεία στα δεξιά της κύριας διαγωνίου να είναι μηδέν, αφού και τα στοιχεία στα αριστερά της κύριας διαγωνίου είναι μηδέν, λόγω του γεγονότος ότι ο T είναι άνω τριγωνικός. Άρα $M = 0$. Επομένως, $\|M\|_F = 0$.

Αντίστροφα, αν $\|M\|_F = 0$, δηλαδή αν $M = 0$, τότε $T = D$. Τότε ο T έχει στοιχεία μόνο στην κύρια διαγώνιο και άρα είναι κανονικός. Ακολουθώντας την αντίστροφη διαδικασία στις παραπάνω ισοδυναμίες προκύπτει ότι και ο A είναι κανονικός. \square

Δεδομένου ότι ο A είναι κανονικός αν και μόνο αν $\|M\|_F = 0$, έπεται ότι η σχέση $\delta_{\|\cdot\|_F}(A) = \|M\|_F$ είναι μηδέν αν και μόνο αν ο A είναι κανονικός πίνακας. Από την ιδιότητα (II) παραπάνω, αυτό αληθεύει για $\delta_{\|\cdot\|}(A)$, όπου $\|\cdot\|$ είναι οποιαδήποτε νόρμα.

3.1.3 Ένα φράγμα για το $\delta_{\|\cdot\|_F}(A)$. Σ' αυτήν την παράγραφο θα “κατασκευάσουμε” ένα φράγμα για το $\delta_{\|\cdot\|_F}(A)$, το οποίο μπορεί να υπολογιστεί με ένα στοιχειώδη τρόπο από τα στοιχεία του A και το οποίο τείνει στο μηδέν για A κανονικό πίνακα. Από την ιδιότητα (II) παραπάνω μπορούμε να κατασκευάσουμε φράγματα για το $\delta_{\|\cdot\|}(A)$ για κάθε νόρμα $\|\cdot\|$.

Θεώρημα 3.1: Για έναν αυθαίρετο πίνακα A έχουμε

$$\|M\|_F = \delta_{\|\cdot\|_F}(A) \leq \sqrt[4]{\frac{n^3-n}{12}} \sqrt{\|A^* A - A A^*\|_F}. \quad (3.1.7)$$

Η ισότητα ισχύει στην (3.1.7) αν και μόνο αν ο A είναι ορθομοναδιαία παρόμοιος με έναν πίνακα της μορφής

$$\begin{pmatrix} \lambda & \alpha_1 & & & 0 \\ & \lambda & \alpha_2 & & \\ & & \lambda & \ddots & \\ & & & \ddots & \lambda & \alpha_{n-1} \\ 0 & & & & & \lambda \end{pmatrix},$$

όπου λ είναι μια μιγαδική σταθερά και $|\alpha_k|^2 = \alpha k(n-k)$, $k = 1, 2, \dots, n-1$, για κάποια $\alpha \geq 0$.

Παρατήρηση: Το παραπάνω φράγμα στη σχέση (3.1.7) διαπιστώνουμε ότι δεν είναι συναρτήσε των ιδιοτιμών.

Απόδειξη: Έστω $T = D + M$, όπου ο πίνακας T είναι τριγωνική μορφή Schur του A , δηλαδή υπάρχει ορθομοναδιαίος πίνακας U ώστε $A = UTU^*$ και ο πίνακας M είναι γνήσια άνω τριγωνικός. Έστω ακόμα ο πίνακας $\Gamma = (\gamma_{ij})$, τέτοιος ώστε $\Gamma = T^*T - TT^*$. Αρχικά θα δείξουμε ότι

$$\|M\|_F^2 \leq \gamma_{22} + 2\gamma_{33} + \dots + (n-1)\gamma_{nn}. \quad (3.1.8)$$

Πράγματι, παρατηρούμε ότι για $M = (m_{ij})$, έχουμε

$$\gamma_{ii} = \sum_{k<i} |m_{ki}|^2 - \sum_{k>i} |m_{ik}|^2, \quad i = 1, 2, \dots, n. \quad (3.1.9)$$

Θα αποδείξουμε τη σχέση (3.1.8) με επαγωγή. Προφανώς, η ανισότητα (3.1.8) είναι αληθής για πίνακες τάξης 1. Υποθέτουμε πως ισχύει για n και θα δείξουμε ότι ισχύει για $n+1$, όπου έχει τον πίνακα M ως πάνω αριστερά υποπίνακα τάξης n , με

$$\gamma'_{ii} = \gamma_{ii} - |m_{i,n+1}|^2, \quad i = 1, 2, \dots, n.$$

{ Πράγματι, ο παραπάνω τύπος ισχύει και μπορούμε να τον επαληθεύσουμε μ' ένα παράδειγμα:

Έστω άνω τριγωνικός $T = \begin{bmatrix} 1 & 2 \\ 0 & 3 \end{bmatrix}$.

Επομένως, $T^* = \begin{bmatrix} 1 & 0 \\ 2 & 3 \end{bmatrix}$, $T^*T = \begin{bmatrix} 1 & 2 \\ 2 & 13 \end{bmatrix}$, $TT^* = \begin{bmatrix} 5 & 6 \\ 6 & 9 \end{bmatrix}$ και $\Gamma = \begin{bmatrix} -4 & -4 \\ -4 & 4 \end{bmatrix}$.

Ακόμη έστω $T' = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 3 & 0 \\ 0 & 0 & 1 \end{bmatrix}$.

Άρα $T'^* = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 3 & 0 \\ 1 & 0 & 1 \end{bmatrix}$, $T'^*T' = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 13 & 2 \\ 1 & 2 & 2 \end{bmatrix}$, $T'(T')^* = \begin{bmatrix} 6 & 9 & 1 \\ 6 & 9 & 0 \\ 1 & 0 & 1 \end{bmatrix}$ και $\Gamma' = \begin{bmatrix} -5 & -7 & 0 \\ -4 & 4 & 2 \\ 0 & 2 & 1 \end{bmatrix}$.

Παρατηρούμε ότι ισχύει

$$\gamma_{11} - |m_{1,3}|^2 = -4 - 1^2 = -5 = \gamma'_{11} \text{ και } \gamma_{22} - |m_{2,3}|^2 = 4 - 0 = 4 = \gamma'_{22}. \}$$

Έχουμε

$$\begin{aligned} \|M'\|_F^2 &= \|M\|_F^2 + |m_{1,n+1}|^2 + |m_{2,n+1}|^2 + \dots + |m_{n,n+1}|^2 \\ &\leq \gamma_{22} + 2\gamma_{33} + \dots + (n-1)\gamma_{nn} + \gamma'_{n+1,n+1} \\ &= \gamma'_{22} + |m_{2,n+1}|^2 + 2\gamma'_{33} + 2|m_{3,n+1}|^2 + \dots + \\ &\quad + (n-1)\gamma'_{nn} + (n-1)|m_{n,n+1}|^2 + \gamma'_{n+1,n+1} \\ &= \gamma'_{22} + 2\gamma'_{33} + \dots + (n-1)\gamma'_{nn} + \sum_{k=2}^n (k-1)|m_{k,n+1}|^2 + \gamma'_{n+1,n+1} \\ &\leq \gamma'_{22} + 2\gamma'_{33} + \dots + (n-1)\gamma'_{n+1,n+1} + \gamma'_{n+1,n+1} \\ &\leq \gamma'_{22} + 2\gamma'_{33} + \dots + n\gamma'_{n+1,n+1}, \end{aligned}$$

από την (3.1.9), αποδεικνύοντας την (3.1.8) με n αυξανόμενο κατά ένα.

Η ισότητα της σχέσης (3.1.8) ισχύει αν και μόνο αν $m_{ij} = 0, \forall j > i + 1$, δηλαδή εάν ο τριγωνικός πίνακας T είναι “σχεδόν διαγώνιος”.

Μια συνέπεια της (3.1.9) είναι

$$\gamma_{11} + \gamma_{22} + \gamma_{33} + \dots + \gamma_{nn} = 0.$$

Έχουμε

$$\|M\|_F^2 - \frac{n-1}{2} (\gamma_{11} + \gamma_{22} + \gamma_{33} + \dots + \gamma_{nn}) \leq \|M\|_F^2 \leq \gamma_{22} + 2\gamma_{33} + \dots + (n-1)\gamma_{nn}$$

και συνεπώς,

$$\|M\|_F^2 - \frac{n-1}{2}\gamma_{11} - \frac{n-1}{2}\gamma_{22} - \frac{n-1}{2}\gamma_{33} - \dots - \frac{n-1}{2}\gamma_{nn} \leq \|M\|_F^2 \leq \gamma_{22} + 2\gamma_{33} + \dots + (n-1)\gamma_{nn}.$$

Επομένως βρίσκουμε

$$\|M\|_F^2 \leq \frac{1-n}{2}\gamma_{11} + \frac{3-n}{2}\gamma_{22} + \dots + \frac{n-1}{2}\gamma_{nn}$$

και χρησιμοποιώντας την ανισότητα Cauchy προκύπτει

$$\begin{aligned} \|M\|_F^4 &\leq \left(\frac{1-n}{2}\gamma_{11} + \frac{3-n}{2}\gamma_{22} + \dots + \frac{n-1}{2}\gamma_{nn} \right)^2 \leq \\ &\leq \left[\left(\frac{1-n}{2} \right)^2 + \left(\frac{3-n}{2} \right)^2 + \dots + \left(\frac{n-1}{2} \right)^2 \right] [\gamma_{11}^2 + \gamma_{22}^2 + \dots + \gamma_{nn}^2]. \end{aligned} \quad (3.1.10)$$

Στο σημείο αυτό υπενθυμίζουμε το Θεώρημα Cauchy.

Θεώρημα Cauchy: Αν a_1, \dots, a_n και b_1, \dots, b_n είναι δύο n -άδες μιγαδικών αριθμών, τότε

$$(a_1 b_1 + a_2 b_2 + \dots + a_n b_n)^2 \leq (a_1^2 + a_2^2 + \dots + a_n^2) (b_1^2 + b_2^2 + \dots + b_n^2)$$

ή, ισοδύναμα,

$$|a_1 b_1 + a_2 b_2 + \dots + a_n b_n| \leq \sqrt{a_1^2 + a_2^2 + \dots + a_n^2} \sqrt{b_1^2 + b_2^2 + \dots + b_n^2}.$$

Ένας στοιχειώδης υπολογισμός δείχνει ότι

$$\left(\frac{1-n}{2}\right)^2 + \left(\frac{3-n}{2}\right)^2 + \dots + \left(\frac{n-1}{2}\right)^2 = \frac{n^3 - n}{12}.$$

Έτσι προκύπτει το αποτέλεσμα (3.1.7) λόγω του γεγονότος ότι

$$\begin{aligned} A^* A - A A^* &= (U T U^*)^* U T U^* - U T U^* (U T U^*)^* = U T^* U^* U T U^* - U T U^* U T^* U^* = \\ &= U T^* T U^* - U T T^* U^* = U (T^* T - T T^*) U^* = U \Gamma U^* \end{aligned}$$

και ως εκ τούτου έχουμε

$$\gamma_{11}^2 + \gamma_{22}^2 + \dots + \gamma_{nn}^2 \leq \|\Gamma\|_F^2 = \|U \Gamma U^*\|_F^2 = \|A^* A - A A^*\|_F^2. \quad (3.1.11)$$

Η ισότητα ισχύει στην (3.1.7) αν και μόνο αν η ισότητα ισχύει ταυτόχρονα στην (3.1.8), στην (3.1.10) και στην (3.1.11). Η ισότητα στην (3.1.11) απαιτεί ο πίνακας $\Gamma = T^* T - T T^*$ να είναι διαγώνιος πίνακας. Για τους σχεδόν διαγώνιους πίνακες T (όπως απαιτείται από την ισότητα στην (3.1.8)), ο Γ είναι διαγώνιος αν και μόνο αν όλα τα διαγώνια στοιχεία του T ταυτίζονται. Τα διαγώνια στοιχεία του Γ (με βάση και την (3.1.9)) ικανοποιούν τη σχέση

$$|m_{k,k+1}|^2 = \gamma_{11} + \gamma_{22} + \gamma_{33} + \dots + \gamma_{kk}, \quad k = 1, 2, \dots, n-1. \quad (3.1.12)$$

Η ισότητα στην (3.1.10) τελικά προϋποθέτει ότι $\gamma_{kk} = \alpha (1 + n - 2k)$, $k = 1, 2, \dots, n$, όπου α είναι μια πραγματική σταθερά. Η σχέση (3.1.12) ικανοποιεί την υπόθεση αν και μόνο αν, για ορισμένους πραγματικούς $\alpha \geq 0$,

$$|m_{k,k+1}|^2 = \alpha k (n - k), \quad k = 1, 2, \dots, n-1,$$

δηλαδή, εάν ο T έχει τη μορφή που αναφέρεται στο θεώρημα. □

Θεώρημα 3.2: Εάν λ_j είναι μια απλή ιδιοτιμή του A , τότε

$$\kappa_A(\lambda_j) \leq \{1 + (n-1)^{-1} \delta_j^{-2} (\delta_{\|\cdot\|_F}(A))^2\}^{(n-1)/2}, \quad (3.1.13)$$

όπου $\delta_j = \min |\lambda_j - \lambda_i|$ για όλα τα $i \neq j$ με $1 \leq i \leq n$.

