

Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών



**Η ΚΑΤΑ HENRICI ΑΠΟΜΑΚΡΥΝΣΗ ΕΝΟΣ ΤΥΧΑΙΟΥ
ΤΕΤΡΑΓΩΝΙΚΟΥ ΠΙΝΑΚΑ ΑΠΟ ΤΗΝ ΚΑΝΟΝΙΚΟΤΗΤΑ**

Διπλωματική Εργασία

ΚΟΥΤΣΗ ΦΕΛΗΣΙΑ

Επιβλέπων: Παναγιώτης Ψαρράκος

Καθηγητής Ε.Μ.Π.

Αθήνα, Οκτώβριος 2018

Εθνικό Μετσόβιο Πολυτεχνείο
Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών



**Η ΚΑΤΑ HENRICI ΑΠΟΜΑΚΡΥΝΣΗ ΕΝΟΣ ΤΥΧΑΙΟΥ
ΤΕΤΡΑΓΩΝΙΚΟΥ ΠΙΝΑΚΑ ΑΠΟ ΤΗΝ ΚΑΝΟΝΙΚΟΤΗΤΑ**

Διπλωματική Εργασία

ΚΟΥΤΣΗ ΦΕΛΗΣΙΑ

Επιβλέπων: Παναγιώτης Ψαρράκος

Καθηγητής Ε.Μ.Π.

Εγκρίθηκε από την τριμελή επιτροπή την 3^η Οκτωβρίου 2018.

Β. Κανελλόπουλος

Ν. Γιαννακάκης

Π. Ψαρράκος

Αναπληρωτής Καθηγητής

Αναπληρωτής Καθηγητής

Καθηγητής

Αθήνα, Οκτώβριος 2018

Περίληψη

Στη παρούσα εργασία επικεντρωνόμαστε στην απομάκρυνση ενός τυχαίου τετραγωνικού πίνακα από την κανονικότητα κατά Henrici, η οποία είναι άμεσα υπολογίσιμη μόνο για τη νόρμα Frobenius, και θα δείξουμε πώς μπορεί να χρησιμοποιηθεί στην εκτίμηση άλλων ποσοτήτων.

Η απομάκρυνση κατά Henrici συνδέεται με την έννοια του πλησιέστερου στον A κανονικού πίνακα και μετρά την απόσταση του A από τους κανονικούς πίνακες που έχουν τις ίδιες ακριβώς ιδιοτιμές με τον A , λαμβάνοντας υπόψη και τις πολλαπλότητές τους.

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 ένα φράγμα για την απομάκρυνση από την κανονικότητα, αναφερόμαστε 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. Επίσης, περιγράφουμε το αριθμητικό πεδίο ενός πίνακα και δίνουμε κάποια χρήσιμα θεωρήματα για την απόσταση ενός σημείου του αριθμητικού πεδίου ενός μη κανονικού πίνακα από την κυρτή θήκη των ιδιοτιμών του.

Η απομάκρυνση από την κανονικότητα ενός πίνακα είναι ένα πραγματικό βαθμωτό μέγεθος, το οποίο δεν είναι πρακτικό να υπολογιστεί αν ο πίνακας έχει μεγάλες διαστάσεις και οι ιδιοτιμές του είναι άγνωστες. Θα παρουσιάσουμε έναν απλό τύπο για τον υπολογισμό ενός καινούργιου άνω φράγματος για την απομάκρυνση από την κανονικότητα. Αυτό το καινούργιο άνω φράγμα φθηνότερο να υπολογιστεί από το άνω φράγμα κατά Henrici. Τέλος, δίνουμε τα καλύτερα δυνατά φράγματα για την απομάκρυνση από την κανονικότητα. Αυτά τα νέα άνω και κάτω φράγματα είναι βέλτιστα για ερμιτιανούς πίνακες, αντιερμιτιανούς πίνακες, και γενικά για κάθε πίνακα με ιδιοτιμές οριζόντια ή κάθετα ευθυγραμμισμένες στο μιγαδικό επίπεδο.

ΛΕΞΕΙΣ ΚΛΕΙΔΙΑ: μη-κανονικός πίνακας, απομάκρυνση από την κανονικότητα, αριθμητικό πεδίο, ιδιοτιμές, φασματική μεταβολή, μεταβολή ιδιοτιμής.

Abstract

In this thesis, we concentrate our study in the departure from normality of an arbitrary square matrix A of order n with complex elements as defined by Henrici, which is directly computable only for the Frobenius norm, and will show its significance in estimating other quantities.

The departure from normality of matrix A is associated with the concept of the closest normal matrix to A and measures the distance of A from all the normal matrices which have exactly the same eigenvalues as A , considering their multiplicities.

Firstly, we mention the Schur Theorem, which is a very helpful tool in the triangulation of matrices and which we will use in some proofs of theorems. Then, we give a bound derived by Henrici for the departure from normality, describe iterated matrices, their norms and some approximate solutions of linear systems. Later, we refer to spectral variation and eigenvalue variation of matrices, which are two concepts that combine the eigenvalues of two different matrices. Also, we describe the field of values of a matrix and give some theorems about the distance of the boundary of the field of values of a non-normal matrix from the convex hull of its eigenvalues.

The departure from normality of a matrix is a real scalar that is impractical to compute if a matrix is large and its eigenvalues are unknown. A simple formula is presented for computing an upper bound for departure from normality in the Frobenius norm. This new upper bound is cheaper to compute than the upper bound derived by Henrici. Finally, the best available bounds for departure from normality are given. Moreover, these new lower and upper bounds are sharp for Hermitian matrices, skew-Hermitian matrices and, in general, any matrix with eigenvalues that are horizontally or vertically aligned in the complex plane.

Περιεχόμενα

ΠΕΡΙΛΗΨΗ.....	4
ABSTRACT.....	5
ΣΥΜΒΟΛΙΣΜΟΙ.....	8
ΚΕΦΑΛΑΙΟ 1	12
Εισαγωγικές Έννοιες και Αποτελέσματα της Γραμμικής Άλγεβρας.....	12
1.1. Γενικοί Ορισμοί.....	12
1.2 Στοιχεία Ανάλυσης Πινάκων	14
ΚΕΦΑΛΑΙΟ 2	18
Απομάκρυνση από την κανονικότητα.....	18
2.1. Ένα μέτρο για τη μη κανονικότητα.....	19
2.1.1. Η απομάκρυνση από την κανονικότητα ως προς τη $\ \cdot\ $	20
2.1.2. Ένα φράγμα για το $\delta(A)$	22
2.2. Δυνάμεις πινάκων	26
2.2.1. Ένα νέο φράγμα	27
2.3. Προσεγγιστικές λύσεις για γραμμικά συστήματα.....	28
2.3.1 Εισαγωγή.....	28
2.3.2. Ένα νέο φράγμα για τη νόρμα του αντιστρόφου.....	27
2.4. Φασματική μεταβολή και μεταβολή ιδιοτιμών	28
2.4.1. Κλασικά αποτελέσματα.....	28
2.4.2. Ένα νέο αποτέλεσμα για το $sA(B)$	29
2.4.3. Αριθμητικό παράδειγμα	35
2.4.4. Σχετικά αποτελέσματα για το $\nu(A, B)$	37
2.4.5. Ένα αποτέλεσμα σε προσεγγιστικές ιδιοτιμές.....	37

2.5. Αριθμητικό πεδίο	38
2.5.1. Κλασικά αποτελέσματα.....	39
2.5.2. Η απόσταση του συνόρου του αριθμητικού πεδίου από τη κυρτή θήκη των ιδιοτιμών του πίνακα A	40
2.5.3. Ένα βελτιωμένο φράγμα για τα αριθμητικά πεδία.....	43
 ΚΕΦΑΛΑΙΟ 3	 46
Φράγματα για την απομάκρυνση από την κανονικότητα.....	46
3.1. Ένα βέλτιστο άνω φράγμα.....	48
3.2. Φράγματα για την απομάκρυνση από την κανονικότητα και η Frobenius νόρμα των ιδιοτιμών ενός πίνακα	52
3.3. Φράγματα για τις ιδιοτιμές και την απομάκρυνση από την κανονικότητα	54
3.3.1. Βελτιωμένα κάτω φράγματα	56
3.3.2. Βελτιωμένα άνω φράγματα.....	58
3.4. Βασικά αποτελέσματα.....	59
 ΚΕΦΑΛΑΙΟ 4	 63
ΠΑΡΑΔΕΙΓΜΑΤΑ.....	63
 ΒΙΒΛΙΟΓΡΑΦΙΑ	 69

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
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^*	Αναστροφοσυζυγής του πίνακα A
x^T	Ανάστροφος του διανύσματος x
x^*	Αναστροφοσυζυγής του διανύσματος x
$H(A)$	Το ερμιτιανό μέρος ενός τετραγωνικού πίνακα A
$S(A)$	Το αντιερμιτιανό μέρος ενός τετραγωνικού πίνακα A
A^r	Δύναμη πίνακα
$\text{trace}(A)$ ή $\text{tr}(A)$	Ίχνος ενός τετραγωνικού πίνακα A
$\det(A)$	Ορίζουσα ενός τετραγωνικού 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
\tilde{A}	Κεντραρισμένος πίνακας
$\text{Co}(\Omega)$	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 Ευκλείδεια νόρμα
$\ \cdot\ _\infty$	Η ∞ -νόρμα διανυσμάτων ή μέγιστη νόρμα
$\ \cdot\ $	Νόρμα πινάκων
$\ \cdot\ _F$	Η νόρμα Frobenius
$\ \cdot\ _1$	Νόρμα μεγίστου αθροίσματος στηλών πίνακα
$\ \cdot\ _2$	Η φασματική (τελεστική) νόρμα πινάκων
$\ \cdot\ _\infty$	Νόρμα μεγίστου αθροίσματος γραμμών πίνακα
$\kappa(A)$ ή $\kappa_{\ \cdot\ }(A)$ ή $\kappa_\alpha(A)$	Δείκτης κατάστασης του πίνακα A τυχαίας νόρμας
$\kappa_2(A)$	Δείκτης κατάστασης του πίνακα A Ευκλείδειας νόρμας

Κεφάλαιο 1

Εισαγωγικές Έννοιες και Αποτελέσματα της Γραμμικής Άλγεβρας

Σκοπός αυτού του κεφαλαίου είναι να ορίσουμε μερικές βασικές έννοιες που αφορούν την Γραμμική Άλγεβρα και την Ανάλυση Πινάκων και θα μας χρειαστούν παρακάτω.

Σε όλα τα παρακάτω θα συμβολίζουμε το σύνολο των μιγαδικών $m \times n$ πινάκων με $\mathbb{C}^{m \times n}$ και το σύνολο των πραγματικών $m \times n$ πινάκων θα συμβολίζεται με $\mathbb{R}^{m \times n}$.

1.1. Γενικοί Ορισμοί

Ορισμός 1.1.1. 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 1.1.2. 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 ικανοποιεί τη σχέση $AI_n = I_n A = A$, για κάθε $n \times n$ πίνακα A .

Ορισμός 1.1.3. 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 1.14. 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 1.1.5. *Τχνος* ενός πίνακα $A=(a_{ij})$, που συμβολίζουμε με $\text{trace}(A)$ ή απλώς $\text{tr}(A)$, ονομάζουμε το άθροισμα των διαγώνιων στοιχείων του πίνακα,δηλαδή $\text{tr}(A)=\sum_i a_{ii}$.

Ορισμός 1.1.6. 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 *αντίστροφός* του. Γράφουμε τότε $X=A^{-1}$.

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

Ορισμός 1.1.7. 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 1.1.8 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 ονομάζεται *ανάστροφος* πίνακας του A και συμβολίζεται με A^T . Έχουμε δηλαδή ότι $A^T = (a_{ji}) \in \mathbb{C}^{m \times n}$.

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

Ορισμός 1.1.9. Αν $A=(a_{ij}) \in \mathbb{C}^{m \times n}$, τότε ο πίνακας με τα συζυγή στοιχεία του ονομάζεται *συζυγής* πίνακας και συμβολίζεται με $\bar{A} = [\bar{a}_{ij}]$. Ο *ανάστροφος συζυγής* του πίνακα A συμβολίζεται με $A^* = \bar{A}^T$. Εάν ισχύει ότι $A=A^*$, τότε ο πίνακας ονομάζεται *ερμιτιανός*. Ακόμη, εάν $A^* = -A$, τότε ο πίνακας ονομάζεται *αντιερμιτιανός*.

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

Ορισμός 1.1.10. Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ λέγεται *ορθομοναδιαίος* αν ισχύει

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

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

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

Ορισμός 1.1.11. Ένας τετραγωνικός πίνακας A καλείται *κανονικός* όταν αντιμετωπίζεται με τον αναστροφοσυζυγή του, δηλαδή όταν ικανοποιεί τη σχέση

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

Ορισμός 1.1.12. 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.

Υπενθυμίζουμε εδώ ότι κάθε πίνακας μπορεί να γραφτεί ως άθροισμα ενός συμμετρικού και ενός αντισυμμετρικού πίνακα.

Ορισμός 1.1.13. Έστω A ένας $n \times n$ πίνακας. Ονομάζουμε *υποπίνακα* του A , τον πίνακα που προκύπτει από τον A , αν διαγράψουμε ένα αριθμό γραμμών ή έναν αριθμό στηλών ή έναν αριθμό γραμμών και στηλών.

Ορισμός 1.1.14. Έστω A ένας $n \times n$ πίνακας. Ένα μη-μηδενικό διάνυσμα $x \in \mathbb{C}^n$ ονομάζεται *ιδιοδιάνυσμα* του A αν

$$Ax = \lambda x$$

για κάποιο $\lambda \in \mathbb{C}$. Αυτό το λ ονομάζεται *ιδιοτιμή* του A και λέμε ότι το x είναι ένα *ιδιοδιάνυσμα* του A που αντιστοιχεί στη ιδιοτιμή 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 1.1.15. Μια ιδιοτιμή ενός πίνακα A είναι *απλή* (simple), αν αυτή έχει αλγεβρική πολλαπλότητα 1 και *πολλαπλή* (multiple) διαφορετικά.

Ορισμός 1.1.16. Ο πίνακας $A \in \mathbb{C}^{n \times n}$ είναι *διαγωνοποιήσιμος* αν υπάρχει αντιστρέψιμος πίνακας $P \in \mathbb{C}^{n \times n}$ τέτοιος ώστε ο πίνακας $P^{-1}AP$ να είναι ο διαγώνιος πίνακας, δηλαδή αν ο πίνακας A είναι όμοιος με διαγώνιο πίνακα.

1.2. Στοιχεία Ανάλυσης Πινάκων

Ορισμός 1.2.1. Μια συνάρτηση $\|\cdot\|: \mathbb{C}^n \rightarrow \mathbb{R}$ ονομάζεται *νόρμα διανυσμάτων*, αν για κάθε $x, y \in \mathbb{C}^n$ ικανοποιεί τα ακόλουθα:

- 1) $\|x\| \geq 0$ (μη αρνητική).
- 2) $\|x\| = 0$ αν και μόνο αν $x = 0$.
- 3) $\|ax\| = |a|\|x\|$ για κάθε $a \in \mathbb{C}$.
- 4) $\|x + y\| \leq \|x\| + \|y\|$ (τριγωνική ανισότητα).