Απόδειξη: Όπως και στην απόδειξη του Θεωρήματος 2.2 υποθέτουμε χωρίς βλάβη της γενικότητας, ότι ο A είναι ένας άνω τριγωνικός πίνακας με συγκεκριμένη ιδιοτιμή λ_j ως το οδηγό διαγώνιο στοιχείο του. Δηλαδή $A = \Lambda + M$, όπου $\Lambda = \text{diag}(\lambda_i)$ και M είναι ένας πίνακας του οποίου μόνο τα μη-μηδενικά στοιχεία εμφανίζονται πάνω από την κύρια διαγώνιο.

Το i° διαγώνιο στοιχείο του $(\lambda_j I - A)^* (\lambda_j I - A)$ είναι επομένως το

$$|\lambda_j - \lambda_i|^2 + \|m_i\|_2^2 \quad (3.1.14),$$

όπου το διάνυσμα m_i είναι η $i^{\text{η}}$ στήλη του M .

Επίσης

$$\sum \|m_i\|_2^2 = \|A\|_F^2 - \|\Lambda\|_F^2 = (\delta_{\|\cdot\|_F}(A))^2 \quad (3.1.15)$$

από τη σχέση (3.1.6).

Με $C = \lambda_j I - A$ η (2.2.6) γίνεται

$$\{\text{adj}(\lambda_j I - A)\} \{\text{adj}(\lambda_j I - A)\}^* = \text{adj}\{(\lambda_j I - A)^* (\lambda_j I - A)\}. \quad (3.1.16)$$

Εάν το διάνυσμα y^* είναι η πρώτη γραμμή του $\text{adj}(\lambda_j I - A)$, τότε το οδηγό διαγώνιο στοιχείο του πίνακα στην αριστερή πλευρά της (3.1.16) είναι $\|y\|_2^2$. Από τη δεξιά πλευρά του (3.1.16) έχουμε $\|y\|_2^2 = \det B$, όπου B σχηματίζεται διαγράφοντας την πρώτη γραμμή και στήλη από τον πίνακα $(\lambda_j I - A)^* (\lambda_j I - A)$.

Επειδή ο πίνακας B είναι ερμιτιανός και θετικά ορισμένος, η ανισότητα Hadamard δίνει

$$\|y\|_2^2 = \det B \leq b_1 b_2 \dots b_{n-1},$$

όπου b_1, b_2, \dots, b_{n-1} είναι τα διαγώνια στοιχεία του B .

Ο B είναι $(n-1) \times (n-1)$ πίνακας, αφού προκύπτει από τη διαγραφή της πρώτης γραμμής και της πρώτης στήλης του $n \times n$ πίνακα $(\lambda_j I - A)^* (\lambda_j I - A)$.

Δεδομένου ότι αυτά είναι επίσης τα διαγώνια στοιχεία του $(\lambda_j I - A)^* (\lambda_j I - A)$, προκύπτει από την (3.1.14) ότι

$$\begin{aligned} \|y\|_2^2 &\leq \prod_j \{|\lambda_j - \lambda_i|^2 + \|m_i\|_2^2\} \leq \prod_j \left\{ 1 + \frac{\|m_i\|_2^2}{|\lambda_j - \lambda_i|^2} \right\} |\lambda_j - \lambda_i|^2 \leq \\ &\leq \prod_j \left\{ 1 + \frac{\|m_i\|_2^2}{\delta_j^2} \right\} \prod_j |\lambda_j - \lambda_i|^2 \leq \prod_j \{1 + \delta_j^{-2} \|m_i\|_2^2\} \cdot \prod_j |\lambda_j - \lambda_i|^2. \end{aligned} \quad (3.1.17)$$

Εάν \sum_j δηλώνει το άθροισμα πάνω από όλα τα $i \neq j$ με $1 \leq i \leq n$, τότε η σύγκριση των γεωμετρικών και αριθμητικών μέσων δίνει

$$\begin{aligned} \prod_j \{1 + \delta_j^{-2} \|m_i\|_2^2\}^{1/(n-1)} &\leq (n-1)^{-1} \sum_j \{1 + \delta_j^{-2} \|m_i\|_2^2\} \leq \\ &\leq 1 + (n-1)^{-1} \delta_j^{-2} (\delta_{\|\cdot\|_F}(A))^2, \end{aligned}$$

από την (3.1.15).

Αντικαθιστώντας την παραπάνω σχέση στην (3.1.17) παίρνουμε

$$\|y\|_2^2 \leq [1 + (n-1)^{-1} \delta_j^{-2} (\delta_{\|\cdot\|_F}(A))^2]^{n-1} \cdot \prod_j |\lambda_j - \lambda_i|^2.$$

Από αυτό και χρησιμοποιώντας τις σχέσεις (2.2.2) και (2.2.9) συνάγουμε την (3.1.13). Η απόδειξη του Θεωρήματος 3.2 μόλις ολοκληρώθηκε. \square

Μια άμεση συνέπεια της (3.1.13) είναι

$$\kappa_A(\lambda_j) \leq \exp((\delta_{\|\cdot\|_F}(A))^2 / 2\delta_j^2).$$

Για $n = 3$ η σχέση (3.1.13) μας δίνει

$$\kappa_A(\lambda_j) \leq 1 + (\delta_{\|\cdot\|_F}(A))^2 / 2\delta_j^2 \leq \exp((\delta_{\|\cdot\|_F}(A))^2 / 2\delta_j^2),$$

αφού η $\exp((\delta_{\|\cdot\|_F}(A))^2 / 2\delta_j^2) - (\delta_{\|\cdot\|_F}(A))^2 / 2\delta_j^2 - 1$ είναι γνησίως αύξουσα συνάρτηση στο $[0, +\infty)$.

Παρατήρηση: Στο παράδειγμα 1 του τμήματος 2.7 η εκτίμηση (3.1.13) δίνει

$$\kappa_A(\lambda_1) \leq 1 + \frac{1}{2} v^2 \max(|u|^{-2}, |1 + u|^{-2}). \quad (3.1.13)$$

Επομένως, τα $\kappa_A(\lambda_1)$ παραμένουν και αυτά, μαζί με τα κ_2 , οριοθετημένα, όταν $u \rightarrow 0$, καθώς οι δεξιές πλευρές του (3.1.13) και (2.5.6) τείνουν στο ∞ . Όλες οι εκτιμήσεις αυτές ενδέχεται να μην έχουν “υποστεί” επεξεργασία, όταν η διακρίνουσα Δ του παραδείγματος είναι μικρή ή η απομάκρυνση από την κανονικότητα $|v|$ είναι μεγάλη.

3.2 Δυνάμεις Πινάκων

3.2.1 Εισαγωγή.

Ορισμός 29. Ονομάζουμε *δύναμη πίνακα* έναν πίνακα της μορφής A^r , $r = 1, 2, \dots$.

Σημείωση: Μια απλή συνέπεια της Schur τριγωνικής μορφής είναι ότι για έναν κανονικό πίνακα A ισχύει

$$\|A^r\|_2 = \rho^r(A), \quad r = 1, 2, \dots \quad (3.2.1)$$

Φράγματα για νόρμες ορισμένων μη-κανονικών πινάκων (που εμφανίζονται σε πεπερασμένα διαφορετικά σχήματα για την επίλυση υπερβολικών και παραβολικών διαφορικών εξισώσεων) έχουν δοθεί από τους Lax, Richtmyer και Kato. Για

αυθαίρετους πίνακες το πρόβλημα έχει επεξεργαστεί από τον Gautschi και τον Ostrowski. Ο Ostrowski έδειξε ότι για κάθε $\varepsilon > 0$

$$\|A^r\|_\infty \leq \kappa_\infty(\varepsilon) (\rho(A) + \varepsilon)^r, \quad r = 1, 2, \dots$$

Εδώ η συνάρτηση $\kappa_\infty(\varepsilon)$ ορίζεται ως εξής: Εάν ο J είναι μια Jordan κανονική μορφή του πίνακα $2\varepsilon^{-1}A$ και εάν $2\varepsilon^{-1}A = S J S^{-1}$, τότε $\kappa_\infty(\varepsilon) = \|S\|_\infty \|S^{-1}\|_\infty$, όπου $\kappa_\infty(\varepsilon)$ είναι ο δείκτης κατάστασης του S ως προς τη $\|\circ\|_\infty$. Το φράγμα του Gautschi ομοίως απαιτεί κάποια γνώση της Jordan κανονικής μορφής.

3.2.2 Ένα νέο φράγμα. Σε αντίθεση με τα παραπάνω, το Θεώρημα 3.3 παρακάτω μας δίνει μια εκτίμηση για το $\|A^r\|_2$, η οποία εξαρτάται μόνο από τη $\rho(A)$ και το $\delta_{\|\circ\|_2}(A)$ και τείνει στη (3.2.1) για A κανονικό πίνακα.

Θεώρημα 3.3: Εάν $m = \delta_{\|\circ\|_2}(A)$ και

- εάν $\rho(A) > 0$, τότε

$$\|A^r\|_2 \leq \rho^r(A) + \binom{r}{1} \rho^{r-1}(A) m + \dots + \binom{r}{n-1} \rho^{r-n+1}(A) m^{n-1} \quad (3.2.2)$$

- εάν $\rho(A) = 0$, τότε

$$\begin{cases} \|A^r\|_2 \leq m^r, & r = 0, 1, \dots, n-1, \\ \|A^r\|_2 = 0, & r \geq n. \end{cases} \quad (3.2.3)$$

Παρατήρηση: Η έκφραση στη δεξιά πλευρά του (3.2.2) μπορεί να εκτιμηθεί περαιτέρω με διάφορους τρόπους και για μεγάλα r είναι ασυμπτωτικά ίση με

$$\frac{\rho^{r-n+1}(A)(mr)^{n-1}}{(n-1)!}.$$

Απόδειξη: Έστω $A = UTU^*$, όπου U είναι ένας ορθομοναδιαίος πίνακας και $T = D + M$ είναι μια Schur τριγωνική μορφή. Τότε έχουμε

$$A^r = U(D + M)^r U^*.$$

Δεδομένου ότι ο D και ο M , γενικά, δεν αντιμετατίθενται, ο $(D + M)^r$ δεν μπορεί να επεκταθεί σύμφωνα με το διωνυμικό ανάπτυγμα του Newton. Ωστόσο, αφού ο D είναι διαγώνιος, εξακολουθεί να είναι αληθές ότι, αν τον επεκτείνουμε χωρίς αντιμετάθεση, οποιοσδήποτε όρος του που περιέχει περισσότερους από $n - 1$ πίνακες M είναι μηδέν. Υπάρχουν ακριβώς $\binom{r}{q}$ όροι που “αναμειγνύουν” qM και $r - qD$ για $q = 0, 1, 2, \dots, n - 1$. Χρησιμοποιώντας τη νόρμα $\|\circ\|_2$ και τη σχέση $\|D\|_2 = \rho(A)$, αντικαθιστούμε στις σχέσεις (3.2.2) και (3.2.3), όπου m το $\|M\|_2$. Επειδή το αποτέλεσμα είναι αληθές για κάθε M και επειδή οι εκφράσεις στα δεξιά της (3.2.2) και της (3.2.3) εξαρτώνται συνεχώς από το m , διαπιστώνουμε ότι το συμπέρασμα του θεωρήματος προκύπτει αμέσως. \square

3.3 Προσεγγιστικές Λύσεις για Γραμμικά Συστήματα

3.3.1 *Εισαγωγή.* Έστω A είναι ένας αντιστρέψιμος πίνακας, έστω b ένα δοθέν διάνυσμα και έστω \tilde{x} μια υποτιθέμενη λύση της $Ax = b$. Εάν ορίσουμε το υπόλοιπο του \tilde{x} ως $r = A\tilde{x} - b$ και εάν $\|\cdot\|$ είναι μια νόρμα διάνυσμα, το σφάλμα $\tilde{x} - A^{-1}b = A^{-1}r$ του \tilde{x} μπορεί να εκτιμηθεί ως εξής

$$\|\tilde{x} - A^{-1}b\| = \|A^{-1}r\| \leq \|A^{-1}\| \|r\|.$$

Εδώ η $\|\cdot\|$ υποδηλώνει μια νόρμα πίνακα συμβατή με τη $\|\cdot\|$. Παρομοίως, εάν ο \tilde{X} είναι ένας υποτιθέμενος αντίστροφος του A και εάν $\|\cdot\|$ είναι μια οποιαδήποτε νόρμα πίνακα, μπορούμε να υπολογίσουμε ένα φράγμα για τη $\|\tilde{X} - A^{-1}\|$ σε σχέση με τη νόρμα του πίνακα υπολοίπου $R = A\tilde{X} - I$ ως εξής

$$\|\tilde{X} - A^{-1}\| = \|A^{-1}R\| \leq \|A^{-1}\| \|R\|.$$

Έτσι και για τα δύο προβλήματα απαιτούμε ένα φράγμα για το $\|A^{-1}\|$. Ένα τέτοιο φράγμα, κατ' αρχήν, εύκολα “κατασκευάζεται” εάν υποθέσουμε ότι ο A είναι όμοιος με ένα διαγώνιο πίνακα D , δηλαδή

$$A = S D S^{-1}.$$

Υποθέτοντας ότι η νόρμα $\|\cdot\|$ είναι προσανατολισμένη κατά άξονα και βρίσκουμε εύκολα από τη σχέση $A^{-1} = S^{-1} D^{-1} S$ ότι

$$\|A^{-1}\| = \kappa(S) \rho(A^{-1}), \quad (3.3.1)$$

όπου το

$$\kappa(S) = \|S\| \|S^{-1}\| \quad (3.3.2)$$

υποδηλώνει το δείκτη κατάστασης του S ως προς τη $\|\cdot\|$ που εισήχθη από τον Bauer. Εάν ο A είναι κανονικός πίνακας, τότε ο S μπορεί να θεωρηθεί ορθομοναδιαίος και ο φασματικός δείκτης κατάστασης του S ισούται με 1. Έτσι

$$\|A^{-1}\|_2 \leq \|S^{-1}\|_2 \|D^{-1}\|_2 \|S\|_2 = \|D^{-1}\|_2 = \rho(A^{-1}),$$

και ενόψει της (3.1.1) προκύπτει

$$\|A^{-1}\|_2 = \rho(A^{-1}) \quad (3.3.3)$$

για κανονικούς πίνακες A . Για μη-κανονικούς πίνακες το φράγμα (3.3.1), αν σε όλες τις εφαρμογές απαιτείται η πλήρης διαγωνοποίηση του πίνακα A , είναι μια μάλλον υψηλή τιμή για το επιθυμητό αποτέλεσμα.

3.3.2 Ένα νέο φράγμα για τη νόρμα του αντιστρόφου. Για μια πραγματική μεταβλητή $x \geq 0$, έστω η συνάρτηση f οριζόμενη ως εξής

$$f(x) = x + x^2 + \dots + x^n. \quad (3.3.4)$$

Σημειώνουμε ότι και η f και η $x^{-1}f$ γνησίως αύξουσες συναρτήσεις για $x > 0$. Τότε

$$\lim_{x \rightarrow 0^+} x^{-1}f(x) = 1. \quad (3.3.5)$$

Μπορούμε τώρα να αναφέρουμε το ακόλουθο Θεώρημα.

Θεώρημα 3.4: Εάν ο A είναι ένας αντιστρέψιμος και μη-κανονικός πίνακας και εάν $x = \rho(A^{-1}) \delta_{\|\circ\|_2}(A)$, τότε

$$\|A^{-1}\|_2 \leq \frac{f(x)}{x} \rho(A^{-1}) = \frac{f(x)}{\delta_{\|\circ\|_2}(A)}. \quad (3.3.6)$$

Παρατήρηση: Για $\rho(A^{-1})$ σταθερό, καθώς $\delta_{\|\circ\|_2}(A) \rightarrow 0$, το φράγμα (3.3.6) προσεγγίζει την (3.3.3), ενόψει της (3.3.5). Φράγματα για τη $\|A^{-1}\|$, όπου $\|\circ\|$ είναι αυθαίρετη νόρμα μπορούν να βρεθούν από την (3.1.2).

Απόδειξη: Έστω $D + M$ μια Schur τριγωνική μορφή του A . Τότε για κάποιο ορθομοναδιαίο πίνακα U έχουμε

$$A^{-1} = U (D + M)^{-1} U^* = U (I + D^{-1}M)^{-1} D^{-1} U^*.$$

Όμως, με βάση την (3.1.3), ισχύει $(D^{-1}M)^n = O$. Άρα

$$(I + D^{-1}M)^{-1} = I - D^{-1}M + \dots + (-1)^{n-1} (D^{-1}M)^{n-1}.$$

Έτσι έχουμε το αποτέλεσμα (3.3.6) με τη λήψη νορμών και ενόψει της $\|D^{-1}\|_2 = \rho(A^{-1})$ και του γεγονότος ότι η $\|\circ\|_2$ είναι ορθομοναδιαία αναλλοίωτη. \square

3.4 Φασματική Μεταβολή και Μεταβολή Ιδιοτιμών

3.4.1 Κλασικά αποτελέσματα. Έστω ο πίνακας $A = (a_{ij})$ έχει ιδιοτιμές λ_i και έστω ο $B = (b_{ij})$ έχει ιδιοτιμές μ_i ($i = 1, 2, \dots, n$). Η ποσότητα

$$s = s_A(B) = \max_{1 \leq i \leq n} \{ \min_{1 \leq j \leq n} |\mu_i - \lambda_j| \} \quad (3.4.1)$$

ονομάζεται *φασματική μεταβολή* του B σε σχέση με το A . Δεν υπονοείται κάποια ένα-προς-ένα αντιστοιχία μεταξύ των ιδιοτιμών του A και εκείνων του B . Γενικά από απλά παραδείγματα μπορούμε να παρατηρήσουμε ότι $s_A(B) \neq s_B(A)$.

Θεωρούμε, επίσης, την ποσότητα

$$v = v(A, B) = \min_p \{ \max_{1 \leq i \leq n} |\lambda_i - \mu_{p(i)}| \}, \quad (3.4.2)$$

η οποία ονομάζεται *μεταβολή ιδιοτιμής* του A και του B . Εδώ το ελάχιστο λαμβάνεται σε σχέση με όλες τις μεταθέσεις p για $i = 1, 2, \dots, n$. Προφανώς,

$$v(A, B) = v(B, A)$$

και

$$s_A(B) \leq v(A, B) \quad (3.4.3)$$

για όλους τους πίνακες A και B .

Τα καλύτερα “διαθέσιμα” φράγματα για το s και το v , με βάση τον Ostrowski, έχουν ως ακολούθως. Εάν $M = \max_{1 \leq i, j \leq n} (|a_{ij}|, |b_{ij}|)$ και αν η νόρμα l_1 είναι ορισμένη όπως στην Παράγραφο 1.2, τότε

$$s_A(B) \leq (n+2) M \left[\frac{\|A-B\|_{l_1}}{M} \right]^{1/n} \quad (3.4.4)$$

και

$$v(A, B) \leq 2n(n+2) M \left[\frac{\|A-B\|_{l_1}}{M} \right]^{1/n}. \quad (3.4.5)$$

Εξετάζοντας ένα παράδειγμα γίνεται εύκολα αντιληπτό, σύμφωνα με τον G.E. Forsythe, ότι ο εκθέτης $1/n$ σ' αυτά τα φράγματα δεν μπορεί γενικά να βελτιωθεί. Σε ειδικές περιπτώσεις, ωστόσο, οι βελτιώσεις είναι πιθανές. Εάν ο A είναι όμοιος με ένα διαγώνιο πίνακα D έτσι ώστε

$$A = S D S^{-1}$$

και αν $\|\circ\|$ είναι οποιαδήποτε προσανατολισμένη κατά άξονα επαγόμενη νόρμα, τότε ο Bauer και ο Fike έδειξαν ότι

$$s_A(B) \leq \kappa(S) \|A - B\|. \quad (3.4.6)$$

Συγκεκριμένα, εάν ο A είναι κανονικός, βρίσκουμε για οποιαδήποτε νόρμα $\|\circ\|$, θεωρώντας ως μεγαλύτερη τη φασματική νόρμα, ότι

$$s_A(B) \leq \|A - B\|. \quad (3.4.7)$$

Για $\|\circ\| = \|\circ\|_F$ και για A, B κανονικούς πίνακες, από το αποτέλεσμα του Hoffman και του Wielandt προκύπτει ότι η (3.4.7) είναι ακόμη σε ισχύ για τη μεταβολή ιδιοτιμής. Έχουμε

$$v(A, B) \leq \|A - B\|_F. \quad (3.4.8)$$

Για A και B πραγματικούς συμμετρικούς ή ερμιτιανούς πίνακες αυτό το αποτέλεσμα χρησιμοποιείται συχνά.

3.4.2 Ένα νέο αποτέλεσμα για το $s_A(B)$. Για τυχαίους διαγωνοποιήσιμους πίνακες A το αποτέλεσμα (3.4.6) δεν μπορεί να δώσει κάποια αρκετά σαφή εκτίμηση για το δείκτη κατάστασης $\kappa(S)$. Το ακόλουθο θεώρημα περιέχει (ενόψει του Θεωρήματος 3.1) μια άμεση εκτίμηση τόσο γενική όσο η (3.4.4) (και έτσι πιο γενικευμένη από την (3.4.6)) που ανάγεται στην (3.4.7) για A κανονικό πίνακα.

Είναι βολικό να ορίσουμε, για μια πραγματική μεταβλητή $y \geq 0$, τη συνάρτηση $g = g(y)$ ως τη (μοναδική) μη-αρνητική λύση της εξίσωσης

$$g + g^2 + \dots + g^n = y.$$

Η συνάρτηση g είναι η αντίστροφη της συνάρτησης f που έχουμε ορίσει στο τμήμα 3.3.2. Για μεταγενέστερη χρήση σημειώνουμε τις ιδιότητες:

$$\lim_{y \rightarrow 0^+} y^{-1} g(y) = 1, \quad (3.4.9)$$

$$n^{-1} y \leq g(y) \leq y, \quad 0 \leq y \leq n, \quad (3.4.10)$$

$$g(n) = 1, \quad (3.4.11)$$

$$(n^{-1} y)^{1/n} \leq g(y) \leq y^{1/n}, \quad y \geq n, \quad (3.4.12)$$

$$\lim_{y \rightarrow \infty} y^{-1/n} g(y) = 1. \quad (3.4.13)$$

Θεώρημα 3.5: Έστω A ένας μη-κανονικός πίνακας και έστω $B - A \neq 0$. Εάν $\|\cdot\|$ είναι οποιαδήποτε νόρμα μεγαλύτερη της φασματικής νόρμα και εάν

$$y = \frac{\delta_{\|\cdot\|}(A)}{\|B - A\|},$$

τότε

$$s_A(B) \leq \frac{y}{g(y)} \|B - A\| = \frac{\delta_{\|\cdot\|}(A)}{g(y)}. \quad (3.4.14)$$

Παρατηρήσεις: Οι σχέσεις (3.4.10), (3.4.11) και (3.4.12) μπορούν να χρησιμεύσουν για να καταστεί το φράγμα στο (3.4.14) πιο σαφές. Η σχέση (3.4.9) δείχνει ότι, για $\delta_{\|\cdot\|}(A) \rightarrow 0$ και το όριο του $\|B - A\|$ να απομακρύνεται από το μηδέν, η εκτίμηση (3.4.14) προσεγγίζει την (3.4.7). (Ωστόσο, δεν μπορούμε να λάβουμε την (3.4.8) μ' αυτόν τον τρόπο.) Η σχέση (3.4.13) δείχνει ότι για ένα σταθερό μη-κανονικό πίνακα A και για $B \rightarrow A$ το φράγμα στη σχέση (3.4.14) είναι της ίδιας τάξης με την (3.4.4). Επίσης, ο αριθμητικός συντελεστής μπορεί να είναι μεγαλύτερος ή μικρότερος από τον συντελεστή στην (3.4.4) ανάλογα με την απομάκρυνσή του από την κανονικότητα του A .