Μία συνάρτηση $\|\cdot\|: \mathbb{C}^n \rightarrow \mathbb{R}$ που ικανοποιεί τα (1),(3) και (4) του παραπάνω ορισμού καλείται ημι-νόρμα διανυσμάτων. Η ημι-νόρμα αποτελεί μία γενίκευση της έννοιας της νόρμας, η οποία επιτρέπει σε μη μηδενικά διανύσματα να έχουν μηδενικό μέτρο.

Ορισμός 1.2.2. Μία 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:

- 1) $\|A\| \geq 0$ (μη αρνητική) και $\|A\| = 0$ αν και μόνο αν $A = 0$.
- 2) $\|\alpha A\| = |\alpha| \|A\|$ για κάθε $\alpha \in \mathbb{C}$.
- 3) $\|A+B\| \leq \|A\| + \|B\|$ (τριγωνική ανισότητα).
- 4) $\|AB\| \leq \|A\| \|B\|$ (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 νόρμες διανυσμάτων αποτελούν νόρμες πινάκων στον διανυσματικό χώρο $\mathbb{C}^{n \times n}$. Κάποια παραδείγματα είναι τα εξής:

- 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 A^* είναι ο αναστροφосуζυγής 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, υπολογίζουμε:

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

Με απλές πράξεις μπορούμε να επαληθεύσουμε ότι οι δύο πρώτες νόρμες αποτελούν πράγματι νόρμες πινάκων, ενώ η τελευταία όχι, αφού δεν ικανοποιεί τη σχέση (4) του παραπάνω ορισμού.

Ορισμός 1.2.3. Έστω $\|\cdot\|$ μία νόρμα διανυσμάτων στο \mathbb{C}^n . Η επαγόμενη από την $\|\cdot\|$ νόρμα στο $\mathbb{C}^{n \times n}$ ορίζεται ως

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

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

Στη συνέχεια θα αναφέρουμε τα σημαντικότερα παραδείγματα νορμών πινάκων που επάγονται από τις γνωστές l_p -νόρμες. Σε όλα τα παρακάτω θεωρώ έναν τετραγωνικό πίνακα $A = (a_{ij}) \in \mathbb{C}^{n \times n}$.

- Η νόρμα πινάκων μεγίστου αθροίσματος κατά στήλη στο ορίζεται ως

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

Η νόρμα αυτή 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

$$\begin{aligned} \|Ax\|_1 &= \|x_1 a_1 + x_2 a_2 + \dots + x_n a_n\|_1 \\ &\leq |x_1| \|a_1\|_1 + |x_2| \|a_2\|_1 + \dots + |x_n| \|a_n\|_1 \end{aligned}$$

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

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, $\max_{\|x\|_1} \|Ax\|_1 = \|A\|_1$. Συνεπώς, η νόρμα $\|\cdot\|_1$ επάγεται από τη διανυσματική l_1 -νόρμα και αποτελεί νόρμα πινάκων.

- Η νόρμα πινάκων μεγίστου αθροίσματος κατά γραμμή $\mathbb{C}^{n \times n}$ στο 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 της τελευταίας σχέσης είναι παρόμοια με την απόδειξη του προηγούμενου παραδείγματος. Προφανώς, και η νόρμα $\|\cdot\|_\infty$ ως επαγόμενη αποτελεί νόρμα πινάκων.

- Η φασματική (τελεστική) νόρμα πινάκων στο ορίζεται ως

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

Επίσης, ένας ισοδύναμος ορισμός είναι και ο παρακάτω:

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

όπου με $\sigma(A)$ συμβολίζουμε το φάσμα του πίνακα A , δηλαδή το σύνολο των ιδιοτιμών του A και με A^* συμβολίζουμε τον αναστροφοσυζυγή του A . Η νόρμα αυτή είναι μια νόρμα πινάκων που συνεπάγεται από τη διανυσματική νόρμα l_2 , δηλαδή

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

Σημείωση: Μερικές νόρμες πινάκων, όπως η νόρμα Frobenius και η L_1 -νόρμα, δεν επάγονται.

Ορισμός 1.2.4. Έστω ένας πίνακας A 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.

Ορισμός 1.2.5. Μια νόρμα 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 ότι $\|Ux\| = \|x\|$.

Ορισμός 1.2.6. 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

$$\kappa(A) = \kappa_\alpha(A) = \begin{cases} \|A\|_\alpha \|A^{-1}\|_\alpha, & \text{αν } A \text{ αντιστρέψιμος πίνακας} \\ 0, & \text{διαφορετικά} \end{cases}$$

ονομάζεται δείκτης κατάστασης (condition number) του πίνακα A .

Για παράδειγμα, για τη 2-νόρμα ο δείκτης κατάστασης είναι

$$\kappa_2(A) = \|A\|_2 \|A^{-1}\|_2 = \frac{\sigma_1}{\sigma_2}.$$

Εάν ο A δεν είναι αντιστρέψιμος, τότε ορίζουμε $\kappa_2(A) = \infty$.

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 (illconditioned). 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.

1.2.1.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.

Ορισμός 1.2.7. Ένας τετραγωνικός πίνακας $A \in \mathbb{C}^{n \times n}$ καλείται *κανονικός* όταν αντιμετωπίζεται με τον αναστροφосуζυγή του, δηλαδή όταν ικανοποιεί τη σχέση $AA^* = A^*A$.

Κάποιες από τις συνθήκες κανονικότητας που θα μας χρειαστούν:

Πρόταση 1.2.8. Ένας αντιστρέψιμος πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν ο A^{-1} είναι κανονικός.

Πρόταση 1.2.9. Ένας αντιστρέψιμος πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν ο $A^{-1}A^*$ είναι ορθομοναδιαίος.

Πρόταση 1.2.10. Ένας αντιστρέψιμος πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν ο A αντιμετωπίζεται με τον πίνακα $A^{-1}A^*$.

Πρόταση 1.2.11. 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 1.2.12. 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 $U^*AU = \Lambda$, δηλαδή όταν είναι ορθομοναδιαία όμοιος με διαγώνιο πίνακα.

Πρόταση 1.2.13. Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν ο A είναι διαγωνοποιήσιμος και υπάρχει ορθοκανονική βάση του $\mathbb{C}^{n \times n}$ που αποτελείται από τα ιδιοδιανύσματα του A .

Πρόταση 1.2.14. 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 1.2.15. 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 1.2.16. Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι κανονικός αν και μόνο αν $\|A\|_F = \|A^*\|_F = \sqrt{\sum |\lambda_k|^2}$.

Κεφάλαιο 2

Απομάκρυνση από την κανονικότητα

Στο παρόν κεφάλαιο θα εξετάσουμε τα ακόλουθα προβλήματα της γραμμικής άλγεβρας που αφορούν:

- 1) Την εκτίμηση των νορμών πινάκων A^n , $n = 1, 2, 3, \dots$, σε σχέση με τη $\rho(A)$.
- 2) 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 $\rho(A^{-1})$.
- 3) Την εκτίμηση της απόστασης του φάσματος ενός πίνακα B από το φάσμα ενός πίνακα A σε σχέση με τη νόρμα του $B - A$.
- 4) 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

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

Οι λύσεις στα παραπάνω προβλήματα είναι κλασσικές (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 $A = S J S^{-1}$, 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 κλασσικές 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.

2.1. Ένα μέτρο για την μη κανονικότητα

Ορισμός 2.1.1. 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 την ορίζει.

Ορισμός 2.1.2. Μια νόρμα πίνακα $\|\cdot\|$ θα καλείται *ορθομοναδιαία αναλλοίωτη* (unitarily invariant), εάν $\|UAU^*\| = \|A\|$ για κάθε A και για κάθε ορθομοναδιαίο U . Η νόρμα Frobenius και η φασματική νόρμα είναι ορθομοναδιαία αναλλοίωτες.

Ορισμός 2.1.3. Μια επαγόμενη νόρμα $\|\cdot\|$ καλείται *προσανατολισμένη κατά άξονα* (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 $\|\cdot\|_\infty$, και είναι προσανατολισμένες κατά άξονα.

Ορισμός 2.1.4. Μια νόρμα $\|\cdot\|_\alpha$ λέμε ότι είναι μεγαλύτερη μιας άλλης νόρμας $\|\cdot\|_\beta$ εάν $\|A\|_\alpha \geq \|A\|_\beta$ για όλους τους πίνακες A . Ξεκάθαρα, η $\|A\|_{l_1}$ 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.

Θεώρημα 2.1.5. Έστω $\|\cdot\|$ μία νόρμα πινάκων στο $\mathbb{C}^{n \times n}$. Τότε για κάθε $A \in \mathbb{C}^{n \times n}$, ισχύει

$$\rho(A) \leq \|A\| \quad (2.1.1)$$

Η φασματική ακτίνα ενός πίνακα $A \in \mathbb{C}^{n \times n}$ ορίζεται ως $\rho(A) = \max \{ |\lambda| : \lambda \in \sigma(A) \}$, όπου $\sigma(A)$ συμβολίζουμε το φάσμα του A , δηλαδή το σύνολο των ιδιοτιμών 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 $x_0 \in \mathbb{C}^n$ 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,

$$\rho(A) = \lambda_0 \leq \|A\|.$$

□

Πρόταση 2.1.6. Έστω $\|\cdot\|_a$ και $\|\cdot\|_b$ δύο οποιοσδήποτε νόρμες πίνακα. Τότε υπάρχει μια σταθερά M_{ab} που εξαρτάται μόνο από αυτές τις δύο νόρμες, τέτοια ώστε

$$\|A\|_a \leq M_{ab} \|A\|_b \quad (2.1.2)$$

για όλους τους πίνακες A .

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

Λήμμα 2.1.7 (Λήμμα 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

$$A = UTU^*.$$

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 $M = (m_{ij})$ είναι γνήσια άνω τριγωνικός και έτσι τότε

$$M^r = 0, r \geq n. \quad (2.1.3)$$

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

$$\delta_{\|\cdot\|}(A) = \inf\{\|M\|: U^*AU = D + M, U^*U = UU^* = I_n\}, \quad (2.1.4)$$

όπου το inf λαμβάνεται σε σχέση με όλους του M που μπορούν να εμφανιστούν στη Schur τριγωνική μορφή.

Παρατήρηση: Χρησιμοποιούμε τον όρο “απομάκρυνση” αντί του όρου “απόσταση”, γιατί συνδέεται με την έννοια του πλησιέστερου στον A κανονικού πίνακα και μετρά την απόσταση του A από τους κανονικούς πίνακες που έχουν τις ίδιες ακριβώς ιδιοτιμές με τον A, λαμβάνοντας υπόψη και τις πολλαπλότητες.

Αφού η νόρμα Frobenius είναι ορθομοναδιαία αναλλοίωτη, θα ισχύει ότι $\|A\|_F^2 = \|UTU^*\|_F^2 = \|T\|_F^2 = \|D + M\|_F^2 = \|D\|_F^2 + \|M\|_F^2$.

Άρα προκύπτει ότι η σχέση

$$\|M\|_F = \{\|A\|_F^2 - \sum_{i=1}^n |\lambda_i|^2\}^{\frac{1}{2}} = \{\sum_{i,j=1}^n |a_{ij}|^2 - \sum_{i=1}^n |\lambda_i|^2\}^{\frac{1}{2}} \quad (2.1.5)$$

όπου $\lambda_i, i=1,2,\dots,n$ οι ιδιοτιμές του A (όχι κατ' ανάγκη διακεκριμένες), είναι ανεξάρτητη από την ειδική επιλογή της Schur τριγωνικής μορφής. Χρησιμοποιώντας τις παραπάνω σχέσεις παρατηρούμε ότι $\delta_{\|\cdot\|_F}(A) = \|M\|_F$.

Πρόταση 2.1.8. 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 $A^*A = AA^*$. Όπως είδαμε παραπάνω, με βάση το Θεώρημα Schur ο A γράφεται στη μορφή $A = UTU^*$, όπου ο U είναι ορθομοναδιαίος πίνακας (δηλαδή, $U^*U = UU^* = I_n$). Έτσι προκύπτει ότι

$$\begin{aligned} A^*A = AA^* &\Leftrightarrow (UTU^*)^*(UTU^*) = (UTU^*)(UTU^*)^* \Leftrightarrow \\ UT^*U^*UTU^* &= UTU^*UT^*U^* \Leftrightarrow UT^*I_nTU^* = UTI_nT^*U^* \Leftrightarrow \\ UT^*TU^* &= UTT^*U^* \Leftrightarrow U^{-1}UT^*TU^*(U^*)^{-1} = U^{-1}UTT^*U^*(U^*)^{-1} \Leftrightarrow \\ I_nT^*TI_n &= I_nTT^*I_n \Leftrightarrow T^*T = TT^*. \end{aligned}$$

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 με στοιχεία τις ιδιοτιμές του πίνακα A και ο M είναι αυστηρά άνω τριγωνικός πίνακας.

Η μόνη, όμως, περίπτωση στην οποία ισχύει $T^*T = TT^*$ είναι όταν τα στοιχεία στα δεξιά της κύριας διαγωνίου είναι μηδέν, αφού και τα στοιχεία στα αριστερά της κύριας διαγωνίου θα 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.

□

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.

2.1.2. Ένα φράγμα για το $\delta_{\|\cdot\|_F}(A)$

Σε αυτή την παράγραφο θα βρούμε 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 (2.1.2) παραπάνω μπορούμε να κατασκευάσουμε φράγματα για το $\delta_{\|\cdot\|_F}(A)$ για κάθε νόρμα $\|\cdot\|$.