Απόδειξη του Θεωρήματος 3.5: Έστω $A = UTU^*$, όπου $T = D + M$ είναι μια Schur τριγωνική μορφή του A . Θέτοντας $E = B - A$, $U^*BU = B_1$ και $U^*EU = F$, έχουμε

$$B_1 = U^*BU = U^*(E + A)U = U^*EU + U^*AU = F + U^*(UTU^*)U = F + T = F + D + M.$$

Άρα

$$B_1 = D + M + F.$$

Έστω μ είναι μια ιδιοτιμή του B (ή του B_1) η οποία δεν είναι ιδιοτιμή του A . Τότε ο πίνακας $D + M - \mu I$ είναι αντιστρέψιμος και έχουμε

$$\begin{aligned} 0 &= \det(B_1 - \mu I) = \det(D + M - \mu I + F) \\ &= \det(D + M - \mu I) \det[I + (D + M - \mu I)^{-1}F]. \end{aligned}$$

Αφού $\det(D + M - \mu I) \neq 0$, ισχύει

$$\det[I + (D + M - \mu I)^{-1}F] = 0$$

και προκύπτει ότι το -1 είναι μια ιδιοτιμή του πίνακα $(D + M - \mu I)^{-1}F$. Από τη θεμελιώδη ανισότητα (3.1.1) προκύπτει ότι

$$\|(D + M - \mu I)^{-1}F\|_2 \geq |\lambda| = 1$$

ή, από τη (iv), αφού $\|F\|_2 = \|E\|_2$, θα πάρουμε

$$\|(D + M - \mu I)^{-1}\|_2 \|E\|_2 \geq \|(D + M - \mu I)^{-1}\|_2 \geq 1$$

Άρα

$$\|(D + M - \mu I)^{-1}\|_2 \geq \frac{1}{\|E\|_2}. \quad (3.4.15)$$

Τώρα θα πρέπει να εκτιμήσουμε τη $\|(D + M - \mu I)^{-1}\|_2$ από παραπάνω. Επειδή ο $D - \mu I$ είναι αντιστρέψιμος, έχουμε

$$(D + M - \mu I)^{-1} = [I + (D - \mu I)^{-1}M]^{-1}(D - \mu I)^{-1}. \quad (3.4.16)$$

Επειδή ο $(D - \mu I)^{-1}$ είναι διαγώνιος, ο πίνακας $(D - \mu I)^{-1}M$ “μοιράζεται” με τον M την ιδιότητα ότι τα στοιχεία κάτω από την κύρια διαγώνιο είναι μηδέν και συνεπώς ότι

$$[(D - \mu I)^{-1}M]^r = 0 \text{ για } r \geq n.$$

Ως εκ τούτου,

$$[I + (D - \mu I)^{-1}M]^{-1} = I - (D - \mu I)^{-1}M + \dots + (-1)^{n-1}[(D - \mu I)^{-1}M]^{n-1}. \quad (3.4.17)$$

Θέτουμε για λόγους συντομίας

$$\|(D - \mu I)^{-1}\|_2 = p, \quad \|M\|_2 = m (\neq 0), \quad \|E\|_2 = e$$

και έχουμε από την (3.4.16) και την (3.4.17) ότι

$$\|(D + M - \mu I)^{-1}\|_2 \leq p + p^2 m + \dots + p^n m^{n-1} = m^{-1} f(mp)$$

όπου η f ορίζεται από την (3.3.4).

Συνδυάζοντας το τελευταίο αποτέλεσμα με την (3.4.15) βρίσκουμε ότι $f(mp) \geq me^{-1}$ και έτσι, από τον ορισμό της g , έχουμε $mp \geq g(me^{-1})$ ή

$$\frac{1}{p} \leq \frac{m}{g(me^{-1})}. \quad (3.4.18)$$

Επειδή η $\|\circ\|$ είναι μια προσανατολισμένη κατά άξονα επαγόμενη νόρμα,

$$p = \|(D - \mu I)^{-1}\|_2 = \max_{1 \leq j \leq n} |\lambda_j - \mu|^{-1}$$

και

$$\frac{1}{p} = \min_{1 \leq j \leq n} |\lambda_j - \mu|.$$

Έτσι, από την (3.4.18), προκύπτει ότι

$$\min_{1 \leq j \leq n} |\lambda_j - \mu| \leq \frac{m}{g(me^{-1})}. \quad (3.4.19)$$

Αυτή η σχέση έχει αποδειχθεί για μια αυθαίρετη ιδιοτιμή $\mu = \mu_i$ του B η οποία δεν αποτελεί, όμως, μια ιδιοτιμή του A . Ισχύει επίσης για τετριμμένα μ_i που είναι ιδιοτιμές του A . Έτσι, παίρνουμε

$$\begin{aligned} s_A(B) &= \max_{1 \leq i \leq n} \{ \min_{1 \leq j \leq n} |\lambda_j - \mu_i| \} \\ &\leq \frac{m}{g(me^{-1})} = \frac{y_0}{g(y_0)} \|B - A\|_2, \quad y_0 = \frac{\|M\|_2}{\|B - A\|_2}. \end{aligned}$$

Αυτό είναι αληθές για οποιαδήποτε επιλογή της Schur τριγωνικής μορφής $D + M$. Αφού η συνάρτηση g είναι συνεχής, ακολουθεί η δήλωση του Θεωρήματος 3.5 για την ειδική νόρμα $\|\circ\| = \|\circ\|_2$.

Έστω, τώρα, $\|\circ\|$ να είναι μια νόρμα μεγαλύτερη της $\|\circ\|_2$. Δεδομένου ότι η συνάρτηση g είναι μη-αρνητική και αυξάνεται μονοτονικά, έχουμε επίσης

$$s_A(B) \leq \frac{\|M\|_2}{g\left(\frac{\|M\|_2}{\|B - A\|}\right)} = \|B - A\| \frac{y_1}{g(y_1)}, \quad \text{με } y_1 = \frac{\|M\|_2}{\|B - A\|}. \quad (3.4.20)$$

Έστω $0 < y_1 < y_2$ και ορίζουμε $x_i = g(y_i)$, $i = 1, 2$. Από την μονοτονία του $x^{-1}f(x)$ προκύπτει ότι

$$\frac{y_1}{g(y_1)} = \frac{f(x_1)}{x_1} < \frac{f(x_2)}{x_2} = \frac{y_2}{g(y_2)}.$$

Επίσης η συνάρτηση $y[g(y)]^{-1}$ είναι γνησίως αύξουσα και αντικαθιστώντας τη $\|M\|_2$ με $\|M\|$ στην (3.4.20), παίρνουμε

$$s_A(B) \leq \frac{y_2}{g(y_2)} \|B - A\|, \quad y_2 = \frac{\|M\|}{\|B - A\|}.$$

Η πλήρης δήλωση του Θεωρήματος 3.5 προκύπτει, όπως παραπάνω, από τη συνέχεια της g . □

3.4.3 Αριθμητικό παράδειγμα. Απεικονίζουμε την αριθμητική παράσταση των διαφόρων φραγμάτων που συζητήθηκαν παραπάνω εκτιμώντας τη φασματική μεταβολή των δύο πινάκων

$$A_1 = \begin{pmatrix} 1 & 10^{-4} \\ 0 & 1 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 - 10^{-4} \end{pmatrix}$$

υπό την προϋπόθεση ότι

$$\|B_i - A_i\|_F \leq 10^{-2}, \quad i = 1, 2. \quad (3.4.21)$$

Θέτουμε

$$s_i = \sup s_{A_i}(B_i), \quad i = 1, 2,$$

όπου το supremum (= το ελάχιστο άνω φράγμα) λαμβάνεται σε σχέση με όλους τους πίνακες B_i που ικανοποιούν την (3.4.21).

Τα παραδείγματα

$$B_1 - A_1 = \begin{pmatrix} 2^{-\frac{1}{2}} 10^{-2} & 0 \\ 0 & -2^{-\frac{1}{2}} 10^{-2} \end{pmatrix}$$

$$B_2 - A_2 = \begin{pmatrix} 0 & 0 \\ -10^{-2} & 0 \end{pmatrix}$$

δείχνουν ότι

$$s_1 \geq 0.00707 \quad \text{και} \quad s_2 \geq 0.10000.$$

Χρησιμοποιώντας $M = 1.02$, $p_{\|\cdot\|_1, \|\cdot\|_F} = 2$, το φράγμα (3.4.4) αποδίδει $s_i \leq 0.57132$, $i = 1, 2$ και στις δύο περιπτώσεις.

Το φράγμα στην (3.4.6) δεν εφαρμόζεται στο A_1 , αφού ο πίνακας έχει έναν μη-γραμμικό στοιχειώδη διαιρέτη. Η σχέση $A_2 = S D S^{-1}$ ικανοποιείται για

$$S = \begin{pmatrix} \xi & \eta \\ 0 & -10^{-4}\eta \end{pmatrix},$$

όπου ξ και η είναι αυθαίρετες μη-μηδενικές σταθερές. Επιλέγοντας ξ και η έτσι ώστε ο δείκτης κατάστασης $\kappa_F(S)$ να ελαχιστοποιηθεί, η (3.4.6) αποδίδει

$$s_2 \leq 20.00001.$$

Στρεφόμενοι προς την (3.4.14), βρίσκουμε $\delta_{\|\cdot\|_F}(A_1) = 10^{-4}$, $y = 10^{-2}$, $g(y) = 0.009902$ και ως εκ τούτου

$$s_1 \leq 0.010099.$$

Για το διαγωνοποιήσιμο πίνακα A_2 , βρίσκουμε $\delta_{\|\cdot\|_F}(A_2) = 1$, $y = 10^2$, $g(y) = 9.5125$ και ως εκ τούτου

$$s_2 \leq 0.105125.$$

Είναι φανερό ότι, στο παράδειγμα που συζητήθηκε παραπάνω, το φράγμα που δόθηκε από την (3.4.14) συγκρίνεται ικανοποιητικά με τα άλλα φράγματα. Εξάλλου και το παράδειγμα δείχνει ότι αν ένας πίνακας έχει ένα μη-γραμμικό στοιχειώδες διαιρέτη, δε σημαίνει ότι θα έχει μια μεγάλη φασματική μεταβολή.

3.4.4 Σχετικά αποτελέσματα για το $v(A, B)$. Για δεδομένους πίνακες A, B , που ικανοποιούν τις υποθέσεις του Θεωρήματος 3.5, έστω δ η ποσότητα στα δεξιά της (3.4.14), δηλαδή $\delta = \frac{\delta_{\|\cdot\|}(A)}{g(y)}$. Η εκφώνηση του Θεωρήματος τότε μπορεί να ερμηνευθεί γεωμετρικά λέγοντας ότι το φάσμα του B περιέχεται στην ένωση L_δ των δίσκων

$$D_i = \{\lambda \mid |\lambda - \lambda_i| \leq \delta\}, \quad i = 1, 2, \dots, n, \quad \text{με } L_\delta = \bigcup_{i=1}^n D_i.$$

Αφού $\delta \rightarrow 0$ μονότονα καθώς $B \rightarrow A$, μπορούμε να συμπεράνουμε από το επιχείρημα συνοχής ότι κάθε συνιστώσα του L_δ περιέχει τόσες πολλές ιδιοτιμές του B όσες και του A . Από το γεγονός αυτό μπορούμε να επιτύχουμε, χρησιμοποιώντας ξανά το γνωστό επιχείρημα, το ακόλουθο αποτέλεσμα.

Θεώρημα 3.6: Έστω A, B και $\|\cdot\|$ ικανοποιούν τις υποθέσεις του Θεωρήματος 3.5 και έστω y ορίζεται όπως προηγουμένως. Τότε

$$v(A, B) \leq (2n - 1) \frac{y}{g(y)} \|B - A\| = \frac{(2n - 1) \delta_{\|\cdot\|}(A)}{g(y)} \quad (\text{αφού } y = \frac{\delta_{\|\cdot\|}(A)}{\|B - A\|}). \quad (3.4.22)$$

Θα πρέπει να σημειωθεί ότι αυτό το αποτέλεσμα δε συνεπάγεται τον τύπο (3.4.8) του Hoffmann-Wielandt.

3.4.5 Ένα αποτέλεσμα σε προσεγγιστικές ιδιοτιμές. Το ακόλουθο ερώτημα θεωρείται συχνό σε υπολογισμό πίνακα: Έστω $\|\cdot\|$ μια νόρμα διάνυσμα και έστω x ένα διάνυσμα τέτοιο ώστε $\|x\| = 1$. Έστω, επίσης, λ ένας αριθμός τέτοιος ώστε η $\|Ax - \lambda x\|$ να είναι μικρή. Το ερώτημα που δημιουργείται είναι πόσο κοντά, λοιπόν, είναι το λ σε μια ιδιοτιμή του A . (Αυτή η ερώτηση ήταν το σημείο εκκίνησης της μελέτης του Wielandt για περιοχές έγκλησης.) Ο Householder και ο Bauer απέδειξαν ένα αποτέλεσμα το οποίο ως μια σημαντική ειδική περίπτωση περιέχει το ακόλουθο: Εάν ο A είναι όμοιος μ' ένα διαγώνιο πίνακα D , δηλαδή $A = S D S^{-1}$ και εάν $\|\circ\|$ υποδηλώνει τη νόρμα πίνακα επαγόμενη από τη $\|\cdot\|$, τότε

$$|\lambda_i - \lambda| \leq \kappa(S) \|Ax - \lambda x\| \quad (3.4.23)$$

για τουλάχιστον μία ιδιοτιμή λ_i . Στη συγκεκριμένη περίπτωση το κ υποδηλώνει το δείκτη κατάστασης της $\|\circ\|$ νόρμας πίνακα που επάγεται από τη $\|\cdot\|$, όπως έχει οριστεί στην Παράγραφο 3.3.1. Το παραπάνω αποτέλεσμα έχει και αυτό μειονεκτήματα, αφού ισχύει μόνο για διαγωνοποιήσιμους πίνακες και δεν υπάρχει σαφής εκτίμηση για το δείκτη κατάστασης που έχει δοθεί. Σ' αυτήν την κατεύθυνση ο Morrison, χρησιμοποιώντας μια μέθοδο που δεν έρχεται σε αντίθεση με αυτή που χρησιμοποιήθηκε στην απόδειξη του Θεωρήματος 3.5, έλαβε το ακόλουθο αποτέλεσμα.