Θεώρημα 2.1.9. Για έναν αυθαίρετο πίνακα A έχουμε

$$\|M\|_F = \delta_{\|\cdot\|_F}(A) \leq \sqrt{\frac{n^3 - n}{12}} \sqrt{\|A^*A - AA^*\|_F}. \quad (2.1.6)$$

Η ισότητα ισχύει αν και μόνο αν ο A είναι ορθομοναδιαίομοιος με έναν πίνακα της μορφής

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

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

Παρατήρηση: Για το παραπάνω φράγμα διαπιστώνουμε ότι δεν είναι συναρτήσεων 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

$$\|M\|_F^2 \leq \gamma_{22} + 2\gamma_{33} + \dots + (n-1)\gamma_{nn}, \quad (2.1.7)$$

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

$$\gamma_{ii} = \sum_{k < i} |m_{ki}|^2 - \sum_{k > i} |m_{ki}|^2, i = 1, 2, \dots, n. \quad (2.1.8)$$

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

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

Άρα, χρησιμοποιώντας αυτή την σχέση (με το n αυξανόμενο κατά 1) έχουμε

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

Σημειώνουμε ότι η ισότητα ισχύει στην (2.1.7) αν και μόνο αν $m_{ij} = 0$ για κάθε $j > i+1$ (δηλαδή, αν ο τριγωνικός πίνακας T είναι “σχεδόν διαγώνιος”). Αφαιρώντας από την (2.1.7) $\frac{n-1}{2}$ φορές την ταυτότητα $\gamma_{11} + \gamma_{22} + \dots + \gamma_{nn} = 0$ (η οποία είναι συνέπεια της σχέσης (2.1.8)), βρίσκουμε ότι

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

και χρησιμοποιώντας το θεώρημα Cauchy, έχουμε:

$$\|M\|_F^4 \leq \left[\left(\frac{1-n}{2}\right)^2 + \left(\frac{3-n}{2}\right)^2 + \dots + \left(\frac{n-1}{2}\right)^2 \right] [\gamma_{11}^2 + \gamma_{22}^2 + \dots + \gamma_{nn}^2]. \quad (2.1.9)$$

Σε αυτό το σημείο θα υπενθυμίσουμε το θεώρημα Cauchy:

Θεώρημα Cauchy: Αν $\alpha_1, \dots, \alpha_n$ και b_1, \dots, b_n είναι δύο n -άδες μιγαδικών αριθμών, τότε

$$(a_1b_1 + a_2b_2 + \dots + a_nb_n)^2 \leq (a_1^2 + a_2^2 + \dots + a_n^2)(b_1^2 + b_2^2 + \dots + b_n^2)$$

ή, ισοδύναμα,

$$|\alpha_1 b_1 + \alpha_2 b_2 + \dots + \alpha_n b_n| \leq \left(\sqrt{a_1^2 + a_2^2 + \dots + a_n^2} \sqrt{b_1^2 + b_2^2 + \dots + b_n^2} \right).$$

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 (2.1.6) λόγω του γεγονότος ότι

$$\begin{aligned} A^*A - AA^* &= (UTU^*)^*UTU^* - UTU^*(UTU^*)^* \\ &= UT^*U^*UTU^* - UTU^*UT^*U^*UT^*TU^* - UTT^*U^* \\ &= U(T^*T - TT^*)U^* \\ &= U\Gamma U^* \end{aligned}$$

και ως εκ τούτου έχουμε

$$\gamma_{11}^2 + \gamma_{22}^2 + \dots + \gamma_{nn}^2 \leq \|\Gamma\|_F^2 = \|U\Gamma U^*\|_F^2 = \|A^*A - AA^*\|_F^2. \quad (2.1.10)$$

Η ισότητα ισχύει στην (2.1.6) αν και μόνο αν η ισότητα ισχύει ταυτόχρονα στην (2.1.7), στην (2.1.9) και στην (2.1.10). Η ισότητα στην (2.1.10) απαιτεί ο πίνακας $\Gamma = T^*T - TT^*$ να είναι διαγώνιος πίνακας. Για τους σχεδόν διαγώνιους πίνακες T (όπως απαιτείται από την ισότητα στην (2.1.7)), ο Γ είναι διαγώνιος αν και μόνο αν όλα τα διαγώνια στοιχεία του T ταυτίζονται. Τα διαγώνια στοιχεία του Γ (με βάση και την (2.1.8)) ικανοποιούν τη σχέση

$$|m_{k,k+1}|^2 = \gamma_{11} + \gamma_{22} + \dots + \gamma_{kk}, \quad k=1,2,\dots,n-1. \quad (2.1.11)$$

Η ισότητα στην (2.1.9) τελικά προϋποθέτει ότι $\gamma_{kk} = \alpha(1+n-2k)$, $k=1,2,\dots,n$, όπου α είναι μια πραγματική σταθερά.

Η σχέση (2.1.11) ικανοποιεί την υπόθεση αν και μόνο αν, για ορισμένους πραγματικούς $\alpha \geq 0$, $|m_{k,k+1}|^2 = \alpha k(n-k)$, $k=1,2,\dots,n-1$, δηλαδή, εάν ο T έχει τη μορφή που αναφέρεται στο θεώρημα.

□

2.2. Δυνάμεις πινάκων

Ορισμός 2.2.1. Ονομάζουμε *δύναμη πίνακα* έναν πίνακα της μορφής

$$A^r, \quad r = 1, 2, \dots$$

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

$$\|A^r\|_2 = \rho^r(A), \quad r = 1, 2, \dots \quad (2.2.1)$$

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 $\|\cdot\|_\infty$. Το φράγμα του Gautschi ομοίως απαιτεί κάποια γνώση της Jordan κανονικής μορφής.

2.2.1. Ένα νέο φράγμα

Σε αντίθεση με τα παραπάνω, το Θεώρημα 2.2.2. παρακάτω μας δίνει μια εκτίμηση για το $\|A^r\|_2$, η οποία εξαρτάται μόνο από τη $\rho(A)$ και τη $\delta_{\|\cdot\|_2}(A)$ και τείνει στη (2.2.1) αν ο A είναι κανονικός πίνακας.

Θεώρημα 2.2.2. Εάν $m = \delta_{\|\cdot\|_2}(A)$ και εάν $\rho(A) > 0$, τότε

$$\|A^r\|_2 \leq \rho^r(A) + \binom{r}{1} \rho^{r-1}(A)m + \dots + \binom{r}{n-1} \rho^{r-n+1}(A)m^{n-1}, \quad (2.2.2)$$

εάν $\rho(A) = 0$, τότε

$$\begin{cases} \|A^r\|_2 \leq m^r, & r = 0, 1, \dots, n-1, \\ \|A^r\|_2 = 0, & r \geq n. \end{cases} \quad (2.2.3)$$

Παρατήρηση: Η έκφραση στη δεξιά πλευρά του (2.2.2) 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 περιέχουν qM και $r - qD$ για $q = 0, 1, 2, \dots, n - 1$. Χρησιμοποιώντας τη νόρμα $\|\cdot\|_2$ και τη σχέση $\|D\|_2 = \rho(A)$, αντικαθιστούμε στις σχέσεις (2.2.2) και (2.2.3), όπου m το $\|M\|_2$. Επειδή το αποτέλεσμα είναι αληθές για κάθε M και επειδή οι εκφράσεις στα δεξιά της (2.2.2) και της (2.2.3) εξαρτώνται συνεχώς από το m , διαπιστώνουμε ότι το συμπέρασμα του θεωρήματος προκύπτει αμέσως.

□

2.3. Προσεγγιστικές λύσεις για γραμμικά συστήματα

2.3.1 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 τον \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 μία προσέγγιση του αντιστρόφου 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 τη $\|A^{-1}\|$. 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 βρίσκουμε εύκολα από τη σχέση $A^{-1} = S D^{-1} S^{-1}$ ότι

$$\|A^{-1}\| = \kappa(S) \rho(A^{-1}), \quad (2.3.1)$$

όπου το

$$\kappa(S) = \|S\| \|S^{-1}\| \quad (2.3.2)$$

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 άρα, από την (2.1.1) προκύπτει ότι

$$\|A^{-1}\|_2 = \rho(A^{-1}) \quad (2.3.3)$$

για κανονικούς πίνακες A . Για μη-κανονικούς πίνακες το φράγμα (2.3.1) απαιτεί την πλήρη διαγωνοποίηση του πίνακα A .

2.3.2. Ένα νέο φράγμα για τη νόρμα του αντιστρόφου

Για μια πραγματική μεταβλητή $x \geq 0$, έστω η συνάρτηση f οριζόμενη ως εξής:

$$f(x) = x + x^2 + \dots + x^n. \quad (2.3.4)$$

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

$$\lim_{x \rightarrow 0^+} x^{-1} f(x) = 1. \quad (2.3.5)$$

Μπορούμε τώρα να αναφέρουμε το ακόλουθο Θεώρημα.

Θεώρημα 2.3.1. Εάν ο A είναι ένας αντιστρέψιμος και μη-κανονικός πίνακας και εάν $x = \rho(A^{-1})\delta_{\|\cdot\|_2}(A)$, τότε

$$\|A^{-1}\|_2 \leq \frac{f(x)}{x} \rho(A^{-1}) = \frac{f(x)}{\delta_{\|\cdot\|_2}(A)}. \quad (2.3.6)$$

Παρατήρηση: Για $\rho(A^{-1})$ σταθερό, καθώς $\delta_{\|\cdot\|_2}(A) \rightarrow 0$, το φράγμα (2.3.6) προσεγγίζει την (2.3.3), ενόψει της (2.3.5). Φράγματα για τη $\|A^{-1}\|$, όπου $\|\cdot\|$ είναι αυθαίρετη νόρμα μπορούν να βρεθούν από την (2.1.2).

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 (2.1.3), ισχύει $(D^{-1}M)^n = 0$. Άρα

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

Έτσι έχουμε το αποτέλεσμα (2.3.6) με τη λήψη νορμών και ενόψει της $\|D^{-1}\|_2 = \rho(A^{-1})$ και του γεγονότος ότι η $\|\cdot\|$ είναι ορθομοναδιαία αναλλοίωτη.

□

2.4. Φασματική μεταβολή και μεταβολή ιδιοτιμών

2.4.1. 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

$$s = s_A(B) = \max_{1 \leq i \leq n} \left\{ \min_{1 \leq j \leq n} |\mu_i - \lambda_j| \right\} \quad (2.4.1)$$

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

$$v = v(A,B) = \min_p \left\{ \max_{1 \leq i \leq n} |\lambda_i - \mu_{p(i)}| \right\}, \quad (2.4.2)$$

η οποία ονομάζεται *μεταβολή ιδιοτιμής* του A και του B. Εδώ το ελάχιστο λαμβάνεται σε σχέση με όλες τις μεταθέσεις p του συνόλου (1, 2, ..., n). Προφανώς, $v(A,B) = v(B, A)$ και

$$s_A(B) \leq v(A,B) \quad (2.4.3)$$

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 1.1., τότε

$$s_A(B) \leq (n+2) M \left[\frac{\|A-B\|_{l_1}}{M} \right]^{1/n} \quad (2.4.4)$$

και

$$v(A,B) \leq 2n(n+2) M \left[\frac{\|A-B\|_{l_1}}{M} \right]^{1/n}. \quad (2.4.5)$$

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 $A=SDS^{-1}$ και αν $\|\cdot\|$ είναι οποιαδήποτε προσανατολισμένη κατά άξονα επαγόμενη νόρμα, τότε ο Bauer και ο Fike έδειξαν ότι

$$s_A(B) \leq \kappa(S) \|A - B\| \quad (2.4.6)$$

Συγκεκριμένα, εάν ο A είναι κανονικός, βρίσκουμε για οποιαδήποτε νόρμα $\|\cdot\|$, θεωρώντας ως μεγαλύτερη τη φασματική νόρμα, ότι

$$s_A(B) \leq \|A - B\|. \quad (2.4.7)$$

Για $\|\cdot\| = \|\cdot\|_F$ και για A, B κανονικούς πίνακες, από το αποτέλεσμα του Hoffman και του Wielandt προκύπτει ότι η (2.4.7) είναι ακόμη σε ισχύ για τη μεταβολή ιδιοτιμής. Έχουμε

$$u(A, B) \leq \|A-B\|_F. \quad (2.4.8)$$

Για A και B πραγματικούς συμμετρικούς ή ερμιτιανούς πίνακες αυτό το αποτέλεσμα χρησιμοποιείται συχνά.

2.4.2. Ένα νέο αποτέλεσμα για το $s_A(B)$

Για τυχαίους διαγωνοποιήσιμους πίνακες A το αποτέλεσμα (2.4.6) δεν μπορεί να δώσει κάποια αρκετά σαφή εκτίμηση για το δείκτη κατάστασης $\kappa(S)$. Το ακόλουθο θεώρημα περιέχει (ενόψει του Θεωρήματος 2.1.9.) μια άμεση εκτίμηση τόσο γενική όσο η (2.4.4) (και έτσι πιο γενικευμένη από την (2.4.6)) που ανάγεται στην (2.4.7) 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 3.2. Για μεταγενέστερη χρήση σημειώνουμε τις ιδιότητες:

$$\lim_{y \rightarrow 0^+} y^{-1} g(y) = 1, \quad (2.4.9)$$

$$n^{-1} y \leq g(y) \leq y, \quad 0 \leq y \leq n, \quad (2.4.10)$$

$$g(n) = 1, \quad (2.4.11)$$

$$(n^{-1} y)^{1/n} \leq g(y) \leq y^{1/n}, \quad y \geq n, \quad (2.4.12)$$

$$\lim_{y \rightarrow \infty^+} y^{-1} g(y) = 1. \quad (2.4.13)$$

Θεώρημα 2.4.1. Έστω A ένας μη-κανονικός πίνακας και έστω $B - A \neq 0$. Εάν $\|\cdot\|$ είναι οποιαδήποτε νόρμα μεγαλύτερη από την φασματική νόρμα και εάν

$$y = \frac{\delta_{\|\cdot\|}(A)}{\|B-A\|},$$

τότε

$$s_A(B) \leq \frac{y}{g(y)} \|B - A\| = \frac{\delta_{\|\cdot\|}(A)}{g(y)}. \quad (2.4.14)$$

Παρατηρήσεις: Οι σχέσεις (2.4.10), (2.4.11) και (2.4.12) μπορούν να χρησιμεύσουν για να καταστεί το φράγμα στο (2.4.14) πιο σαφές. Η σχέση (2.4.9) δείχνει ότι, για $\delta_{\|\cdot\|}(A) \rightarrow 0$ και το όριο του $\|B - A\|$ να απομακρύνεται από το μηδέν, η εκτίμηση (2.4.14) προσεγγίζει την (2.4.7). (Ωστόσο, δεν μπορούμε να λάβουμε την (2.4.8) μ' αυτόν τον τρόπο.) Η σχέση (2.4.13) δείχνει ότι για ένα σταθερό μη-κανονικό πίνακα A και για $B \rightarrow A$ το φράγμα στη σχέση (2.4.14) είναι της ίδιας τάξης με την (2.4.4). Επίσης, ο αριθμητικός συντελεστής μπορεί να είναι μεγαλύτερος ή μικρότερος από τον συντελεστή στην (2.4.4) ανάλογα με την απομάκρυνσή του από την κανονικότητα του A .

Απόδειξη του Θεωρήματος 2.4.1. Έστω $A = U T U^*$, όπου $T = D + M$ είναι μια Schur τριγωνική μορφή του A . Θέτοντας $E = B - A$, $U^* B U = B_1$ και $U^* E U = F$, έχουμε

$$B_1 = U^* B U = U^*(E+A)U = U^* E U + U^* A U = F + U^*(U T U^*)U = F + T = F + D + M. \text{ Άρα}$$

$$B_1 = D + M + F.$$

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

$$0 = \det(B_1 - \mu I) = \det(D + M - \mu I + F) = \det(D + M - \mu I) \det[I + (D + M - \mu I)^{-1} F].$$

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 (2.1.1) προκύπτει ότι

$$\|(D + M - \mu I)^{-1} F\|_2 \geq |\lambda| = 1$$

ή, από την (iv), αφού $\|F\|_2 = \|E\|_2$, θα πάρουμε

$$\|(D + M - \mu I)^{-1}\|_2 \|E\|_2 \geq \|D + M - \mu I\|_2^{-1} \geq 1$$

Άρα,

$$\|(D+M-\mu I)^{-1}\|_2 \geq \frac{1}{\|E\|_2} \quad (2.4.15)$$

Τώρα θα πρέπει να εκτιμήσουμε τη $\|(D+M-\mu I)^{-1}\|_2$ από παραπάνω. Επειδή ο πίνακας $D - \mu I$ είναι αντιστρέψιμος, έχουμε

$$(D+M-\mu I)^{-1} = [I + (D-\mu I)^{-1}M]^{-1}(D-\mu I)^{-1}. \quad (2.4.16)$$

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,

$$[I + (D-\mu I)^{-1}M]^{-1} = I - (D-\mu I)^{-1}M + \dots + (-1)^{n-1} [(D-\mu I)^{-1} M]^{n-1}. \quad (2.4.17)$$

Θέτουμε για λόγους συντομίας

$$\|(D-\mu I)^{-1}\|_2 = p, \|M\|_2 = m (\neq 0), \|E\|_2 = e$$

και έχουμε από την (2.4.16) και την (2.4.17) ότι

$$\|D+M-\mu I\|_2 \leq p + p^2 m + \dots + p^n m^{n-1} = m^{-1} f(mp)$$

όπου η f ορίζεται από την (2.3.4).