Θεώρημα 3.8 (D.D. Morrison): Έστω A ένας μη-κανονικός πίνακας, έστω $\|\cdot\|$ μια νόρμα διάνυσμα και έστω $\|\circ\|$ η επαγόμενη νόρμα πινάκων από τη $\|\cdot\|$. Εάν x είναι ένα διάνυσμα και λ είναι ένας μιγαδικός αριθμός τέτοιος ώστε $\|x\| = 1$ και $\|Ax - \lambda x\| \neq 0$ και εάν

$$y = \frac{\delta_{\|\circ\|}(A)}{\|Ax - \lambda x\|},$$

τότε υπάρχει μια ιδιοτιμή λ_i του A τέτοια ώστε

$$|\lambda_i - \lambda| \leq \frac{y}{g(y)} \|Ax - \lambda x\|. \quad (3.4.24)$$

3.5 Αριθμητικό Πεδίο

3.5.1 Κλασικά αποτελέσματα. Το αριθμητικό πεδίο $F(A)$ ενός πίνακα A ορίστηκε στο Κεφάλαιο 1. Με βάση τον Toeplitz έχουμε ένα κλασικό αποτέλεσμα που αναφέρει ότι για A κανονικό πίνακα ισχύει $F(A) = \text{Co}(\sigma(A))$, όπου το $\text{Co}(\sigma(A))$ δηλώνει την κυρτή θήκη των ιδιοτιμών του A . Είναι γνωστό ότι το αριθμητικό πεδίο ενός μη-κανονικού πίνακα είναι ακόμη κυρτό (Hausdorff), αν και μπορεί να επεκταθεί πέραν του $\text{Co}(\sigma(A))$. Φράγματα για το αριθμητικό πεδίο των μη-κανονικών πινάκων έχουν δοθεί από τον Farnell και από τον Parker. Αυτά τα φράγματα δεν περιορίζουν την κυρτή θήκη των ιδιοτιμών, εάν ο πίνακας προσεγγίζει έναν κανονικό πίνακα.

Η ακριβής εξίσωση του συνόρου του αριθμητικού πεδίου ενός μη-κανονικού πίνακα έχει δοθεί από τον Murnaghan και, σε πιο σαφή μορφή, από τον Kirpenhahn. Ο Kirpenhahn επίσης δίνει φράγματα για τη διάμετρο και για την περιοχή του αριθμητικού πεδίου.

3.5.2 Η απόσταση του συνόρου του αριθμητικού πεδίου από τη κυρτή θήκη των ιδιοτιμών του πίνακα A . Σ' αυτό το τμήμα οι προτάσεις που ακολουθούν θα δείξουν, για συγκεκριμένες νόρμες, την ύπαρξη μιας απλής σύνδεσης μεταξύ της maximum απόστασης του συνόρου του $F(A)$ από την κυρτή θήκη των ιδιοτιμών και της απομάκρυνσης από την κανονικότητα.

Θεώρημα 3.8: Εάν ζ είναι ένα σημείο του αριθμητικού πεδίου ενός πίνακα A , τότε υπάρχει ένα σημείο η στην κυρτή θήκη των ιδιοτιμών του A , $\text{Co}(\sigma(A))$, τέτοιο ώστε

$$|\zeta - \eta| \leq \sqrt{\frac{1-n^{-1}}{2}} \delta_{\|\cdot\|_F}(A) = \sqrt{\frac{n-1}{2n}} \delta_{\|\cdot\|_F}(A) \quad (3.5.1)$$

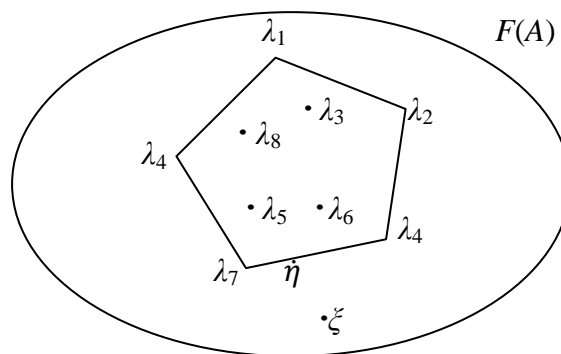
ή αν συνδυαστεί με την (3.1.7) έχουμε

$$|\zeta - \eta| \leq \sqrt{\frac{n-1}{2n}} \|M\|_F = \sqrt{\frac{n-1}{2n}} \sqrt{\|A\|_F^2 - \sum_{i=1}^n |\lambda_i|^2}.$$

Η σταθερά $\left[\frac{1-n^{-1}}{2}\right]^{\frac{1}{2}}$ δεν μπορεί να αντικατασταθεί από οποιαδήποτε άλλη μικρότερη σταθερά.

(Σημείωση: Το φράγμα (3.5.1) έχει βρεθεί ξεχωριστά από τον P.J. Eberlein.)

Για να κατανοήσουμε καλύτερα το θεώρημα “κατασκευάζουμε” το παρακάτω σχήμα



Σχήμα 11: Το αριθμητικό πεδίο $F(A)$ με την κυρτή θήκη και τις ιδιοτιμές του A , λ_i , $i = 1, 2, \dots, 8$.

Απόδειξη: Έστω $A = U T U^*$, όπου U είναι ορθομοναδιαίος πίνακας και $T = D + M$ είναι μια Schur τριγωνική μορφή του A . Έστω, επίσης, $\zeta = x^* A x$, όπου $x^* x = 1$. Θέτοντας $U^* x = y$, έχουμε $y^* y = (U^* x)^* (U^* x) = x^* U U^* x = x^* x = 1$ και

$$\zeta = x^* (U T U^*) x = (U^* x)^* (D + M) (U^* x) = y^* (D + M) y = y^* D y + y^* M y.$$

Άρα

$$\zeta = \eta + y^* M y, \quad (3.5.2)$$

όταν $\eta = y^* D y$ είναι ένα σημείο στο $\text{Co}(\sigma(A))$. Ως εκ τούτου για αυτόν τον αριθμό η ισχύει

$$|\zeta - \eta| = |y^* M y|.$$

Σημείωση: Το η δεν είναι το πλησιέστερο στο $\zeta \in \text{Co}(\sigma(A))$ σημείο της κυρτής θήκης του φάσματος.

Θέτοντας $M = (m_{ij})$ και $y = (y_i)$, βρίσκουμε από την ανισότητα του Cauchy

$$|y^* M y|^2 = |\sum_{i < j} m_{ij} \bar{y}_i y_j|^2 \leq \sum_{i < j} |m_{ij}|^2 \sum_{i < j} |\bar{y}_i y_j|^2 = [\delta_{\|\cdot\|_F}(A)]^2 \sum_{i < j} |\bar{y}_i y_j|^2.$$

Για τον τελευταίο παράγοντα βρίσκουμε την εκτίμηση

$$\begin{aligned} \sum_{i < j} |\bar{y}_i y_j|^2 &= \sum_{i < j} |y_i|^2 |y_j|^2 = \frac{1}{2} \{(\sum_{i=1}^n |y_i|^2)^2 - \sum_{i=1}^n |y_i|^4\} \\ &\leq \frac{1}{2} \{(\sum_{i=1}^n |y_i|^2)^2 - \frac{1}{n} (\sum_{i=1}^n |y_i|^2)^2\} = \frac{1}{2} \left(1 - \frac{1}{n}\right). \end{aligned}$$

Έτσι, θα βρούμε

$$|y^* M y| \leq \sqrt{\frac{1-n^{-1}}{2}} \delta_{\|\cdot\|_F}(A)$$

αποδεικνύοντας την (3.5.1).

Προκειμένου να δείξουμε ότι η ισότητα επιτυγχάνεται θεωρούμε οποιοδήποτε πίνακα της μορφής $A = \lambda I + M$, όπου λ είναι ένας μιγαδικός αριθμός και $m_{ij} = c$, $i < j$, όπου $c > 0$. Έχουμε $\delta_{\|\cdot\|_F}(A) = \|M\|_F = (\sum_{i < j} c^2)^{\frac{1}{2}} = [\frac{1}{2} n(n-1)]^{\frac{1}{2}} c$. Η κυρτή θήκη $\text{Co}(\sigma(A))$ τείνει στο σημείο λ . Επιλέγοντας $y_i = n^{-\frac{1}{2}}$ ($i = 1, \dots, n$) βρίσκουμε ότι το σημείο $y^* A y$ του $F(A)$ απέχει από το $\text{Co}(\sigma(A))$ την απόσταση

$$y^* M y = \sum_{i < j} \frac{c}{n} = \frac{c}{n} \frac{1}{2} n(n-1) = \frac{c}{2} (n-1) = \sqrt{\frac{1-n^{-1}}{2}} \delta_{\|\cdot\|_F}(A),$$

όπως επιθυμούσαμε. □

Ένα παρόμοιο αποτέλεσμα επίσης ισχύει για τη l_1 -νόρμα.

Θεώρημα 3.9: Εάν ζ είναι ένα σημείο του αριθμητικού πεδίου ενός πίνακα A , τότε υπάρχει ένα σημείο η στην κυρτή θήκη των ιδιοτιμών του A , $\text{Co}(\sigma(A))$, έτσι ώστε να ισχύει η ακόλουθη ανισότητα

$$|\zeta - \eta| \leq \frac{1}{2} \delta_{\|\cdot\|_{l_1}}(A). \quad (3.5.3)$$

Η σταθερά $\frac{1}{2}$ δεν μπορεί να αντικατασταθεί από οποιαδήποτε άλλη μικρότερη σταθερά. Η σταθερά αυτή είναι η βέλτιστη δυνατή.

Απόδειξη: Αναπαριστούμε ξανά το ζ στη μορφή (3.5.2), αλλά αυτή τη φορά εκτιμούμε το $y^* M y$ ως εξής

$$|y^* M y| = \left| \sum_{i < j} m_{ij} \bar{y}_i y_j \right| \leq \|M\|_{l_1} \max_{i < j} |\bar{y}_i y_j|.$$

Αφού $\sum_{i=1}^n |y_i|^2 = 1$, το $\max_{i < j} |\bar{y}_i y_j|$ γίνεται μέγιστο, εάν το διάνυσμα y έχει ακριβώς δύο μη-μηδενικές συνιστώσες. Από την ανισότητα αριθμητικού-γεωμετρικού μέσου έχουμε

$$\begin{aligned} \sqrt{|\bar{y}_i y_j|} &\leq \frac{|\bar{y}_i| + |y_j|}{2} \Rightarrow |\bar{y}_i y_j| \leq \left(\frac{|\bar{y}_i| + |y_j|}{2} \right)^2 \Rightarrow \\ \Rightarrow |\bar{y}_i y_j| &\leq \frac{|\bar{y}_i|^2 + |y_j|^2 + 2|\bar{y}_i y_j|}{4} \Rightarrow 4|\bar{y}_i y_j| \leq |\bar{y}_i|^2 + |y_j|^2 + 2|\bar{y}_i y_j| \Rightarrow \\ &\Rightarrow 2|\bar{y}_i y_j| \leq |\bar{y}_i|^2 + |y_j|^2 \Rightarrow 2|\bar{y}_i y_j| \leq 1 \Rightarrow \max_{i < j} |\bar{y}_i y_j| \leq \frac{1}{2}. \end{aligned}$$

Η τιμή του maximum τότε είναι $\frac{1}{2}$. Έτσι, παίρνουμε

$$|y^* M y| \leq \frac{1}{2} \|M\|_{l_1}. \quad (3.5.4)$$

Η ανισότητα (3.5.3) τώρα ακολουθεί ενόψει του γεγονότος ότι η (3.5.4) είναι αληθής για όλα τα M που μπορούν να εμφανιστούν σε μια Schur τριγωνική μορφή στο A .

Προκειμένου να δείξουμε ότι η ισότητα επιτυγχάνεται στην (3.5.3) για μη-μηδενικά $\delta_{\|\cdot\|_{l_1}}(A)$, έστω A οποιοσδήποτε πίνακας της μορφής $A = \lambda I + M$, όπου λ είναι ένας μιγαδικός αριθμός, $m_{12} = m > 0$ και $m_{ij} = 0$ για κάθε άλλη περίπτωση.