Συνδυάζοντας το τελευταίο αποτέλεσμα με την (2.4.15) βρίσκουμε ότι

$$f(mp) \geq me^{-1}$$

και έτσι, από τον ορισμό της g , έχουμε $mp \geq g(me^{-1})$ ή

$$\frac{1}{p} \leq \frac{m}{g(me^{-1})}. \quad (2.4.18)$$

Επειδή η $\|\cdot\|$ 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 (2.4.18), προκύπτει ότι

$$\min_{1 \leq j \leq n} |\lambda_j - \mu| \leq \frac{m}{g(me^{-1})}. \quad (2.4.19)$$

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

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

$$\leq \frac{m}{g(me^{-1})} = \frac{y_0}{g(y_0)} \|B - A\|_2, \quad y_0 = \frac{\|M\|_2}{\|B - A\|_2}.$$

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

$$s_A(B) \leq \frac{\|M\|_2}{g\left(\frac{\|M\|_2}{\|B - A\|_2}\right)} = \|B - A\| \frac{y_1}{g(y_1)}, \quad \text{με } y_1 = \frac{\|M\|_2}{\|B - A\|}. \quad (2.4.20)$$

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

$$\frac{y_1}{g(y_1)} = \frac{f(y_1)}{x_1} < \frac{f(y_2)}{x_2} = \frac{y_2}{g(y_2)}.$$

Επίσης η συνάρτηση $y [g(y)]^{-1}$ είναι γνησίως αύξουσα και αντικαθιστώντας τη με $\|M\|_2$ στην (2.4.20), παίρνουμε

$$s_A(B) \leq \frac{y_2}{g(y_2)} \|B - A\|, \quad y_2 = \frac{\|M\|}{\|B - A\|}.$$

Η πλήρης απόδειξη του Θεωρήματος 2.4.1. προκύπτει, όπως παραπάνω, από τη συνέχεια της g .

□

2.4.3. Αριθμητικό παράδειγμα

Απεικονίζουμε την αριθμητική παράσταση των διαφόρων φραγμάτων που συζητήθηκαν παραπάνω εκτιμώντας τη φασματική μεταβολή των δύο πινάκων

$$A_1 = \begin{pmatrix} 1 & 10^{-4} \\ 0 & 1 \end{pmatrix} \text{ και } A_2 = \begin{pmatrix} 1 & 1 \\ 0 & 1 - 10^{-4} \end{pmatrix}$$

υπό την προϋπόθεση ότι

$$\|B_i - A_i\|_F \leq 10^{-2}, i = 1, 2. \quad (2.4.21)$$

Θέτουμε

$$s_i = \sup s_{A_i}(B_i), i = 1, 2.$$

όπου το supremum (= το ελάχιστο άνω φράγμα) λαμβάνεται σε σχέση με όλους τους πίνακες B_i που ικανοποιούν την (2.4.21).

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 $s_1 \geq 0.00707$ και $s_2 \geq 0.10000$.

Χρησιμοποιώντας $M=1.02$, $p_{\|\cdot\|_{L_1}, \|\cdot\|_F} = 2$, το φράγμα (2.4.4) δίνει $s_i \leq 0.57132$, $i = 1, 2$ και στις δύο περιπτώσεις.

Το φράγμα στην (2.4.6) δεν εφαρμόζεται στο A_1 , αφού ο πίνακας έχει έναν μη-γραμμικό στοιχειώδη διαιρέτη. Η σχέση $A_2 = SDS^{-1}$ 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 (2.4.6) αποδίδει

$$s_2 \leq 20.00001.$$

Στρεφόμενοι προς την (2.4.14), βρίσκουμε $\delta_{\|\cdot\|_F}(A_2) = 10^{-4}$, $y = 10^{-2}$, $g(y) = 0.009902$ 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 (2.4.14) 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.

2.4.4. Σχετικά αποτελέσματα για το $\nu(A, B)$

Για δεδομένους πίνακες A, B , που ικανοποιούν τις υποθέσεις του Θεωρήματος 2.4.1., έστω δ η ποσότητα στα δεξιά της (2.4.14), 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

$$D_i = \{\lambda \mid |\lambda - \lambda_i| \leq \delta\}, i = 1, 2, \dots, n, \text{ με } L_\delta = \bigcup_{i=1}^n D_i.$$

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 2.4.2. Έστω A, B και $\|\cdot\|$ ικανοποιούν τις υποθέσεις του Θεωρήματος 2.4.1. και έστω y ορίζεται όπως προηγουμένως. Τότε

$$\nu(A, B) \leq (2n-1) \frac{y}{g(y)} \|B - A\| = \frac{(2n-1)\delta_{\|\cdot\|}(A)}{g(y)} \quad (\text{αφού } y = \frac{\delta_{\|\cdot\|}(A)}{\|B-A\|}). \quad (2.4.22)$$

Θα πρέπει να σημειωθεί ότι αυτό το αποτέλεσμα δε συνεπάγεται τον τύπο (2.4.8) του Hoffmann-Wielandt.

2.4.5. 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 $A=SDS^{-1}$ και εάν $\|\circ\|$ υποδηλώνει τη νόρμα πίνακα επαγόμενη από τη $\|\cdot\|$, τότε

$$|\lambda_i - \lambda| \leq \kappa(S) \|Ax - \lambda x\| \quad (2.4.23)$$

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 3.1. 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 2.4.1., έλαβε το ακόλουθο αποτέλεσμα.

Θεώρημα 2.4.3 (D.D. Morrison). Έστω A ένας μη-κανονικός πίνακας, έστω $\|\cdot\|$ μια νόρμα διανύσματος 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

$$|\lambda_i - \lambda| \leq \frac{y}{g(y)} \|Ax - \lambda x\|. \quad (2.4.24)$$

2.5. Αριθμητικό πεδίο

Ορισμός 2.5.1. 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

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

Παραδείγμα 2.5.2.

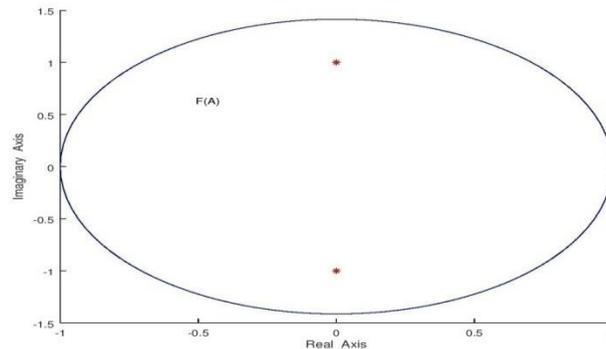
$$F\left(\begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix}\right) = \left\{ [\bar{x}_1 \ \bar{x}_2] \begin{bmatrix} 0 & 2 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} : |x_1|^2 + |x_2|^2 = 1 \right\}$$

$$= \{ 2x_2\bar{x}_1 : |x_1|^2 + |x_2|^2 = 1 \},$$

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

Παράδειγμα 2.5.3.

Έστω ο πίνακας $A = \begin{bmatrix} i & 2 \\ 0 & -i \end{bmatrix}$. Τότε, γραφικά το αριθμητικό πεδίο του A θα είναι:



2.5.1. Κλασικά αποτελέσματα

Πρόταση 2.5.4. Για κάθε πίνακα $A \in \mathbb{C}^{n \times n}$, το $F(A)$ είναι ένα συμπαγές (δηλαδή, κλειστό και φραγμένο) υποσύνολο του \mathbb{C} .

Απόδειξη: Το $F(A)$, όπως παρατηρήσαμε είναι η εικόνα της μοναδιαίας σφαίρας $\{x \in \mathbb{C}^n, x^* x = 1\}$ στο μιγαδικό επίπεδο, μέσω της απεικόνισης $x \rightarrow x^* A x$. Όμως η σφαίρα είναι συμπαγές σύνολο και η εν λόγω απεικόνιση συνεχής. Από το Θεώρημα

του Weierstrass το $F(A)$ είναι συμπαγές υποσύνολο του μιγαδικού επιπέδου (ως συνεχής εικόνα συμπαγούς συνόλου).

□

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 .

Θεώρημα 2.5.5 (Ελλειπτικό Θεώρημα). Έστω $A \in C^{2 \times 2}$ 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 $\sqrt{\text{trace}(A^*A) - |\lambda_1|^2 - |\lambda_2|^2}$.

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.

2.5.2. Η απόσταση του συνόρου του αριθμητικού πεδίου από τη κυρτή θήκη των ιδιοτιμών του πίνακα A

Σε αυτό το μέρος οι προτάσεις που ακολουθούν θα δείξουν, για συγκεκριμένες νόρμες, την ύπαρξη μιας απλής σύνδεσης μεταξύ της μέγιστης απόστασης του συνόρου του $F(A)$ από την κυρτή θήκη των ιδιοτιμών και της απομάκρυνσης από την κανονικότητα.

Θεώρημα 2.5.6. Εάν ξ 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

$$|\xi - \eta| \leq \sqrt{\frac{1-n^{-1}}{2}} \delta_{\|\cdot\|_F}(A) = \sqrt{\frac{n-1}{2n}} \delta_{\|\cdot\|_F}(A) \quad (2.5.1)$$

ή αν συνδυαστεί με την (2.1.7) έχουμε

$$|\xi - \eta| \leq \sqrt{\frac{n-1}{2n}} \|M\|_F = \sqrt{\frac{n-1}{2n}} \sqrt{\|A\|_F^2 - \sum_{i=1}^n |\lambda_i|^2}.$$

Η σταθερά $\sqrt{\frac{1-n^{-1}}{2}}$ δεν μπορεί να αντικατασταθεί από οποιαδήποτε άλλη μικρότερη σταθερά. (Σημείωση: Το φράγμα (2.5.1) έχει βρεθεί ξεχωριστά από τον P.J. Eberlein.)

Απόδειξη: Έστω $A = UTU^*$, όπου U είναι ορθομοναδιαίος πίνακας και $T = D + M$ είναι μια Schur τριγωνική μορφή του A . Έστω, επίσης, $\xi = x^*Ax$, όπου $x^*x = 1$. Θέτοντας, $U^*x = y$, έχουμε $y^*y = (U^*x)^*U^*x = x^*UU^*x = x^*x = 1$ και

$$\xi = x^*(UTU^*)x = (U^*x)^*(D + M)(U^*x) = y^*(D + M)y = y^*Dy + y^*My.$$

Άρα,

$$\xi = \eta + y^*My, \quad (2.5.2)$$

όπου $\eta = y^*Dy$, είναι ένα σημείο στο $\text{Co}(\sigma(A))$. Ως εκ τούτου για αυτόν τον αριθμό η ισχύει

$$|\xi - \eta| = |y^*My|.$$

Σημείωση: Το η δεν είναι το πλησιέστερο στο $\xi \in \text{Co}(\sigma(A))$ σημείο της κυρτής θήκης του φάσματος.

Θέτοντας $M = (m_{ij})$ και $y = y_i$, βρίσκουμε από την ανισότητα του Cauchy

$$|y^*My|^2 = |\sum_{i<j} m_{ij} \bar{y}_i y_j|^2 \leq \sum_{i<j} |m_{ij}|^2 \sum_{i<j} |\bar{y}_i y_j|^2 = [\delta_{\|\cdot\|_F}(A)]^2 \sum_{i<j} |\bar{y}_i y_j|^2.$$

Για τον τελευταίο παράγοντα βρίσκουμε την εκτίμηση

$$\begin{aligned} \sum_{i<j} |\bar{y}_i y_j|^2 &= \sum_{i<j} |y_i|^2 |y_j|^2 \leq \frac{1}{2} \{(\sum_{i=1}^n |y_i|^2)^2 - \sum_{i=1}^n |y_i|^4\} \\ &\leq \frac{1}{2} \left\{ (\sum_{i=1}^n |y_i|^2)^2 - \frac{1}{n} (\sum_{i=1}^n |y_i|^2)^2 \right\} = \frac{1}{2} \left(1 - \frac{1}{n}\right). \end{aligned}$$

Έτσι, θα βρούμε

$$|y^*My| \leq \sqrt{\frac{1-n^{-1}}{2}} \delta_{\|\cdot\|_F}(A)$$

αποδεικνύοντας την (2.5.1).

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

$$\delta_{\|\cdot\|_F}(A) = \|M\|_F = \left(\sum_{i < j} c^2\right)^{\frac{1}{2}} = \left[\frac{1}{2}n(n-1)\right]^{\frac{1}{2}} c.$$

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

$$y^*M y = \sum_{i < j} \frac{c}{n} = \frac{c}{n} \frac{1}{2} n(n-1) = \frac{c}{2} (n-1) = \sqrt{\frac{1-n^{-1}}{2}} \delta_{\|\cdot\|_F}(A)$$

□

Ένα παρόμοιο αποτέλεσμα επίσης ισχύει για τη l_1 -νόρμα.

Θεώρημα 2.5.7. Εάν ξ 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

$$|\xi - \eta| \leq \frac{1}{2} \delta_{\|\cdot\|_{l_1}}(A). \quad (2.5.3)$$

Η σταθερά $\frac{1}{2}$ δεν μπορεί να αντικατασταθεί από οποιαδήποτε άλλη μικρότερη σταθερά. Η σταθερά αυτή είναι η βέλτιστη δυνατή.

Απόδειξη. Γράφουμε ξανά το ξ στη μορφή (2.5.2), αλλά αυτή τη φορά θα εκτιμήσουμε 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.

Η τιμή του maximum τότε είναι $\frac{1}{2}$. Έτσι, παίρνουμε

$$|y^*M y| \leq \frac{1}{2} \|M\|_{l_1}. \quad (2.5.4)$$

Η ανισότητα (2.5.3) τώρα ακολουθεί ενόψει του γεγονότος ότι η (2.5.4) 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 (2.5.3) 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

$$y^* M y = \frac{1}{2} m = \frac{1}{2} \|M\|_{l_1}. \quad (2.5.5)$$

Έστω τώρα $M' = (m'_{ij})$ τα εκτός διαγωνίου στοιχεία οποιασδήποτε άλλης Schur τριγωνικής μορφής του A . Ενόψει της σχέσης: $\|M'\|_{l_1} \geq \|M'\|_F = \|M\|_F = m$, έχουμε $\delta_{\|\cdot\|_{l_1}}(A) = m$ και η (2.5.5) μπορεί να αντικατασταθεί με

$$y^* M y = \frac{1}{2} \delta_{\|\cdot\|_{l_1}}(A),$$

δείχνοντας ότι η ισότητα μπορεί να ισχύσει στην (2.5.3).

□

2.5.3. 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 2.5.6 και 2.5.7. Όλα τα άλλα σημεία έχουν μια μικρότερη απόσταση.

Το παραπάνω γεγονός γενικεύεται από το Θεώρημα 2.5.8. Για $\lambda, \mu \in \mathbb{C}$ και $c \in \mathbb{R}$, όπου $c \geq 0$, συμβολίζουμε με $E(\lambda, \mu, c)$ το συμπαγές σύνολο σημείων που στο μιγαδικό επίπεδο ορίζεται από την έλλειψη με εστίες λ, μ και μικρό ημιάξονα c .

Θεώρημα 2.5.8. Έστω $\lambda_i, i = 1, 2, \dots, n$, είναι 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

$$E(\lambda_i, \lambda_j, \frac{1}{2}\alpha), 1 \leq i < j \leq n.$$

Απόδειξη. Έστω ένας πίνακας A που ικανοποιεί την υπόθεση και έστω ξ, y_i, D και $M = (m_{ij})$ ορίζονται όπως στην απόδειξη του Θεωρήματος 2.5.7. 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. 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

$$\begin{aligned} \xi &= \sum_{k=1}^n \lambda_k |y_k|^2 + \sum_{i < j} m_{ij} \bar{y}_i y_j \\ &= \sum_{i < j} |m_{ij}| \left\{ \theta_{ij} \bar{y}_i y_j + \frac{1}{m} \sum_{k=1}^n \lambda_k |y_k|^2 \right\} \\ &= \frac{1}{m} \sum_{i < j} |m_{ij}| \left\{ m \theta_{ij} \bar{y}_i y_j + \sum_{k=1}^n \lambda_k |y_k|^2 \right\}, \end{aligned}$$

δείχνοντας ότι το ξ ανήκει στην κυρτή θήκη των σημείων

$$\eta_{ij} = m \theta_{ij} \bar{y}_i y_j + \sum_{k=1}^n \lambda_k |y_k|^2, 1 \leq i < j \leq n.$$

Θα δείξουμε στη συνέχεια ότι τα σημεία η_{ij} ανήκουν στο $\text{Co}(\sigma_m(A))$, δηλαδή στην κυρτή θήκη των συνόλων

$$E(\lambda_k, \lambda_i; \frac{1}{2}m), 1 \leq k < l = n.$$

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 $\text{Co}(\sigma_m(A))$.

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,

$$\begin{aligned} \eta_{ij} &= (|y_i|^2 + |y_j|^2) (\lambda_i |z_i|^2 + \lambda_j |z_j|^2 + \theta_{ij} m \bar{z}_i z_j) \\ &+ (1 - |y_i|^2 - |y_j|^2) \sum_{\substack{k=1 \\ k \neq i, j}}^n \lambda_k |z_k|^2. \end{aligned}$$

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 $\lambda_k (k \neq i, j)$ ανήκουν όλα στο $\text{Co}(\sigma_m(A))$. Έτσι ξανά το η_{ij} ανήκει στο $\text{Co}(\sigma_m(A))$. Προκύπτει επίσης ότι το ξ ανήκει στο $\text{Co}(\sigma_m(A))$.

Το παραπάνω είναι σε ισχύ για οποιοδήποτε πίνακα M που μπορεί να εμφανιστεί σε μια Schur τριγωνική μορφή του A . Έτσι το ξ ανήκει στη τομή όλων των πιθανών συνόλων $\text{Co}(\sigma_m(A))$. Αφού τα σύνολα $\text{Co}(\sigma_m(A))$ είναι κλειστά και αφού εξαρτώνται συνεχώς από το m , προκύπτει ότι το ξ ανήκει στο $\text{Co}(\sigma_{\inf m}(A))$, δηλαδή, στο $\text{Co}(\sigma(A))$.

□

Παρατήρηση: Η ένωση των αριθμητικών πεδίων όλων των πινάκων A με ιδιοτιμές $\lambda_i, i = 1, 2, \dots, n$ και δοθέν $\delta_{\|\cdot\|_1}(A) = m$, δεν συμπληρώνει το σύνολο $\text{Co}(\sigma_m(A))$.

Κεφάλαιο 3

Φράγματα για την απομάκρυνση από την κανονικότητα

Η απομάκρυνση από την κανονικότητα ενός πίνακα είναι ένας πραγματικόβαθμωτό μέγεθος που μπορεί να χρησιμοποιηθεί για να υπολογίσουμε φράγματα για διάφορους υπολογισμούς πινάκων. Για παράδειγμα, η απομάκρυνση από την κανονικότητα μπορεί να χρησιμοποιηθεί για να φράξουμε δυνάμεις, αντιστρόφους, φάσμα και αριθμητικά πεδία μη κανονικών πινάκων. Δυστυχώς, η απομάκρυνση από την κανονικότητα είναι δύσκολο να υπολογισθεί αν ο πίνακας έχει μεγάλες διαστάσεις και οι ιδιοτιμές του είναι άγνωστες.

Σ' αυτό το κεφάλαιο θα υπολογίσουμε ένα καινούργιο άνω φράγμα για την απομάκρυνση από την κανονικότητα ενός τετραγωνικού πίνακα που απαιτεί λιγότερους υπολογισμούς από το φράγμα για την απομάκρυνση από την κανονικότητα του Herfici και είναι βέλτιστο για κάθε πίνακα με ιδιοτιμές οριζόντια ή κάθετα ευθυγραμμισμένες στο μιγαδικό επίπεδο.

Έστω A ένας $n \times n$ μιγαδικός πίνακας και έστω A^* ο ανάστροφος συζυγής του A . Γνωρίζουμε ότι A λέγεται *ερμιτιανός* αν και μόνον αν $A^* = A$ και A λέγεται *αντιερμιτιανός* αν και μόνον αν $A^* = -A$. Έστω H και S το ερμιτιανό και αντιερμιτιανό μέρος του A , αντίστοιχα. Πιο συγκεκριμένα, έστω ότι οι συναρτήσεις $H(\circ)$ και $S(\circ)$ παριστάνουν το ερμιτιανό και το αντιερμιτιανό μέρος, αντίστοιχα, για κάθε τετραγωνικό πίνακα. Τότε, για

$$H(A) := \frac{1}{2}(A + A^*) \equiv H \text{ και } S(A) := \frac{1}{2}(A^* - A) \equiv S,$$

ο πίνακας A μπορεί να γραφτεί ως:

$$A = H - S.$$

Έστω, όπως ορίσαμε πριν, T είναι ένας άνω τριγωνικός πίνακας, M είναι ένας γνήσια άνω τριγωνικός πίνακας και D είναι διαγώνιος πίνακας με στοιχεία τις ιδιοτιμές λ_i του A . Έστω $\text{Re}(D)$ και $\text{Im}(D)$ είναι διαγώνιοι πίνακες με στοιχεία τα πραγματικά και φανταστικά μέρη των ιδιοτιμών του A , αντίστοιχα. Επίσης, έστω η τριγωνοποίηση Schur του πίνακα A : $U^*AU = T$, όπου U οποιοσδήποτε

ορθομοναδιαίος πίνακας και $T = D + M$. Ξανά θα χρησιμοποιήσουμε τις νόρμες $\|\cdot\|_2$ και $\|\cdot\|_F$ επειδή είναι *ορθομοναδιαία αναλλοίωτες*.

Από τον ορισμό (γνωστές προτάσεις), οι κανονικοί πίνακες περιλαμβάνουν ερμιτιανούς, αντιερμιτιανούς, ορθομοναδιαίους πίνακες και γενικά κάθε τετραγωνικό πίνακα που είναι ορθομοναδιαία όμοιος με διαγώνιο πίνακα. Άρα, από Λήμμα του Schur, για κάθε κανονικό πίνακα A ισχύει: $U^*AU = T = D + M$ με $M = 0$. Συνεπώς, για ένα μη κανονικό πίνακα είναι λογικό να εκφράσουμε την απομάκρυνσή του από την κανονικότητα σε σχέση με την νόρμα του πίνακα M .

Ορισμός 3.0.1 (Απομάκρυνση από την κανονικότητα). Για κάθε $n \times n$ πίνακα A ,

$$\delta_{\|\cdot\|_F}(A) := \|M\|_F = (\|A\|_F^2 - \|D\|_F^2)^{1/2}. \quad (3.0.1)$$

Είναι επίσης εύκολο να δείξουμε ότι η $\delta_{\|\cdot\|_F}(A)$ είναι ανεξάρτητη από μιγαδικές μετατοπίσεις. Για παράδειγμα, αν κάνουμε την τριγωνοποίηση Schur του πίνακα $A - aI$, όπου $U^*(A - aI)U = T - aI = D + M - aI$, παρατηρούμε ότι η μετατόπιση αυτή αλλάζει μόνο τα διαγώνια στοιχεία του T και η $\delta_{\|\cdot\|_F}(A) := \|M\|_F$ δεν αλλάζει.

Άρα,

$$\delta_{\|\cdot\|_F}(A) = \delta_{\|\cdot\|_F}(e^{-i\theta}(A - aI)), \quad (3.0.2)$$

για κάθε μιγαδικό βαθμωτό a και $0 \leq \theta < 2\pi$.

Στη συνέχεια, θα δείξουμε και την σημαντικότητα αυτής της παρατήρησης.

Παράδειγμα 3.0.2. Έστω οι δύο διαφορετικές παραγοντοποιήσεις Schur του πίνακα

$$A = \begin{pmatrix} 4 & \sqrt{1/2} & -3 \\ \sqrt{2} & 2 & -\sqrt{2} \\ 1 & \sqrt{1/2} & 0 \end{pmatrix}$$

όπου U_1 και U_2 είναι ορθομοναδιαίοι πίνακες τέτοιοι ώστε

$$U_1^*AU_1 = T_1 = \begin{pmatrix} 1 & 1 & 4 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix} \text{ και } U_2^*AU_2 = T_2 = \begin{pmatrix} 2 & -1 & 3\sqrt{2} \\ 0 & 1 & \sqrt{2} \\ 0 & 0 & 3 \end{pmatrix}.$$

Οι τριγωνοποιήσεις Schur μας δείχνουν ότι ο πίνακας A έχει ιδιοτιμές $\lambda(A) = \{1, 2, 3\}$ και ότι ο A δεν είναι κανονικός πίνακας αφού οι T_1 και T_2 δεν είναι διαγώνιοι

πίνακες. Επιπλέον, τα αυστηρά άνω τριγωνικά στοιχεία των παραπάνω τριγωνοποιήσεων μας δείχνουν ότι $\delta_{\|\cdot\|_F}(A) := \|M\|_F = \sqrt{21}$.

Γενικά, η τιμή της $\delta_{\|\cdot\|_F}(A)$ είναι δύσκολη να υπολογιστεί όταν ο πίνακας A έχει μεγάλες διαστάσεις και δεν γνωρίζουμε τις ιδιοτιμές του. Μερικά κάτω φράγματα για την $\delta_{\|\cdot\|_F}(A)$ έχουν βρεθεί από τον Eberlein:

$$\delta_{\|\cdot\|_F}(A) \geq \frac{\|A^*A - AA^*\|_F}{\sqrt{6}\|A\|_F} \quad (3.0.4)$$

και από τον Loizou:

$$\delta_{\|\cdot\|_F}(A) \geq \frac{\|A^*A - AA^*\|_F}{\beta^{1/2} + (\beta + \sqrt{2}\|A^*A - AA^*\|_F)^{1/2}}, \quad (3.0.5)$$

όπου $\beta = \|A\|_F^2 - \frac{1}{n} |\text{trace}(A)|^2$ και $A \neq 0$.

3.1. Ένα βέλτιστο άνω φράγμα

Ένα τετριμμένο άνω φράγμα για την $\delta_{\|\cdot\|_F}(A)$ είναι συνέπεια του ορισμού της:

$$\delta_{\|\cdot\|_F}(A) = (\|A\|_F^2 - \|D\|_F^2)^{1/2} \leq \|A\|_F. \quad (3.1.1)$$

Μπορούμε, όμως, να βρούμε καλύτερα άνω φράγματα για την $\delta_{\|\cdot\|_F}(A)$, χρησιμοποιώντας εκφράσεις που προκύπτουν αφού χωρισούμε τον πίνακα A στο ερμιτιανό του μέρος H και στο αντιερμιτιανό του μέρος S .

Λήμμα 3.1.1. Αν ο $A = H - S$ έχει την τριγωνοποίηση κατά Schur $U^*AU = D + M$, τότε

$$\|H\|_F^2 = \|\text{Re}(D)\|_F^2 + \frac{1}{2} \|M\|_F^2 \quad (3.1.2)$$

και
$$\|S\|_F^2 = \|\text{Im}(D)\|_F^2 + \frac{1}{2} \|M\|_F^2. \quad (3.1.3)$$

Απόδειξη. Οι παρακάτω ισότητες είναι άμεση συνέπεια των ορισμών της νόρμας Frobenius και των συναρτήσεων $H(\cdot)$ και $S(\cdot)$:

$$\|M\|_F^2 = \|H(A)\|_F^2 = \|H(T)\|_F^2 = \|H(D)\|_F^2 + \|H(M)\|_F^2 = \|\operatorname{Re}(D)\|_F^2 + \frac{1}{2}\|M\|_F^2 \quad (3.1.4)$$

$$\|N\|_F^2 = \|S(A)\|_F^2 = \|S(T)\|_F^2 = \|S(D)\|_F^2 + \|S(M)\|_F^2 = \|\operatorname{Im}(D)\|_F^2 + \frac{1}{2}\|M\|_F^2 \quad (3.1.5)$$

□

Οι εξισώσεις (3.1.4) και (3.1.5) δείχνουν ότι η $\delta_{\|\cdot\|_F}(A)$ μπορεί να οριστεί σε δύο διαφορετικούς, αλλά ισοδύναμους τρόπους:

$$\delta_{\|\cdot\|_F}(A) = \sqrt{2} (\|H\|_F^2 - \|\operatorname{Re}(D)\|_F^2)^{1/2} \quad (3.1.6)$$

ή

$$\delta_{\|\cdot\|_F}(A) = \sqrt{2} (\|S\|_F^2 - \|\operatorname{Im}(D)\|_F^2)^{1/2}. \quad (3.1.7)$$

Ένα απλό άνω φράγμα για την $\delta_{\|\cdot\|_F}(A)$ προκύπτει άμεσα από αυτές τις ισότητες.

Λήμμα 3.1.2. Αν $A = H - S$, όπου H είναι το ερμιτιανό μέρος του A και S είναι το αντιερμιτιανό του μέρος, τότε

$$\delta_{\|\cdot\|_F}(A) \leq \sqrt{2} \min\{\|H\|_F, \|S\|_F\}. \quad (3.1.8)$$

Απόδειξη. Από τις ισότητες (3.1.6), (3.1.7) έχουμε:

$$\delta_{\|\cdot\|_F}(A) = \sqrt{2} (\|H\|_F^2 - \|\operatorname{Re}(D)\|_F^2)^{1/2} \leq \sqrt{2} \|H\|_F, \quad (3.1.9)$$

$$\delta_{\|\cdot\|_F}(A) = \sqrt{2} (\|S\|_F^2 - \|\operatorname{Im}(D)\|_F^2)^{1/2} \leq \sqrt{2} \|S\|_F \quad (3.1.10)$$

και το άνω φράγμα θα είναι

$$\delta_{\|\cdot\|_F}(A) \leq \sqrt{2} \min\{\|H\|_F, \|S\|_F\}. \quad (3.1.11)$$

□

Θεωρούμε τώρα τη χρήση μιγαδικών μετατοπίσεων αλγια να βελτιώσουμε το φράγμα (3.1.8). Όπως στο Λήμμα 3.1.1, μπορούμε να χωρίσουμε τον πίνακα $A - \alpha I$ στο ερμιτιανό του μέρος $S - \operatorname{Re}(\alpha)I$ και στο αντιερμιτιανό του μέρος $S - \operatorname{Im}(\alpha)I$ και παίρνουμε:

$$\delta_{\|\cdot\|_F}(A) = \delta_{\|\cdot\|_F}(A - \alpha I) = \sqrt{2} (\|H - \operatorname{Re}(\alpha)I\|_F^2 - \|\operatorname{Re}(D) - \operatorname{Re}(\alpha)I\|_F^2)^{1/2}, \quad (3.1.12)$$

$$\delta_{\|\cdot\|_F}(A) = \delta_{\|\cdot\|_F}(A - \alpha I) = \sqrt{2} (\|S - \operatorname{Im}(\alpha)I\|_F^2 - \|\operatorname{Im}(D) - \operatorname{Im}(\alpha)I\|_F^2)^{1/2}. \quad (3.1.13)$$

Ένα αυστηρότερο άνω φράγμα θα προκύψει αν ελαχιστοποιήσουμε τους όρους

$$\|Re(D) - Re(\alpha)I\|_F^2 \text{ και } \|Im(D) - Im(\alpha)I\|_F^2 \quad (3.1.14)$$

στις εξισώσεις (3.1.12) και (3.1.13), αντίστοιχα.