Όπως παραπάνω, το $\text{Co}(\sigma(A))$ τείνει στο σημείο λ . Επιλέγουμε τώρα το μοναδιαίο διάνυσμα y έτσι ώστε $y_1 = y_2 = 2^{-\frac{1}{2}}$ και $y_i = 0$ για $i = 3, \dots, n$. Το διάνυσμα αυτό “αποδίδει” ένα σημείο $y^* A y$ του $F(A)$ για το οποίο η απόσταση από το $\text{Co}(\sigma(A))$ είναι

$$y^* M y = \frac{1}{2} m = \frac{1}{2} \|M\|_{l_1}. \quad (3.5.5)$$

Έστω τώρα $M' = (m'_{ij})$ τα εκτός διαγωνίου στοιχεία οποιασδήποτε άλλης Schur τριγωνικής μορφής του A . Ενόψει της

$$\|M'\|_{l_1} \geq \|M'\|_F = \|M\|_F = m,$$

έχουμε $\delta_{\|\cdot\|_{l_1}}(A) = m$ και η (3.5.5) μπορεί να αντικατασταθεί με

$$y^* M y = \frac{1}{2} \delta_{\|\cdot\|_{l_1}}(A),$$

δείχνοντας ότι η ισότητα μπορεί να ισχύσει στην (3.5.3). □

3.5.3 Ένα βελτιωμένο φράγμα για τα αριθμητικά πεδία. Ένας στοιχειώδης υπολογισμός δείχνει ότι για έναν 2×2 πίνακα με Schur τριγωνική μορφή

$$T = \begin{pmatrix} \lambda_1 & m_{12} \\ 0 & \lambda_2 \end{pmatrix}$$

το αριθμητικό πεδίο έχει δοθεί από το εσωτερικό και από το σύνορο της έλλειψης με εστίες λ_1 και λ_2 και μικρό (δευτερεύων) ημιάξονα $\frac{1}{2}|m_{12}|$. Μόνο δύο σημεία του αριθμητικού πεδίου επιτυγχάνουν τη μέγιστη απόσταση $\frac{1}{2}|m_{12}|$ από το $\text{Co}(\sigma(A))$, όπως δείξαμε και στα Θεωρήματα 3.8 και 3.9. Όλα τα άλλα σημεία έχουν μια μικρότερη απόσταση.

Το παραπάνω γεγονός γενικεύεται από το Θεώρημα 3.10. Για $\lambda, \mu \in \mathbb{C}$ και $c \in \mathbb{R}$, όπου $c \geq 0$, συμβολίζουμε με $E(\lambda, \mu; c)$ συμπαγές σύνολο σημείων που στο μιγαδικό επίπεδο ορίζεται από την έλλειψη με εστίες λ, μ και μικρό ημιάξονα c .

Θεώρημα 3.10: Έστω $\lambda_i, i = 1, 2, \dots, n$, n μιγαδικοί αριθμοί και έστω $\alpha \geq 0$. Τότε το αριθμητικό πεδίο οποιουδήποτε πίνακα A με ιδιοτιμές λ_i και με $\delta_{\|\cdot\|_{l_1}}(A) = \alpha$ περιέχεται στην κυρτή θήκη των $\frac{n(n-1)}{2}$ συνόλων

$$E(\lambda_i, \lambda_j, \frac{1}{2}\alpha), \quad 1 \leq i < j \leq n.$$

Απόδειξη: Έστω ένας πίνακας A που ικανοποιεί την υπόθεση και έστω ζ, y_i, D και $M = (m_{ij})$ ορίζονται όπως στην απόδειξη του Θεωρήματος 3.8. Θέτουμε

$$m = \|M\|_{l_1} = \sum_{i < j} |m_{ij}|.$$

Εάν $m = 0$, τότε ο πίνακας A είναι κανονικός και το θεώρημα ισχύει τετριμμένα αφού το αριθμητικό πεδίο ενός κανονικού πίνακα είναι η κυρτή θήκη του φάσματός του (Πρόταση 1.7.5). Ως εκ τούτου, υποθέτουμε ότι $m > 0$ και θέτουμε

$$\theta_{ij} = \begin{cases} \frac{m_{ij}}{|m_{ij}|}, & \text{για } m_{ij} \neq 0 \\ 0, & \text{για } m_{ij} = 0 \end{cases}.$$

Λαμβάνουμε την παράσταση

$$\begin{aligned} \zeta &= \sum_{k=1}^n \lambda_k |y_k|^2 + \sum_{i < j} m_{ij} \bar{y}_i y_j \\ &= \sum_{i < j} |m_{ij}| \left\{ \theta_{ij} \bar{y}_i y_j + \frac{1}{m} \sum_{k=1}^n \lambda_k |y_k|^2 \right\} \\ &= \frac{1}{m} \sum_{i < j} |m_{ij}| \left\{ m \theta_{ij} \bar{y}_i y_j + \sum_{k=1}^n \lambda_k |y_k|^2 \right\}, \end{aligned}$$

δείχνοντας ότι το ζ ανήκει στην κυρτή θήκη των σημείων

$$\eta_{ij} = m \theta_{ij} \bar{y}_i y_j + \sum_{k=1}^n \lambda_k |y_k|^2, \quad 1 \leq i < j \leq n.$$

Θα δείξουμε στη συνέχεια ότι τα σημεία η_{ij} ανήκουν στο $Co(\sigma_m(A))$, δηλαδή στην κυρτή θήκη των συνόλων

$$E(\lambda_k, \lambda_l; \frac{1}{2}m), \quad 1 \leq k < l = n.$$

Εάν $|y_i|^2 + |y_j|^2 = 0$, τότε είναι προφανές.

Εάν $|y_i|^2 + |y_j|^2 = 1$, τότε $y_k = 0, k \neq i, j$ και το η_{ij} ανήκει στην $E(\lambda_i, \lambda_j; \frac{1}{2}m)$ και έτσι στο $Co(\sigma_m(A))$.

Εάν $0 < |y_i|^2 + |y_j|^2 < 1$, ορίζουμε τις μεταβλητές

$$z_k = \begin{cases} \frac{y_k}{\sqrt{|y_i|^2 + |y_j|^2}}, & \text{για } k = i, j, \\ \frac{y_k}{\sqrt{1 - |y_i|^2 - |y_j|^2}}, & \text{για } k \neq i, j \end{cases}.$$

Μπορούμε να γράψουμε τότε

$$\eta_{ij} = (|y_i|^2 + |y_j|^2) (\lambda_i |z_i|^2 + \lambda_j |z_j|^2 + \theta_{ij} m \bar{z}_i z_j) + (1 - |y_i|^2 - |y_j|^2) \sum_{k \neq i, j}^n \lambda_k |z_k|^2.$$

Η τελευταία σχέση δείχνει ότι το η_{ij} ανήκει στην κυρτή θήκη του συνόλου $E(\lambda_i, \lambda_j; \frac{1}{2}m)$ και τα σημεία λ_k ($k \neq i, j$) ανήκουν όλα στο $Co(\sigma_m(A))$. Έτσι ξανά το η_{ij} ανήκει στο $Co(\sigma_m(A))$. Προκύπτει επίσης ότι το ζ ανήκει στο $Co(\sigma_m(A))$.

Το παραπάνω είναι σε ισχύ για οποιοδήποτε πίνακα M που μπορεί να εμφανιστεί σε μια Schur τριγωνική μορφή του A . Έτσι το ζ ανήκει στη τομή όλων των πιθανών συνόλων $Co(\sigma_m(A))$. Αφού τα σύνολα $Co(\sigma_m(A))$ είναι κλειστά και αφού εξαρτώνται συνεχώς από το m , προκύπτει ότι το ζ ανήκει στο $Co(\sigma_{\inf m}(A))$, δηλαδή, στο $Co(\sigma(A))$. □

Κεφάλαιο 4: Πίνακες με Ιδιοπρόβλημα Κακής - Κατάστασης

4.1 Παρατηρήσεις σε Πίνακες με Κακής - Κατάστασης Ιδιοπρόβλημα

4.1.1 *Περίληψη.* Σ' αυτό το κεφάλαιο θα δώσουμε ένα φράγμα για την απόσταση ενός πίνακα που έχει ένα κακής-κατάστασης πρόβλημα ιδιοτιμής από έναν πίνακα που έχει πολλαπλή ιδιοτιμή.

Όπως αναφέραμε στο Αλγεβρικό Πρόβλημα Ιδιοτιμής, ο Wilkinson είχε συζητήσει την ευαισθησία μιας απλής ιδιοτιμής λ_i ενός πίνακα A όσον αφορά τα αντίστοιχα δεξιά και αριστερά ιδιοδιανύσματα x_i και y_i . Μια ποσότητα $|s_i|$ ορίζεται από τη σχέση

$$|s_i| = |y_i^* x_i| / \|y_i\|_2 \|x_i\|_2. \quad (4.1.1)$$

Θα ήταν βολικό να υποθέσουμε ότι $\|y_i\|_2 = \|x_i\|_2 = 1$ έτσι ώστε $s_i = |y_i^* x_i|$ και σαφώς τα διανύσματα μπορούν να επιλεγούν έτσι ώστε $s_i = y_i^* x_i$.

Ο Wilkinson διατύπωσε την εικασία ότι όταν το s_i είναι “μικρό”, τότε υπάρχει ένας πίνακας $A + F$ με πολλαπλή ιδιοτιμή τέτοιος ώστε το $\|F\|_2 / \|A\|_2$ να είναι επίσης μικρό. Λίγο αργότερα, ο Ruhe έδειξε ότι αυτό είναι όντως αλήθεια. Στη συνέχεια, θα δώσουμε ένα φράγμα για την απόσταση του πλησιέστερου πίνακα που έχει μια πολλαπλή ιδιοτιμή.

Παρατηρούμε ότι μέσω μιας διαγώνιας ομοιότητας ένας πίνακας $B = D^{-1} A D$ θα μπορούσε να αποκτήσει αυθαίρετα μικρά s_i , ενώ το αποτέλεσμα μας θα εξακολουθεί να ισχύει για τον πίνακα B . Εννοείται φυσικά ότι, όταν ένα μικρό s_i “επάγεται” μ' αυτόν τον τρόπο, η $\|B\|_2$ θα είναι τεχνητά μεγάλη. Το αποτέλεσμα είναι πιο έγκυρο, όταν ο δεδομένος πίνακας A “εξισορροπείται” με την έννοια ότι $\|A\|_2 \leq \|D^{-1} A D\|_2$ για όλους τους αντιστρέψιμους διαγώνιους D .

{ Έστω x, y το δεξιό και αριστερό ιδιοδιάνυσμα του πίνακα A , αντίστοιχα. Έχουμε

$$Ax = \lambda x \Rightarrow DBD^{-1}x = \lambda x \Rightarrow B(D^{-1}x) = \lambda(D^{-1}x) \text{ και}$$

$$y^*A = \lambda y^* \Rightarrow y^*DBD^{-1} = \lambda y^* \Rightarrow (y^*D)B = \lambda(y^*D).$$

Θέτουμε

$$\hat{y}^* = y^*D \Leftrightarrow \hat{y} = D^*y \text{ και } \hat{x} = D^{-1}x.$$

Τότε

$$\langle \hat{x}, \hat{y} \rangle = \hat{y}^* \hat{x} = y^* D D^{-1} x = y^* x.$$

Ο δείκτης κατάστασης της απλής ιδιοτιμής λ για τον πίνακα B για τη 2-νόρμα είναι

$$\kappa_B(\lambda) = \frac{\|\hat{x}\|_2 \|\hat{y}\|_2}{|\hat{y}^* \hat{x}|} = \frac{\|D^{-1}x\|_2 \|D^*y\|_2}{|y^*x|}.$$

Για $\|D^{-1}x\|_2 \rightarrow \infty$ και για $\|D^*y\|_2 \rightarrow \infty$ παρατηρούμε ότι και ο δείκτης κατάστασης θα τείνει στο ∞ . }

Η απόδειξη ξεκινά εξετάζοντας την περίπτωση $s = y^*x = 0$, όπου

$$Ax = \lambda x, \quad y^*A = \lambda y^*. \quad (4.1.2)$$

Έστω P ένας ορθομοναδιαίος πίνακας τέτοιος ώστε

$$Px = e_1, \quad (4.1.3)$$

όπου e_1 είναι η πρώτη στήλη του I . Γράφοντας

$$B = P A P^* \quad (4.1.4)$$

έχουμε

$$P A P^* P x = \lambda P x, \quad y^* P^* P A P^* = \lambda y^* P^*. \quad (4.1.5)$$

Η παραπάνω σχέση δίνει

$$B e_1 = \lambda e_1 \text{ και } z^* B = \lambda z^*, \quad (4.1.6)$$

όπου

$$z = P y. \quad (4.1.7)$$

Η σχέση $y^*x = 0$ συνεπάγεται ότι

$$0 = y^* P^* P x = z^* e_1 \quad (4.1.8)$$

και ως εκ τούτου

$$z_1 = 0. \quad (4.1.9)$$

Από την εξίσωση (4.1.6), ο B είναι της μορφής

$$B = \left(\begin{array}{c|c} \lambda & b^* \\ \hline 0 & B_1 \end{array} \right), \quad (4.1.10)$$

ενώ η εξίσωση (4.1.9) συνεπάγεται ότι το z^* είναι της μορφής

$$z^* = (0 | w^*), \quad (4.1.11)$$

όπου w είναι ένα μοναδιαίο διάνυσμα με $n - 1$ συνιστώσες. Από την εξίσωση (4.1.6) έχουμε

$$(0 | w^*) \left(\frac{\lambda}{0} \left| \begin{array}{c} b^* \\ B_1 \end{array} \right. \right) = \lambda (0 | w^*) \quad (4.1.12)$$

δίνοντας

$$w^* B_1 = \lambda w^*. \quad (4.1.13)$$

Ως εκ τούτου, το λ είναι μια ιδιοτιμή του B_1 δείχνοντας ότι το λ είναι μια ιδιοτιμή του B πολλαπλότητας τουλάχιστον δύο. Το ίδιο ισχύει επομένως και για το A .