Πιο συγκεκριμένα,

$$f_1(Re(a)) = \|Re(D) - Re(\alpha)I\|_F^2 = \sum (Re(\lambda_i) - Re(a))^2 \quad (3.1.15)$$

και

$$f_2(Im(a)) = \|Im(D) - Im(\alpha)I\|_F^2 = \sum (Im(\lambda_i) - Im(a))^2 \quad (3.1.16)$$

είναι τετραγωνικές συναρτήσεις που μπορούν να ελαχιστοποιηθούν χρησιμοποιώντας βασικούς τρόπους ανάλυσης. Λύνοντας τις εξισώσεις $f_1' = 0$ και $f_2' = 0$, έχουμε:

$$Re(a) = \frac{\sum Re(\lambda_i)}{n} \text{ και } Im(a) = \frac{\sum Im(\lambda_i)}{n}. \quad (3.1.17)$$

Αυτές οι τιμές ελαχιστοποιούν τις εξισώσεις (3.1.15) και (3.1.16) αντίστοιχα, αφού οι f_1'' και f_2'' είναι θετικές. Συνεπώς, και οι δύο όροι της (3.1.14) ελαχιστοποιούνται για

$$\alpha = Re(a) + Im(a) = \frac{\sum Re(\lambda_i)}{n} + \frac{\sum Im(\lambda_i)}{n} = \frac{\sum \lambda_i}{n} = \frac{trace(A)}{n}. \quad (3.1.18)$$

Θεώρημα 3.1.3. Αν $A = H - S$, τότε

$$\delta_{\|\cdot\|_F}(A) \leq \sqrt{2} \min \{\|H - Re(a)I\|_F, \|S - Re(a)I\|_F\}, \quad (3.1.19)$$

όπου το άνω φράγμα ελαχιστοποιείται για

$$\alpha = \frac{trace(A)}{n}. \quad (3.1.20)$$

Επιπλέον, το φράγμα είναι βέλτιστο, δηλαδή η ισότητα ισχύει, αν και μόνο αν οι ιδιοτιμές του A είναι οριζόντια ή κάθετα ευθυγραμμισμένες στο μιγαδικό επίπεδο.

Απόδειξη. Το άνω φράγμα προκύπτει από τις εξισώσεις (3.1.12) και (3.1.13) απαλοίφοντας τους όρους $\|Re(D) - Re(\alpha)I\|_F^2$ και $\|Im(D) - Im(\alpha)I\|_F^2$, αντίστοιχα. Το φράγμα είναι βέλτιστο αν και μόνο αν

$$\|Re(D) - Re(\alpha)I\|_F^2 = 0 \quad (3.1.21)$$

και

$$\|\operatorname{Im}(D) - \operatorname{Im}(\alpha)I\|_F^2 = 0. \quad (3.1.22)$$

Η πρώτη συνθήκη (3.1.21) μας δείχνει ότι τα πραγματικά μέρη των ιδιοτιμών του A είναι σταθερά (δηλαδή, είναι κάθετα ευθυγραμμισμένα). Η δεύτερη συνθήκη (3.1.22) μας δείχνει ότι τα φανταστικά μέρη των ιδιοτιμών του A είναι σταθερά (δηλαδή, είναι οριζόντια ευθυγραμμισμένα).

Οι εξισώσεις (3.1.12) και (3.1.13) μας δείχνουν επίσης ότι

$$\delta_{\|\cdot\|_F}(A) \approx \sqrt{2} \min\{\|H - \operatorname{Re}(\alpha)I\|_F, \|S - \operatorname{Im}(\alpha)I\|_F\} \quad (3.1.23)$$

αν και μόνο αν

$$\|\operatorname{Re}(D) - \operatorname{Re}(\alpha)I\|_F^2 \ll \|H - \operatorname{Re}(\alpha)I\|_F^2, \quad (3.1.24)$$

ή

$$\|\operatorname{Im}(D) - \operatorname{Im}(\alpha)I\|_F^2 \ll \|S - \operatorname{Im}(\alpha)I\|_F^2. \quad (3.1.25)$$

Συνεπώς, το καινούργιο φράγμα (3.1.19) είναι μία καλή προσέγγιση όταν οι ιδιοτιμές του A είναι σχετικά κοντά στο να είναι οριζόντια ή κάθετα ευθυγραμμισμένες. Αλλιώς, το φράγμα είναι αδύναμο.

□

Παράδειγμα 3.1.4. Θα κάνουμε τώρα μία σύγκριση μεταξύ του άνω φράγματος κατά Henrici (2.1.6) και του καινούργιου άνω φράγματος (3.1.19) για τον πίνακα A του Παραδείγματος 3.0.2.

Χρησιμοποιώντας τις ποσότητες $\|A^*A - AA^*\|_F = \sqrt{996}$, $\alpha = 2$, $\|M - 2I\|_F = \sqrt{25/2}$ και $\|N\|_F = \sqrt{21/2}$, το φράγμα κατά Henrici μας δίνει

$\delta_{\|\cdot\|_F}(A) \leq (2)^{1/4} (\sqrt{996})^{1/2} = (\sqrt{1992})^{1/2} \approx \sqrt{44.6318}$ και το καινούργιο άνω φράγμα μας δίνει τιμή $\delta_{\|\cdot\|_F}(A) \leq \min\{\sqrt{\frac{25}{2}}, \sqrt{\frac{21}{2}}\} = \sqrt{21}$.

Παρατηρούμε ότι το καινούργιο φράγμα είναι αυστηρότερο από το φράγμα κατά Henrici. Επιπλέον, το καινούργιο φράγμα είναι βέλτιστο, αφού όλες οι ιδιοτιμές του A είναι πραγματικές.

Και αν πάρουμε άλλα παραδείγματα, θα παρατηρήσουμε ότι το καινούργιο φράγμα είναι αυστηρότερο από το φράγμα του Henrici. Γενικά, το καινούργιο φράγμα είναι

προτιμότερο, αφού το φράγμα του Henricία απαιτεί $O(n^3)$ υπολογισμούς, ενώ αυτό απαιτεί $O(n^2)$ υπολογισμούς.

Μερικές φορές είναι δυνατό να βελτιώσουμε περαιτέρω το καινούργιο φράγμα στρέφοντας το $A - \alpha I$. Για μιγαδικούς πίνακες, οι ιδιοτιμές του πίνακα $A - \alpha I$ μπορεί να προκύψουν αυθαίρετα και η καλύτερη στροφή θ δεν μπορεί να βρεθεί θεωρητικά. Για πραγματικούς πίνακες, οι ιδιοτιμές του $A - \alpha I$ προκύπτουν σε μιγαδικά-συζυγή ζεύγη και το καινούργιο φράγμα ελαχιστοποιείται για $\theta \in \{0, \frac{\pi}{2}, \pi, \frac{3\pi}{2}\}$. Παρατηρούμε ότι η χρήση του $\theta=0$ έχει το πλεονέκτημα ότι το καινούργιο φράγμα γίνεται μηδέν για ερμιτιανούς και αντιερμιτιανούς πίνακες. Δυστυχώς, για κανονικούς πίνακες οι ιδιοτιμές των οποίων δεν είναι οριζόντια ή κάθετα ευθυγραμμισμένες, το καινούργιο φράγμα δεν μειώνεται σε μηδέν. Σε αντίθεση με τις βελτιώσεις για το καινούργιο φράγμα, το φράγμα κατά Henrici δεν επηρεάζεται από μιγαδικές μετατοπίσεις ή στροφές του πίνακα A . Αυτό συμβαίνει διότι για μιγαδικές μετατοπίσεις του A έχουμε:

$$\begin{aligned} & \|(A - \alpha I)^*(A - \alpha I) - (A - \alpha I)(A - \alpha I)^*\|_F \\ &= \|(A^*A - a^*A - aA^* + a^*aI) - (AA^* - a^*A - aA^* + aa^*I)\|_F \\ &= \|A^*A - AA^*\|_F. \end{aligned}$$

Το ίδιο αποτέλεσμα θα παίρναμε και στην γενική περίπτωση του $e^{-i\theta}(A - \alpha I)$.

3.2. Φράγματα για την απομάκρυνση από την κανονικότητα και η Frobenius νόρμα των ιδιοτιμών ενός πίνακα

Σ' αυτό το σημείο θα περιγράψουμε καινούργια άνω και κάτω φράγματα για την απομάκρυνση από την κανονικότητα και την νόρμα Frobenius νόρμα των ιδιοτιμών ενός πίνακα. Θα περιγράψουμε επίσης και κάποιες σημαντικές ιδιότητες αυτών των φραγμάτων. Για παράδειγμα, θα παρουσιάσουμε μία βελτίωση του άνω φράγματος των ιδιοτιμών ενός πίνακα που είχαν βρει οι Kress, deVries και Wegmann.

Μερικές σχέσεις που θα μας χρειαστούν:

$$\|D\|_F^2 = \|A\|_F^2 - \delta_{\|\cdot\|_F}^2(A), \quad (3.2.1)$$

$$\|\operatorname{Re}(D)\|_F^2 = \|H\|_F^2 - \frac{1}{2} \delta_{\|\cdot\|_F}^2(A), \quad (3.2.2)$$

$$\|\operatorname{Im}(D)\|_F^2 = \|S\|_F^2 - \frac{1}{2} \delta_{\|\cdot\|_F}^2(A). \quad (3.2.3)$$

Επίσης, όπως είδαμε παραπάνω, η τιμή $a = \frac{\operatorname{tr}(A)}{n}$ ελαχιστοποιεί την τετραγωνική συνάρτηση $\|A - \alpha I\|_F^2$.

Αν $\operatorname{tr}(A) = 0$, θα λέμε ότι ο A είναι ένας *κεντραρισμένος πίνακας*. Κεντραρισμένοι πίνακες, τέτοιοι ώστε

$$\tilde{A} = A - \frac{\operatorname{tr}(A)}{n} I \quad (3.2.4)$$

και

$$\tilde{D} = D - \frac{\operatorname{tr}(A)}{n} I \quad (3.2.5)$$

θα συμβολίζονται με μία περισπωμένη. Παρακάτω δίνουμε ένα λήμμα που συνδέει τη νόρμα των μετατοπισμένων πινάκων $A - \alpha I$ και $D - \alpha I$ με την νόρμα των κεντραρισμένων πινάκων \tilde{A} και \tilde{D} , αντίστοιχα.

Λήμμα 3.2.1. Για κάθε $n \times n$ πίνακα A και για κάθε μιγαδικό αριθμό α ,

$$\|A - \alpha I\|_F^2 = \|\tilde{A}\|_F^2 + \frac{|\operatorname{tr}(A - \alpha I)|^2}{n} I \quad (3.2.6)$$

και

$$\|D - \alpha I\|_F^2 = \|\tilde{D}\|_F^2 + \frac{|\operatorname{tr}(A - \alpha I)|^2}{n} I. \quad (3.2.7)$$

Απόδειξη. Αρχικά σχετίζουμε τη νόρμα του A με τη νόρμα του \tilde{A} . Για $\sigma = \frac{\operatorname{tr}(A)}{n}$ έχουμε ότι

$$\|A\|_F^2 - \|\tilde{A}\|_F^2 = \|A\|_F^2 - \left\| A - \frac{\operatorname{tr}(A)}{n} I \right\|_F^2 \quad (3.2.8)$$

$$= \sum (|a_{ii}|^2) - \sum (|a_{ii} - \sigma|^2) \quad (3.2.9)$$

$$= \sum (a_{ii}^* a_{ii}) - \sum [a_{ii}^* a_{ii} - a_{ii}^* \sigma - \sigma^* a_{ii} + \sigma^* \sigma] \quad (3.2.10)$$

$$= \sigma \sum (a_{ii}^*) + \sigma^* \sum (a_{ii}) - n \sigma^* \sigma \quad (3.2.11)$$

$$= \frac{\operatorname{tr}(A)}{n} \operatorname{tr}^*(A) + \frac{\operatorname{tr}^*(A)}{n} \operatorname{tr}(A) - n \left| \frac{\operatorname{tr}(A)}{n} \right|^2 \quad (3.2.12)$$

$$= \frac{2|\text{tr}(A)|^2}{n} - \frac{|\text{tr}(A)|^2}{n} \quad (3.2.13)$$

$$= \frac{|\text{tr}(A)|^2}{n}. \quad (3.2.14)$$

Αν αντικαταστήσουμε τώρα τον πίνακα A με $A - \alpha I$ στη δεξιά μεριά των σχέσεων (3.2.8) και (3.2.14), παίρνουμε

$$\|A - \alpha I\|_F^2 - \left\| A - \alpha I - \frac{\text{tr}(A - \alpha I)}{n} I \right\|_F^2 = \frac{|\text{tr}(A - \alpha I)|^2}{n} \quad (3.2.15)$$

και ο δεύτερος όρος απλοποιείται μέσω της σχέσης

$$\left\| A - \alpha I - \frac{\text{tr}(A - \alpha I)}{n} I \right\|_F^2 = \left\| A - \alpha I - \frac{[\text{tr}(A) - \text{tr}(\alpha I)]}{n} I \right\|_F^2 \quad (3.2.16)$$

$$= \left\| A - \alpha I - \frac{\text{tr}(A)}{n} I + \frac{n\alpha}{n} I \right\|_F^2 \quad (3.2.17)$$

$$= \left\| A - \frac{\text{tr}(A)}{n} I \right\|_F^2 \quad (3.2.18)$$

$$= \|\tilde{A}\|_F^2 \quad (3.2.19)$$

και έτσι καταλήγουμε στην προς απόδειξη σχέση (3.2.6). Η δεύτερη ισότητα (3.2.7) αποδεικνύεται με παρόμοιο τρόπο. □

Θα παρουσιάσουμε τώρα κάποια φράγματα για την $\|D\|_F^2$ και την $\delta_{\|\cdot\|}^2(A)$, μαζί με κάποιες σημαντικές ιδιότητές τους.

3.3. Φράγματα για τις ιδιοτιμές και την απομάκρυνση από την κανονικότητα

Θεώρημα 3.3.1. Για ένα μη-κανονικό πίνακα A ισχύει ότι

$$\|D\|_F^2 \leq \left(\|A\|_F^4 - \frac{1}{2} \|A^*A - AA^*\|_F^2 \right)^{1/2} \quad (3.3.1)$$

και η ισότητα ισχύει αν και μόνο αν

$$A = \gamma(vw^* + rwnv^*), \quad (3.3.2)$$

όπου γ είναι μη μηδενικό μιγαδικό βαθμωτό, $0 \leq r < 1$ είναι ένας πραγματικός αριθμός και v, w είναι ορθοκανονικά διανύσματα.