Εξετάζουμε τώρα την περίπτωση, όταν

$$y^* x = \varepsilon, \quad (4.1.14)$$

όπου ενδιαφερόμαστε ιδιαίτερα για μικρές τιμές του ε . Έστω ξανά $Px = e_1$ και $Py = z$ και προχωρούμε όπως πριν. Τώρα έχουμε

$$z^* e_1 = \varepsilon \text{ και } z^* = (\varepsilon | w^*), \quad (4.1.15)$$

όπου $\|w\|_2^2 = 1 - \varepsilon^2$. Η εξίσωση (4.1.12), επομένως, γίνεται

$$(\varepsilon | w^*) \left(\frac{\lambda}{0} \left| \begin{array}{c} b^* \\ B_1 \end{array} \right. \right) = \lambda (\varepsilon | w^*) \quad (4.1.16)$$

δίνοντας

$$\varepsilon b^* + w^* B_1 = \lambda w^*, \quad (4.1.17)$$

το οποίο μπορεί να γραφτεί στη μορφή

$$\begin{aligned} w^* \left[B_1 + \frac{\varepsilon w b^*}{w^* w} \right] &= \lambda w^* \Rightarrow \\ \Rightarrow w^* \left[B_1 + \frac{\varepsilon w b^*}{w^* w} - \lambda I \right] &= 0 \Rightarrow \\ \Rightarrow w^* \left[\lambda I - \left(B_1 + \frac{\varepsilon w b^*}{w^* w} \right) \right] &= 0. \quad (4.1.18) \end{aligned}$$

Ως εκ τούτου, το λ είναι μια ιδιοτιμή του $B_1 + \frac{\varepsilon w b^*}{w^* w}$ και

$$\begin{aligned} \|\varepsilon w b^* / w^* w\| &= \varepsilon \|w\|_2 \|b\|_2 / \|w\|_2^2 \\ &\leq \varepsilon \|B\|_2 / (1 - \varepsilon^2)^{\frac{1}{2}} = \varepsilon \|A\|_2 / (1 - \varepsilon^2)^{\frac{1}{2}}. \end{aligned} \quad (4.1.19)$$

Υπάρχει, επομένως, ένας πίνακας $B + E$, έχοντας το λ ως μια ιδιοτιμή πολλαπλότητας τουλάχιστον δύο. Έτσι δείχνουμε ότι ο $A + F$ με

$$F = P^* E P, \quad \|F\|_2 = \|E\|_2 \leq \varepsilon \|A\|_2 / (1 - \varepsilon^2)^{\frac{1}{2}} \quad (4.1.20)$$

έχει την ίδια ιδιότητα.

Παρατηρούμε ότι συχνά θα υπάρχει ένας πολύ πιο κοντινός πίνακας, που θα έχει μια πολλαπλή ιδιοτιμή. Να σημειώσουμε σ' αυτό το σημείο ότι στην εξίσωση (4.1.19) έχουμε χρησιμοποιήσει την ανισότητα

$$\|b\|_2 \leq \|B\|_2 \quad (4.1.21)$$

και αυτό θα μπορούσε να είναι μακριά από ακρίβειες. Πράγματι, εάν ο A είναι κανονικός, το b είναι μηδενικό και το E είναι, επομένως, μηδενικό, υποθέτοντας ότι $\varepsilon \neq 1$. Αυτό είναι προφανές δεδομένου ότι διαφορετικά εάν το λ είναι μια απλή ρίζα ενός κανονικού πίνακα, τότε $y = x$ και το $y^* x$ πρέπει να είναι μονάδα.

Έχουμε παρουσιάσει ένα γειτονικό πίνακα που έχει το λ ως την πολλαπλή ιδιοτιμή. Γενικά θα υπάρξει ένας κοντινότερος πίνακας, ο οποίος θα έχει μια πολλαπλή ιδιοτιμή, όχι ακριβώς ίση με το λ . Για παράδειγμα για τις ιδιοτιμές του ακόλουθου πίνακα

$$A = \begin{bmatrix} 1 & \theta \\ 1 & 1 \end{bmatrix} \quad (4.1.22)$$

έχουμε:

$$\begin{aligned} \det\left[\lambda \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} - \begin{bmatrix} 1 & \theta \\ 1 & 1 \end{bmatrix}\right] &= 0 \Rightarrow \begin{vmatrix} \lambda - 1 & -\theta \\ -1 & \lambda - 1 \end{vmatrix} = 0 \Rightarrow \\ &\Rightarrow (\lambda - 1)^2 - \theta = 0 \Rightarrow \lambda_{1,2} = 1 \pm \theta^{\frac{1}{2}}. \end{aligned}$$

Άρα έχει μια ιδιοτιμή $1 + \theta^{\frac{1}{2}}$. Τα αντίστοιχα αριστερά και δεξιά διανύσματα είναι

$$\begin{bmatrix} 1, & \theta^{\frac{1}{2}} \end{bmatrix} \quad \text{και} \quad \begin{bmatrix} \theta^{\frac{1}{2}}, & 1 \end{bmatrix} \quad (4.1.23)$$

δίνοντας $s = 2\theta^{\frac{1}{2}} / (1 + \theta) \equiv \varepsilon$ από το παραπάνω αποτέλεσμα. Εάν το στοιχείο που βρίσκεται στην πρώτη γραμμή και στη δεύτερη στήλη του πίνακα A αλλάξει από θ σε 0, ο πίνακας θα έχει τη μονάδα ως μια διπλή ιδιοτιμή. Η πολλαπλότητα έχει προέλθει από μια διατάραξη, η οποία είναι της τάξης ε^2 και όχι ε .

4.2 Ιδιότητες ενός Πίνακα με ένα Κακής - Κατάστασης Ιδιοπρόβλημα

4.2.1 *Περίληψη.* Έχει αποδειχθεί ότι ένας πίνακας που έχει έναν πολύ κακής - κατάστασης ιδιοδιάνυσμα πίνακα είναι κοντά σε έναν που έχει πολλαπλές ιδιοτιμές και μια εκτίμηση αυτής της απόστασης που έχει δοθεί με την Ευκλείδεια νόρμα πίνακα.

Πρόκειται να εξάγουμε ένα θεώρημα που αφορά τη λύση του ιδιοπροβλήματος για πίνακες για τους οποίους αυτό το πρόβλημα είναι κακής-κατάστασης.

Είναι πολύ γνωστό το γεγονός ότι ένας πίνακας ο οποίος είναι κοντά σε έναν με πολλαπλές ιδιοτιμές έχει ένα κακής - κατάστασης ιδιοπρόβλημα. Το παρακάτω Θεώρημα θα δείξει μια αντιστροφή σ' αυτό, δηλαδή εάν ένας πίνακας έχει ένα κακής-κατάστασης ιδιοπρόβλημα, είναι επίσης κοντά σ' έναν πίνακα που έχει πολλαπλές ιδιοτιμές. Η κακή-κατάσταση μπορεί να υποδεικνύεται από σχεδόν ορθογώνια αριστερά και δεξιά ιδιοδιανύσματα. Ένας πίνακας έχοντας μερικές ιδιοτιμές πολύ κοντά, είναι, φυσικά, κοντά σ' έναν πίνακα με πολλαπλές ιδιοτιμές. Το θεώρημά μας, όμως, καθιερώνει ότι ακόμη και αν ένας πίνακας έχει καλά διαχωριζόμενες ιδιοτιμές είναι κοντά σ' έναν που έχει πολλαπλές ιδιοτιμές στην περίπτωση που έχουμε ένα κακής-κατάστασης ιδιοδιάνυσμα πίνακα.

4.2.2 *Εισαγωγή και Συμβολισμός.* Έστω A είναι ένας $n \times n$ μιγαδικός πίνακας. Οι ιδιάζουσες τιμές $\sigma_k(A)$ είναι οι ρίζες των ιδιοτιμών του πίνακα A^*A , διατεταγμένες έτσι ώστε $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r \geq 0$. Θα χρησιμοποιήσουμε την Ευκλείδεια νόρμα $\|x\|_2 = (\sum |x_i|^2)^{\frac{1}{2}}$ και τις νόρμες πίνακα

$$\|A\|_2 := \sup_{\|x\|_2=1} \|Ax\|_2 = \sigma_1(A),$$

$$\|A\|_F := \sqrt{\text{trace}(A^*A)} = (\sum |a_{ij}|^2)^{\frac{1}{2}}.$$

Εάν x_1, \dots, x_k είναι αυθαίρετα διανύσματα, συμβολίζουμε τη γραμμική τους θήκη με $[x_1, x_2, \dots, x_k]$. Εάν ο S είναι ένας γραμμικός υπόχωρος του n -διάστατου μοναδιαίου χώρου και x είναι ένα μοναδιαίο διάνυσμα, δηλαδή $\|x\|_2 = 1$, τότε ορίζουμε τη γωνία μεταξύ του διανύσματος και του υποχώρου ως

$$\sin\theta(x, S) := \min_{y \in S} \|x - y\|_2. \quad (4.2.1)$$

Έστω $\Lambda := \text{diag}(\lambda_i)$ είναι ο πίνακας των ιδιοτιμών του A . Εάν ο A είναι διαγωνοποιήσιμος με διακριτές ιδιοτιμές, τότε, για έναν κατάλληλο πίνακα H , ισχύει

$$A = H\Lambda H^{-1}. \quad (4.2.2)$$

Οι στήλες του H (συμβολίζονται με x_k) είναι τα δεξιά ιδιοδιανύσματα του A και οι γραμμές του H^{-1} (συμβολίζονται με y_k^*) είναι τα αριστερά ιδιοδιανύσματα του A . Οι δείκτες κατάστασης συμβολίζονται με $s_k(A) := y_k^* x_k$, όπου τα y_k και x_k είναι βαθμονομημένα έτσι ώστε: $\|y_k\|_2 = \|x_k\|_2 = 1$. Οι ιδιότητες αυτών των αριθμών μπορούν να μελετηθούν. Ειδικότερα μπορεί να αποδειχθεί ότι

$$|s_k(A)| = \sin \theta (x_k, [x_1, \dots, x_{k-1}, x_{k+1}, \dots, x_n]). \quad (4.2.3)$$

Για κάθε πίνακα A υπάρχει ένας ορθομοναδιαίος πίνακας U , που τον μετατρέπει σε Schur τριγωνική μορφή, T , δηλαδή

$$A = UTU^* = U(A+M)U^*. \quad (4.2.4)$$

Ο M είναι αυστηρά άνω τριγωνικός και εάν $M = 0$, ο A είναι ένας κανονικός πίνακας. Επιπλέον οι ιδιοτιμές μπορούν να “κατασκευαστούν” έτσι ώστε να εμφανιστούν σε οποιαδήποτε επιθυμητή σειρά κατά μήκος της διαγωνίου του T .

Τελικά, έστω Φ οποιοδήποτε σύνολο πινάκων. Ορίζουμε την απόσταση σε σχέση με τη νόρμα $\|\cdot\|$ μεταξύ ενός πίνακα A και του συνόλου Φ με

$$\text{dist}_{\|\cdot\|}(A, \Phi) := \inf_{X \in \Phi} \|A - X\|.$$

4.2.3 Το Θεώρημα. Τώρα θα μελετήσουμε την απόσταση μεταξύ ενός δεδομένου πίνακα και του συνόλου \mathcal{M} των πινάκων που έχουν τουλάχιστον μια πολλαπλή ιδιοτιμή. Δηλώνουμε το ακόλουθο θεώρημα το οποίο επισημαίνει μια περίπτωση στην οποία ένας πίνακας είναι κοντά στο \mathcal{M} .