Ένα πρακτικό κάτω φράγμα για τη $\|D\|_F^2$ είναι :

$$|tr(A^2)| \leq \|D\|_F^2. \quad (3.3.3)$$

Η σχέση αυτή αποδεικνύεται εύκολα από την τριγωνική ανισότητα εφαρμοσμένη για τις ιδιοτιμές του A^2 :

$$|tr(A^2)| = |tr(D^2)| = |\lambda_1^2 + \dots + \lambda_n^2| \quad (3.3.4)$$

$$\leq |\lambda_1^2| + \dots + |\lambda_n^2| = |\lambda_1|^2 + \dots + |\lambda_n|^2 = \|D\|_F^2. \quad (3.3.5)$$

Αυτό το κάτω φράγμα είναι βέλτιστο αν και μόνο αν είναι μηδέν και οι ιδιοτιμές του A είναι συγγραμμικές. Επιπλέον, το φράγμα είναι εύκολο για να υπολογιστεί αφού χρειαζόμαστε μόνο την διαγώνιο του A^2 . Αυτή η διαγώνιος μπορεί να υπολογιστεί με (το πολύ) m πολλαπλασιασμούς, όπου m είναι το πλήθος των μη μηδενικών στοιχείων του A .

Ένα κάτω φράγμα και ένα άνω φράγμα για την $\delta_{\|\cdot\|_F}^2(A)$, προκύπτει αντικαθιστώντας τις σχέσεις (3.3.1) και (3.3.3) στην (3.1.1) και είναι:

$$\|A\|_F^2 - \left(\|A\|_F^4 - \frac{1}{2} \|A^*A - AA^*\|_F^2 \right)^{1/2} \leq \delta_{\|\cdot\|_F}^2(A) \leq \|A\|_F^2 - |tr(A^2)|. \quad (3.3.6)$$

Το άνω φράγμα στην (3.3.6) είναι βέλτιστο αν και μόνο αν είναι μηδέν και οι ιδιοτιμές του A είναι συγγραμμικές και μπορεί να υπολογιστεί με (το πολύ) $2m$ πολλαπλασιασμούς. Το κάτω φράγμα είναι ένας $O(n^3)$ υπολογισμός και είναι βέλτιστο αν και μόνο αν ο A είναι κανονικός ή ικανοποιεί τη συνθήκη (3.3.2). Αυτό το κάτω φράγμα κληρονομεί τις ιδιότητες του άνω φράγματος (3.3.1) μέσω της (3.1.1.). Άρα, είναι το καλύτερο δυνατό σε σχέση με τα $\|A\|_F$ και $\|A^*A - AA^*\|_F$.

Παρακάτω, θα δείξουμε πώς μπορούμε να βελτιώσουμε αυτά τα φράγματα.

3.3.1.Βελτιωμένα κάτω φράγματα

Αφού η ποσότητα $\delta_{\|\cdot\|_F}^2(A)$ παραμένει αμετάβλητη για κάθε μιγαδική μετατόπιση α , θα δείξουμε πώς μπορεί αυτή η ελεύθερη παράμετρος να μεγιστοποιήσει το κάτω φράγμα στην (3.3.6). Για κανονικούς πίνακες, το κάτω φράγμα

$$\|A - \alpha I\|_F^2 - \left(\|A - \alpha I\|_F^4 - \frac{1}{2} \|A^*A - AA^*\|_F^2 \right)^{1/2} \leq \delta_{\|\cdot\|_F}^2(A - \alpha I) = \delta_{\|\cdot\|_F}^2(A) \quad (3.3.7)$$

είναι προφανώς μηδέν για κάθε επιλογή της παραμέτρου α . Για μη κανονικούς πίνακες, ωστόσο, υπάρχει μία μοναδική τιμή της α που μεγιστοποιεί την (3.3.7). Πιο συγκεκριμένα, αντικαθιστώντας την (3.2.6) στην (3.3.7), ζητάμε να μεγιστοποιήσουμε την συνάρτηση

$$f(z(\alpha)) = (\beta^2 + z^2(\alpha)) - \left[(\beta^2 + z^2(\alpha))^2 - \frac{1}{2} K^2 \right]^{1/2}, \quad (3.3.8)$$

όπου

$$z^2(\alpha) = \frac{|\text{tr}(A - \alpha I)|^2}{n} \quad (3.3.9)$$

και

$$\beta^2 = \|\tilde{A}\|_F^2, \quad K^2 = \|A^*A - AA^*\|_F^2 > 0. \quad (3.3.10)$$

Λύνοντας την εξίσωση

$$\frac{df}{dz} = 2z(1 - (\beta^2 + z^2)) \left[(\beta^2 + z^2)^2 - \frac{1}{2} K^2 \right]^{-1/2} = 0, \quad (3.3.11)$$

βρίσκουμε ότι η μοναδική λύση $z=0$ είναι ολικό μέγιστο αφού

$$\frac{d^2f}{dz^2}(0) = 2 \left(1 - \frac{\beta^2}{(\beta^4 - \frac{1}{2} K^2)^{1/2}} \right) < 0. \quad (3.3.12)$$

Τέλος, λύνοντας την

$$z(\alpha) = \frac{|\text{tr}(A - \alpha I)|}{n^{1/2}} = 0, \quad (3.3.13)$$

βρίσκουμε ότι το κάτω φράγμα μεγιστοποιείται για $\alpha = \frac{\text{tr}(A)}{n}$.

Σημείωση: Στη συνέχεια θα αποδείξουμε ότι η $\alpha = \frac{\text{tr}(A)}{n}$ βελτιώνει επίσης και το άνω φράγμα

$$\delta_{\|\cdot\|_F}^2(A) = \delta_{\|\cdot\|_F}^2(A - \alpha I) = \|\tilde{A}\|^2 - \|\tilde{D}\|^2 \leq \|\tilde{A}\|^2 - |\text{tr}(\tilde{A}^2)|. \quad (3.3.14)$$

Λήμμα 3.3.2. Για κάθε $n \times n$ πίνακα A ,

$$\delta_{\|\cdot\|_F}^2(A) \geq \|A - \alpha I\|_F^2 - \left(\|A - \alpha I\|_F^4 - \frac{1}{2} \|A^*A - AA^*\|_F^2 \right)^{1/2}, \quad (3.3.15)$$

όπου το κάτω φράγμα μεγιστοποιείται για $\alpha = \frac{\text{tr}(A)}{n}$.

Απόδειξη. Το λήμμα αυτό είναι συνέπεια των σχέσεων (3.3.8)—(3.3.13). □

Λήμμα 3.3.3. Για κάθε $n \times n$ πίνακα A ,

$$\|D\|_F^2 \geq |\text{tr}(\tilde{A}^2)| + \frac{|\text{tr}(A)|^2}{n}, \quad (3.3.16)$$

όπου \tilde{A} κεντραρισμένος πίνακας (τον ορίσαμε παραπάνω).

Απόδειξη. Όπως στις (3.3.4)-(3.3.5), ξεκινάμε την απόδειξη εφαρμόζοντας την τριγωνική ανισότητα στις ιδιοτιμές του πίνακα \tilde{A}^2 ,

$$|\text{tr}(\tilde{A}^2)| = |\text{tr}(\tilde{D}^2)| \leq \|\tilde{D}^2\|_F = \|\tilde{D}\|_F^2. \quad (3.3.17)$$

Αν τώρα αντικαταστήσουμε αυτό το κάτω φράγμα για τη $\|\tilde{D}\|_F^2$ στη σχέση

$$\|D\|_F^2 = \|\tilde{D}\|_F^2 + \frac{|\text{tr}(A)|^2}{n}, \quad (3.3.18)$$

παίρνουμε το αποτέλεσμα του λήμματος.

Παρατηρούμε ότι η σχέση (3.3.18) είναι ίδια με την σχέση (3.2.7) για $\alpha=0$. □

3.3.2. Βελτιωμένα άνω φράγματα

Για το άνω φράγμα της $\|D\|_F^2$, έχουμε

$$\|D\|_F^2 = \|A\|_F^2 - \delta_{\|\cdot\|_F}^2(A) \quad (3.3.19)$$

$$\leq \|A\|_F^2 - \left(\|\tilde{A}\|_F^2 - \left(\|\tilde{A}\|_F^4 - \frac{1}{2} \|A^*A - AA^*\|_F^2 \right)^{1/2} \right). \quad (3.3.20)$$

Η εξίσωση (3.2.6), για $\alpha=0$, δείχνει ότι $\|A\|_F^2 - \|\tilde{A}\|_F^2$ απλοποιείται σε $\frac{|\text{tr}(A)|^2}{n}$.

Λήμμα 3.3.4. Για κάθε $n \times n$ πίνακα A ,

$$\|D\|_F^2 \leq \left(\|\tilde{A}\|_F^4 - \frac{1}{2} \|A^*A - AA^*\|_F^2 \right)^{\frac{1}{2}} + \frac{|\text{tr}(A)|^2}{n}, \quad (3.3.21)$$

όπου $\tilde{A} = A - \frac{\text{tr}(A)}{n} I$.

Για το άνω φράγμα του $\delta_{\|\cdot\|_F}^2(A)$, μπορούμε να αντικαταστήσουμε την ανισότητα

$$\|\tilde{D}\|_F^2 \geq |\text{tr}(\tilde{A}^2)|$$

στη σχέση

$$\delta_{\|\cdot\|_F}^2(A) = \delta_{\|\cdot\|_F}^2(\tilde{A}) = \|\tilde{A}\|_F^2 - \|\tilde{D}\|_F^2 \quad (3.3.22)$$

και έτσι έχουμε το ακόλουθο λήμμα.

Λήμμα 3.3.5. Για κάθε $n \times n$ πίνακα A ,

$$\delta_{\|\cdot\|_F}^2(A) \leq \|\tilde{A}\|_F^2 - |\text{tr}(\tilde{A}^2)|, \quad (3.3.23)$$

όπου $\tilde{A} = A - \frac{\text{tr}(A)}{n} I$.

3.4. Βασικά αποτελέσματα

Αρχικά, υπενθυμίζουμε ότι το κάτω φράγμα για την $\delta_{\|\cdot\|_F}^2(A)$ και το άνω φράγμα για το $\|D\|_F^2$ βελτιώνονται μέσω της μιγαδικής μετατόπισης $\alpha = \frac{\text{tr}(A)}{n}$. Επιπλέον, οι πιο αργότερες σχέσεις (3.3.19)-(3.3.20) δείχνουν ότι αν το κάτω φράγμα $\delta_{\|\cdot\|_F}^2(A)$ είναι βέλτιστο, τότε έτσι είναι και το άνω φράγμα $\|D\|_F^2$. Το κάτω φράγμα (3.3.6) είναι βέλτιστο για κάθε μη-κανονικό πίνακα που ικανοποιεί την συνθήκη (3.3.2).

Το βελτιωμένο κάτω φράγμα (3.3.15) με $a = \frac{\text{tr}(A)}{n}$, δεν επηρεάζεται από οποιαδήποτε μιγαδική μετατόπιση. Οπότε, είναι βέλτιστο για

$$A = \gamma(vw^H + rwn^H) - \sigma I, \quad (3.4.1)$$

για κάθε επιλογή του βαθμωτού μεγέθους σ . Παρατηρούμε ότι

$$\alpha = \frac{\text{tr}(\gamma(vw^* + rwn^*) - \sigma I)}{n} = -\sigma, \quad (3.4.2)$$

έτσι ώστε η τιμή $\alpha = \frac{\text{tr}(A)}{n}$ να ακυρώνει την αυθαίρετη μετατόπιση σ . Το βελτιωμένο φράγμα δεν επηρεάζεται επίσης από οποιαδήποτε στροφή του πίνακα. Συνοψίζουμε τα παραπάνω αποτελέσματα στο παρακάτω θεώρημα.

Θεώρημα 3.4.1. Για κάθε $n \times n$ πίνακα A ,

$$\delta_{\|\cdot\|_F}^2(A) \geq \|\tilde{A}\|_F^2 - \left(\|\tilde{A}\|_F^4 - \frac{1}{2} \|A^*A - AA^*\|_F^2 \right)^{1/2}, \quad (3.4.3)$$

και

$$\|D\|_F^2 \leq \left(\|\tilde{A}\|_F^4 - \frac{1}{2} \|A^*A - AA^*\|_F^2 \right)^{1/2} + \frac{|\text{tr}(A)|^2}{n}, \quad (3.4.4)$$

όπου $\tilde{A} = A - \frac{\text{tr}(A)}{n} I$.

Τα φράγματα είναι βέλτιστα (δηλαδή με ισότητα), αν και μόνο αν

$$A = e^{-i\theta} (\gamma(vw^* + rwn^*) - \sigma I), \quad (3.4.5)$$

όπου γ , r και σ είναι μιγαδικοί αριθμοί, $0 \leq \theta < 2\pi$ και v , w είναι ορθοκανονικά διανύσματα.

Θα αποδείξουμε τώρα ότι τα δύο φράγματα (3.3.16) και (3.3.23) είναι βέλτιστα αν και μόνο αν οι ιδιοτιμές του A είναι συγγραμμικές στο μιγαδικό επίπεδο. Όμως, πρώτα πρέπει να ορίσουμε ένα φυσικό μέτρο της μη συγγραμμικότητας των ιδιοτιμών ενός πίνακα. Μία προσέγγιση αυτού του μέτρου είναι να ορίσουμε την “απομάκρυνση από την συγγραμμικότητα” ως

$$\delta_{col}(A) := \sum |d_k|^2, \quad (3.4.6)$$

όπου $|d_k|$ είναι η κάθετη απόσταση της ιδιοτιμής λ_k από την ευθεία της προσαρμογής των ολικών ελαχίστων τετραγώνων (O.E.T.) των ιδιοτιμών του A .

Υπενθυμίζουμε ότι η παραπάνω προσαρμογή ελαχιστοποιεί το άθροισμα των τετραγώνων των κάθετων αποστάσεων από τα σημεία μέχρι την προσαρμοσμένη ευθεία και ότι $\sum |d_k|^2$ είναι το σφάλμα της παραπάνω προσαρμογής. Χρησιμοποιώντας τον ορισμό (3.4.6), βρίσκουμε ότι $\delta_{col}(A) = 0$ αν και μόνο αν ο πίνακας A έχει συγγραμμικές ιδιοτιμές. Παρακάτω δίνουμε ένα χρήσιμο θεώρημα.

Θεώρημα 3.4.2. Έστω οι μιγαδικοί αριθμοί z_k , $k=1, \dots, n$, και έστω $\bar{z} = \frac{1}{n} \sum z_k$ έτσι ώστε

$$\tilde{z}_k = z_k - \bar{z}. \quad (3.4.7)$$

Το σφάλμα για την προσαρμογή των O.E.T. είναι

$$\sum |d_k|^2 = \frac{1}{2} (\sum |\tilde{z}_k|^2 - |\sum \tilde{z}_k^2|), \quad (3.4.8)$$

όπου $|d_k|$ είναι η κάθετη απόσταση από το σημείο z_k στην προσαρμοσμένη ευθεία.

Σε σχέση με τις ιδιοτιμές ενός πίνακα, από την σχέση (3.4.8) προκύπτει ότι

$$\delta_{col}(A) = \frac{1}{2} (\|\tilde{D}\|_F^2 - |tr(\tilde{A}^2)|). \quad (3.4.9)$$

Αν γράψουμε την σχέση (3.4.9) ως

$$\|\tilde{D}\|_F^2 = |tr(\tilde{A}^2)| + 2\delta_{col}(A), \quad (3.4.10)$$

και την αντικαταστήσουμε στα (3.3.18) και (3.3.22), παίρνουμε τη σχέση

$$\|D\|_F^2 = |tr(\tilde{A}^2)| + 2\delta_{col}(A) + \frac{|tr(A)|^2}{n} \quad (3.4.11)$$

και τελικά προκύπτει ότι

$$\delta_{\|\circ\|_F}^2(A) = \|\tilde{A}\|_F^2 - \|\tilde{D}\|_F^2 = \|\tilde{A}\|_F^2 - (|tr(\tilde{A}^2)| + 2\delta_{col}(A)). \quad (3.4.12)$$

Παρατηρούμε ότι τα φράγματα στα Λήμματα 3.3.3. και 3.3.5. είναι ιδιαίτερες περιπτώσεις των (3.4.11) και (3.4.12), αντίστοιχα.