Θεώρημα 4.1: Έστω A είναι ένας $n \times n$ πίνακας με διακριτές ιδιοτιμές $\lambda_1, \dots, \lambda_n$ και ορίζουμε

$$d := \max |\lambda_i - \lambda_j|. \quad (4.2.5)$$

(Δηλαδή η συγκεκριμένη ποσότητα μπορεί να θεωρηθεί ως διάμετρος του φάσματος.)

Τότε υπάρχει μια γωνία θ τέτοια ώστε

$$\text{dist}_{\|\cdot\|_F}(A, \mathcal{M}) \leq \frac{n}{4} d \tan \theta, \quad (4.2.6)$$

όπου

$$\sin \theta \leq \min_k |s_k(A)|^{\frac{1}{n-1}}.$$

Απόδειξη: Ξεκινώντας με $i = n$, διατάσσουμε τα κανονικοποιημένα ιδιοδιανύσματα x_i έτσι ώστε

$$\theta(x_i, [x_1, \dots, x_{i-1}]) = \min_{1 \leq j \leq i} \theta(x_j, [x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_i]),$$

$$i = n, n-1, \dots, 1. \quad (4.2.7)$$

Τώρα μπορούμε να γράψουμε τον πίνακα ιδιοδιανυσμάτων $H = (x_1, \dots, x_n)$ ως γινόμενο ενός ορθομοναδιαίου και ενός άνω τριγωνικού πίνακα με θετικά διαγώνια στοιχεία, δηλαδή

$$H = QR \quad (4.2.8)$$

και έχουμε μια Schur τριγωνική μορφή T του A στη (4.2.4) που ορίζεται με

$$T := Q^* A Q = R A R^{-1}. \quad (4.2.9)$$

Η παραπάνω σχέση προκύπτει λαμβάνοντας υπόψη και το γεγονός ότι οι πίνακες R , A και R^{-1} είναι άνω τριγωνικοί.

Οφείλουμε να σημειώσουμε ότι

$$r_{ii} = \sin \theta(x_i, [x_1, \dots, x_{i-1}]) = \min_{y \in [x_1, \dots, x_{i-1}]} \|x_i - y\|_2. \quad (4.2.10)$$

Η (4.2.3) και η επιλογή (4.2.7) συνεπάγονται ότι

$$1 = r_{11} \geq r_{22} \geq \dots \geq r_{nn} = |s_n(A)| = \min_k |s_k(A)|. \quad (4.2.11)$$

Πρόκειται, τώρα, να μελετήσουμε τα στοιχεία του T που βρίσκονται πάνω από τη διαγώνιο, να τα συσχετίσουμε με τις ιδιοτιμές και τα ιδιοδιανύσματα και να δείξουμε ότι, μετά την αναδιάταξη μερικών ιδιοτιμών, υπάρχει ένα στοιχείο t_{kk+1} που δεν είναι μικρό. Αυτό σημαίνει ότι μια μικρή διαταραχή E είναι ικανή να παράγει μια διπλή ιδιοτιμή στο $T + E$.

Ορίζουμε $\theta > 0$ με

$$\sin \theta := \min_i r_{i+1, i+1} / r_{i, i}. \quad (4.2.12)$$

Τότε η (4.2.11) συνεπάγεται ότι

$$\sin \theta \leq \left(\frac{r_{22}}{r_{11}} \cdot \dots \cdot \frac{r_{nn}}{r_{n-1, n-1}} \right)^{\frac{1}{n-1}} = \left(\frac{r_{nn}}{r_{11}} \right)^{\frac{1}{n-1}} = (r_{nn})^{\frac{1}{n-1}} = |s_n(A)|^{\frac{1}{n-1}}.$$

Για την παραπάνω σχέση στηριχθήκαμε στο γεωμετρικό μέσο και στο γεγονός ότι $r_{11} = 1$.

Επιλέγω k έτσι ώστε $r_{k+1, k+1} / r_{ii} = \sin \theta$ και διαμερίζουμε τους πίνακες μετά την k^{η} γραμμή και στήλη. Η εξίσωση (4.2.9) θα γραφτεί ως εξής

$$\begin{aligned} \begin{bmatrix} T_{11} & T_{12} \\ 0 & T_{22} \end{bmatrix} &= \begin{bmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{bmatrix} \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \end{bmatrix} \begin{bmatrix} R_{11}^{-1} & -R_{11}^{-1}R_{12}R_{22}^{-1} \\ 0 & R_{22}^{-1} \end{bmatrix} \\ &= \begin{bmatrix} R_{11}\Lambda_1R_{11}^{-1} & -R_{11}(\Lambda_1R_{11}^{-1}R_{12} - R_{11}^{-1}R_{12}\Lambda_2)R_{22}^{-1} \\ 0 & R_{22}\Lambda_2R_{22}^{-1} \end{bmatrix}. \end{aligned}$$

Ορίζοντας $Y_{12} := R_{11}^{-1}R_{12}$ βλέπουμε ότι

$$t_{k, k+1} = -r_{kk} y_{k, k+1} (\lambda_k - \lambda_{k+1}) r_{k+1, k+1}^{-1}. \quad (4.3.13)$$

Συμβολίζουμε την πρώτη στήλη του Y_{12} με y_{k+1} και του R_{12} με r_{k+1} . Εάν ορίσουμε τη γωνία φ με

$$\sin \varphi := r_{k+1, k+1} \quad (4.2.14)$$

τότε

$$\|r_{k+1}\|_2 = \cos \varphi,$$

αφού $\cos \varphi = \sqrt{1 - \sin^2 \varphi} = \sqrt{1 - r_{k+1, k+1}^2} = \|r_{k+1}\|_2$ και

$$\|y_{k+1}\|_2 \geq \|R_{11}\|_2^{-1} \|r_{k+1}\|_2 \geq \frac{1}{\sqrt{k}} \cos \varphi,$$

αφού οι στήλες του R είναι κανονικοποιημένες. Τότε υπάρχει ένας δείκτης j έτσι ώστε

$$|y_{j, k+1}| \geq \frac{1}{\sqrt{k}} \cos \varphi.$$

Μια αναδιάταξη των ιδιοτιμών στη Schur τριγωνική μορφή του (4.2.4) μπορεί να εκφραστεί με τον ακόλουθο τρόπο. Ο διαγώνιος πίνακας Λ αντικαθίσταται από $\Lambda' = P\Lambda P^T$, όπου P είναι ένας πίνακας μετάθεσης και

$$T = RP^T \Lambda' PR^{-1}.$$

Τώρα “τριγωνοποιούμε” ξανά την (4.2.8) και έχουμε

$$RP^T = Q'R' \quad (4.2.15)$$

και παίρνουμε τη νέα Schur μορφή

$$T' := Q'^H T Q' = R' \Lambda' R'^{-1}.$$

Εάν το P περιλαμβάνει μόνο τις πρώτες k ιδιοτιμές, τότε, διαμερίζοντας τη (4.2.15), βλέπουμε ότι

$$\begin{aligned} Y'_{12} &= R_{11}'^{-1} R'_{12} = (Q_{11}'^H R_{11} P_{11}'^T)^{-1} Q_{11}'^H R_{12} \\ &= P_{11} R_{11}^{-1} Q_{11}' Q_{11}'^H R_{12} = P_{11} Y_{12}. \end{aligned}$$

Επομένως μπορούμε να επιλέξουμε P ως I_{jk} και ο στοιχειώδης πίνακας μετάθεσης “κατασκευάζεται”

$$|y'_{k\ k+1}| = |y_{j\ k+1}| \geq \frac{1}{k} \cos \varphi.$$

Τώρα ισχύει $HP^T = QQ'R'$ και έτσι η (4.2.10) εφαρμόζοντας τη νέα παραγοντοποίηση δίνει

$$r'_{kk} = \sin \theta (x_j, [x_1, \dots, x_{j-1}, x_{j+1}, \dots, x_k]).$$

Η επιλογή (4.2.7) τότε θα συνεπάγεται $r'_{kk} \geq r_{kk}$. Οι τελευταίες $n-k$ γραμμές του R δεν επηρεάζονται από τη μετάθεση. Έτσι προκύπτει $r'_{k+1\ k+1} = r_{k+1\ k+1}$ και $\lambda'_{k+1} = \lambda_{k+1}$. Παρατηρούμε ότι $\lambda'_k = \lambda_j$, ώστε ο τύπος (4.2.13) που εφαρμόζεται στη νέα παραγοντοποίηση δίνει την εκτίμηση

$$\begin{aligned} |t'_{k\ k+1}| &= |\lambda'_k - \lambda'_{k+1}| |y'_{k\ k+1}| r'_{kk} / r'_{k+1\ k+1} \\ &\geq |\lambda_j - \lambda_{k+1}| \frac{\cos \varphi}{k} \frac{r_{kk}}{r_{k+1\ k+1}} \geq \frac{|\lambda_j - \lambda_{k+1}|}{k} \cotan \theta, \end{aligned} \quad (4.2.16)$$

αφού $\varphi \leq \theta$ από την (4.2.12) και την (4.2.14).

Προσθέτοντας έναν πίνακα E , που είναι μηδέν εκτός από ένα στοιχείο e στην $(k+1, k)$ θέση του T' , το χαρακτηριστικό πολυώνυμο λαμβάνει ένα τετραγωνικό παράγοντα

$$\det \begin{bmatrix} \lambda_j - \lambda & t'_{k\ k+1} \\ e & \lambda_{k+1} - \lambda \end{bmatrix}$$

το οποίο έχει διακρίνουσα

$$\Delta = (\lambda_j - \lambda_{k+1})^2 + 4 e t'_{k\ k+1}.$$

Εάν ο $T' + E$ έχει μια διπλή ιδιοτιμή, η διακρίνουσα είναι μηδέν δίνοντας

$$e = - \frac{(\lambda_j - \lambda_{k+1})^2}{4 t'_{k\ k+1}}.$$

Από την (4.2.5) και την (4.2.16) προκύπτει

$$|e| \leq \frac{n}{4} d \tan \theta,$$

η οποία αποδεικνύει την (4.2.6). □

ΒΙΒΛΙΟΓΡΑΦΙΑ

- 1) F.L. Bauer and C.T. Fike, Norms and exclusion theorems, *Numeriche Mathematik* 2, 137 - 141, 1960.
- 2) R.A. Smith, The Condition Numbers of the Matrix Eigenvalue Problem, *Numeriche Mathematik* 10, 232 - 240, 1967.
- 3) J.V. Burke, A.S. Lewis ,M.L. Overton, Spectral conditioning and pseudospectral growth, *Numeriche Mathematik*, 2007.
- 4) Pseudospectral gateway examples: <http://www.cs.ox.ac.uk/pseudospectra/index>.
- 5) P. Henrici, Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices, *Numeriche Mathematik* 4, 24 - 40, 1962.
- 6) J.H. Wilkinson, Note on Matrices with a Very Ill-Conditioned Eigenproblem, *Numeriche Mathematik* 19, 176 - 178, 1972.
- 7) A. Ruhe, Properties of a Matrix with a Very Ill-conditioned Eigenproblem, *Numeriche Mathematik* 15, 57 - 60, 1970.
- 8) J.H. Wilkinson, *The Algebraic Eigenvalue Problem*, Oxford: Oxford University Press, 1965
- 9) A.J. Hoffman, H.W. Wielandt, The variation of the spectrum of a normal matrix, *Duke Math. J.* 20, 37 - 39, 1953
- 10) Ole Hansen Hald, A Converse to the Bauer - Fike Theorem, *Linear Algebra and its Applications* 9, 267 - 274, 1974
- 11) Jianguo Huang and Jieyong Zhou, A direct proof and a generalization for a Kantorovich type inequality, *Linear Algebra and its Applications* 397, 185 - 192, 2005
- 12) Ralucca Gera, Matrix Norms and Conditions Numbers, *Numerical Linear Algebra*, Ch 7, <http://faculty.nps.edu/rgera/MA3042/2009/ch7.4.pdf>
- 13) Charles Van Loan, On Estimating the Condition of Eigenvalues and Eigenvectors, *Linear Algebra and its Applications*, 715 - 732, 1987
- 14) Δείκτης Κατάστασης, <http://users.auth.gr/natreas/Num-Math/Num-Math-3.pdf>
- 15) Ε. Γαλλόπουλος, Επιστημονικός Υπολογισμός ΙΙ, *Διάλεξη*, 2013, <http://scgroup.hpclab.ceid.upatras.gr/class/SCII/Lectures/set11.pdf>
- 16) Α.Γ. Φελλούρης, *Γραμμική Άλγεβρα και Αναλυτική Γεωμετρία*, 2006.
- 17) Ι. Β. Μαρουλάς, *Γραμμική Άλγεβρα*, ΕΜΠ, 2005
- 18) Ι. Β. Μαρουλάς, *Σημειώσεις Ανάλυσης Πινάκων*, Αθήνα, 2005
- 19) Π. Ψαρράκος, *Θέματα Ανάλυσης Πινάκων*, 2014.