Θεώρημα 3.4.3 (Lee). Για κάθε $n \times n$ πίνακα A ,

$$\|D\|_F^2 \geq |tr(\tilde{A}^2)| + \frac{|tr(A)|^2}{n} \quad (3.4.13)$$

και

$$\delta_{\|\circ\|_F}^2(A) \leq \|\tilde{A}\|_F^2 - |tr(\tilde{A}^2)|, \quad (3.4.14)$$

όπου $\tilde{A} = A - \frac{tr(A)}{n}I$. Τα φράγματα είναι βέλτιστα, με ισότητα να ισχύει αν και μόνο αν οι ιδιοτιμές του A είναι συγγραμμικές στο μιγαδικό επίπεδο.

Επιπλέον,

$$\|D\|_F^2 \approx |tr(\tilde{A}^2)| + \frac{|tr(A)|^2}{n} \quad (3.4.15)$$

και

$$\delta_{\|\circ\|_F}^2(A) \approx \|\tilde{A}\|_F^2 - |tr(\tilde{A}^2)| \quad (3.4.16)$$

αν και μόνο αν, οι ιδιοτιμές του A είναι περίπου συγγραμμικές.

Απόδειξη. Τα φράγματα (3.4.13)-(3.4.14) προκύπτουν από τις σχέσεις (3.4.11)-(3.4.12) απαλείφοντας τον όρο $2\delta_{col}(A)$. Αυτά τα φράγματα είναι βέλτιστα αν και μόνο αν $\delta_{col}(A) = 0$, δηλαδή ισοδύναμα, αν και μόνο αν οι ιδιοτιμές του A είναι συγγραμμικές. Επίσης, αυτά τα φράγματα δίνουν καλές εκτιμήσεις όταν ο όρος $\delta_{col}(A)$ είναι σχετικά μικρός.

□

Παρατηρήσεις

Παρατηρούμε ότι εκτός της πρακτικότητάς τους, τα φράγματα (3.4.13) και (3.4.14) είναι επίσης πολύ χρήσιμα, γιατί μας επιτρέπουν να υπολογίσουμε την $\delta_{\|\circ\|_F}^2(A)$ και $\|D\|_F^2$ για πίνακες με εξαιρετικά ευαίσθητες ιδιοτιμές.

Για παράδειγμα, θεωρούμε τον $n \times n$ πίνακα

$$\widehat{W}_n = U^* W_n U,$$

μηδέν μόνο για κανονικούς πίνακες με συγγραμμικές ιδιοτιμές (π.χ. ερμιτιανοί και αντιερμιτιανοί πίνακες).

Κεφάλαιο 4

Παραδείγματα

Στα παραδείγματα που ακολουθούν θα κάνουμε μία σύγκριση μεταξύ του φράγματος κατά Henrici για την απομάκρυνση από την κανονικότητα και του καινούργιου άνω φράγματος κατά Lee.

Πιο συγκεκριμένα, θα υπολογίσουμε την απομάκρυνση από την κανονικότητα του πίνακα A με βάση τον ορισμό: $\delta_{\|\cdot\|_F}(A) = \sqrt{\|A\|_F^2 - \sum_{i=1}^n |\lambda_i|^2}$, το φράγμα κατά Henrici(2.1.6):

$$\delta_{\|\cdot\|_F}(A) \leq \sqrt{\frac{n^3 - n}{12}} \sqrt{\|A^*A - AA^*\|_F}$$

και το καινούργιο άνω φράγμα (κατά Lee) (3.4.14)

$$\delta_{\|\cdot\|_F}(A) \leq \sqrt{\|\tilde{A}\|_F^2 - |\text{tr}(\tilde{A}^2)|}$$

για την απομάκρυνση από την κανονικότητα για τους πίνακες:

$$A[h] = D + (h^*R), \quad h=0, 1, \dots, 6$$

όπου D είναι ο διαγώνιος πίνακας με στοιχεία τις ιδιοτιμές του πίνακα A και R είναι ο αυστηρά άνω τριγωνικός.

Όλοι οι υπολογισμοί, καθώς και μία γραφική παράσταση, έχουν γίνει με το λογισμικό Octave-4.2.1.(GUI).

Παράδειγμα 4.0.1.

Έστω ο παρακάτω 8x8 μιγαδικός πίνακας

$$A = \begin{bmatrix} -1/2 & -1 & 1 & 3 & 3 & 0 & 1 & 1 \\ 0 & -2i & 0 & 1/4 & 0 & 0 & \sqrt{2} & 1 \\ 0 & 0 & -1 & 7 & 2 & 1/4 & -1/2 & 1 \\ 0 & 0 & 0 & 1/3 & 2 & 0 & -2 & 1 \\ 0 & 0 & 0 & 0 & 2i & -1 & -2 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1 & \sqrt{2} & i \\ 0 & 0 & 0 & 0 & 0 & 0 & 5 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 8 \end{bmatrix}.$$

Τότε, έχουμε τα εξής αποτελέσματα:

Πίνακας	h	$dep_{\ \cdot\ _F}(A)$	Φράγμα Henrici	Καινούργιο φράγμα
1	0	0	0	0
2	1	8.853	25.691	10.601
3	2	19.17	49.816	20.037
4	3	29.144	74.313	29.722
5	4	39.038	98.912	39.472
6	5	48.902	123.553	49.248
7	6	58.749	148.215	59.038

Πίνακας 4.0.1

Όπως παρατηρούμε στον Πίνακα 4.0.1, και για τους 7 πίνακες που έχουμε επιλέξει στο παράδειγμα, το καινούργιο φράγμα (κατά Lee) είναι αυστηρότερο από αυτό του Henrici, αφού οι τιμές που δίνει πάρα πολύ κοντά στην πραγματική τιμή της απομάκρυνσης από την κανονικότητα (3^η στήλη), όπως αυτή έχει οριστεί παραπάνω.

Στα δύο παραδείγματα που ακολουθούν, θα δείξουμε ότι το καινούργιο φράγμα είναι πράγματι βέλτιστο (με ισότητα) όταν οι ιδιοτιμές του πίνακα A είναι οριζόντια ή κάθετα ευθυγραμμισμένες στο μιγαδικό επίπεδο.

Παράδειγμα 4.0.2.

Έστω ο παρακάτω 5x5 μιγαδικός πίνακας A

$$A = \begin{bmatrix} -1 & 0 & 3i & 2 & 1 \\ 0 & 1 & -1 & 3 & 3 \\ 0 & 0 & 4 & 2i & 1 \\ 0 & 0 & 0 & -7 & 0 \\ 0 & 0 & 0 & 0 & 5 \end{bmatrix}.$$

Τότε έχουμε τα παρακάτω αποτελέσματα:

Πίνακας	h	$dep_{\ \cdot\ _F}(A)$	Φράγμα Henrici	Κανούργιο φράγμα
1	0	0	0	0
2	1	6.164	14.245	6.164
3	2	12.329	34.592	12.329
4	3	18.493	36.301	18.493
5	4	24.658	45.515	24.658
6	5	30.822	56.609	30.822
7	6	36.986	67.794	36.986

Πίνακας 4.0.2.

Από τα αποτελέσματα του Πίνακα 4.0.2, παρατηρούμε ξανά ότι το φράγμα κατά Lee είναι σε όλες τις περιπτώσεις αυστηρότερο από το φράγμα κατά Henrici. Επίσης, είναι εύκολο να παρατηρήσουμε ότι το καινούργιο φράγμα είναι βέλτιστο, αφού σε όλες τις περιπτώσεις η τιμή που δίνει είναι ίση με την πραγματική τιμή της απομάκρυνσης από την κανονικότητα του πίνακα A (3^η στήλη).

Αυτό συμβαίνει διότι οι ιδιοτιμές του πίνακα A είναι όλες πραγματικές, δηλαδή, είναι κάθετα ευθυγραμμισμένες στο μιγαδικό επίπεδο (αποτέλεσμα Θεωρήματος 4.3.3.).

Παράδειγμα 4.0.3.

Έστω ο παρακάτω 7x7 πίνακας

$$\Gamma = \begin{bmatrix} 2i & 4 & \sqrt{3} & 0 & -3 & 1 & 1 \\ 0 & 3i & 0 & 1 & 3 & 2 & 2 \\ 0 & 0 & -i & 2 & 1/2 & 5 & 5 \\ 0 & 0 & 0 & 4i & 3 & 4 & 1 \\ 0 & 0 & 0 & 0 & i & 0 & 5 \\ 0 & 0 & 0 & 0 & 0 & -3i & 4 \\ 0 & 0 & 0 & 0 & 0 & 0 & 2i \end{bmatrix}$$

Τότε, έχουμε τα παρακάτω αποτελέσματα:

Πίνακας	h	$dep_{\ \cdot\ _F}(A)$	Φράγμα Henrici	Καινούργιο φράγμα
1	0	0	0	0
2	1	13.01	28.458	13.01
3	2	26.019	55.018	26.019
4	3	39.028	81.966	39.028
5	4	52.038	109.022	52.038
6	5	65.048	136.123	65.048
7	6	78.058	163.247	78.058

Πίνακας 4.0.3

Ξανά, παρατηρούμε ότι σε όλες τις περιπτώσεις το φράγμα κατά Lee είναι βέλτιστο (με ισότητα). Αυτό συμβαίνει διότι όλες οι ιδιοτιμές του πίνακα Γ είναι φανταστικές, δηλαδή είναι ιοριζόντια ευθυγραμμισμένες στο μιγαδικό επίπεδο (αποτέλεσμα Θεωρήματος 4.3.3.).

Παράδειγμα 4.0.4.

Θα εξετάσουμε τώρα και γραφικά πώς συμπεριφέρονται τα δύο παραπάνω φράγματα, αντίστοιχα, για h από 0 μέχρι 6, με βήμα $\delta h=0.1$.

Δηλαδή, θα υπολογίσουμε τις τιμές αυτών των φραγμάτων για 61 πίνακες και θα κάνουμε την γραφική παράσταση αυτών των τιμών συναρτήσεως του h , για το κάθε φράγμα αντίστοιχα. Επιπλέον, στην ίδια γραφική παράσταση, θα βάλουμε και το γραφική παράσταση της $\delta_{\|\cdot\|_F}(A)$ συναρτήσεως του h , με σκοπό να δούμε και γραφικά ποιό από τα δύο φράγματα είναι αυστηρότερο.

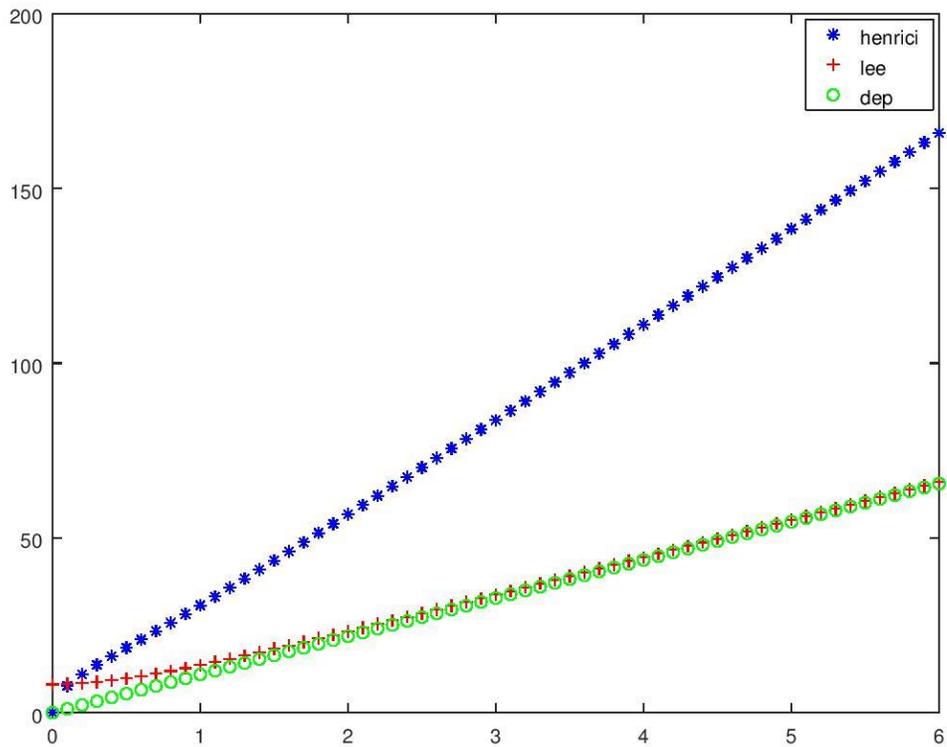
Στο Octave-4.2.1.(GUI), ανοίγουμε ένα new-script και γράφουμε τον παρακάτω κώδικα:

```

clearall
holdoff
clf
%Εισάγουμε τον πίνακα A του παραδείγματος 4.0.1.
a=[-1/2,-1,1,3,3,0,1,1;0,-2i,0,1/4,0,0,sqrt(2),1;0,0,-
1,7,2,1/4,-1/2,1;0,0,0,1/3,2,0,-2,1;0,0,0,0,2i,-1,-
2,0;0,0,0,0,0,-1,sqrt(2),i;0,0,0,0,0,0,5,0;0,0,0,0,0,0,0,8];
%Κατασκευάζουμε τον διαγώνιο πίνακα με στοιχεία τις ιδιοτιμές
του
%A και ύστερα τον αυστηρά άνω τριγωνικό πίνακα.
d=diag(eig(a));
r=triu(a,1);
%Για κάθε τιμή του υπολογίζουμε τονπίνακα α,τον οποίο ορίζουμε
%ως anew
%Μετά υπολογίζουμε το  $A^*A-AA^*$ ,το οποίο το έχουμε ορίσει ως enew
for h=0:0.1:6
rnew=h*r;
anew=d+rnew;
bnew=ctranspose(anew);
cnew=anew*bnew;
dnew=bnew*anew;
enew=dnew-cnew;
%Υπολογίζουμε το φράγμα κατά Henrici
departure_henrici=((42)^(1/4))*(norm(enew,"fro"))^(1/2);
anewcent=anew-((trace(anew)/64)*eye(8)); %κατασκευάσαμε
τον κεντραρισμένο πίνακα
%η Frobenius-νόρμα του κεντραρισμένου πίνακα
normnew=norm(anewcent,"fro");
%το φράγμα κατά Lee στο τετράγωνο
departure_leesq=(normnew**2)-abs(trace(anewcent.*anewcent));
%τοφράγμακατά Lee
departure_lee=(departure_leesq)^(1/2);
%απομάκρυνση από την κανονικότητα
departure=((norm(anew,"fro"))^2)-((norm(d,"fro"))^2))^(1/2);
%τα γραφήματα που ζητάμε
plot(h,departure_henrici,'b*')
hold on
plot(h,departure_lee,'r+')
hold on
plot(h,departure,'go')
legend('henrici','lee','dep')
endfor

```

Αφού αποθήκευσουμε το αρχείο μας, το τρέχουμε και το αποτέλεσμα είναι η παρακάτω γραφική παράσταση:



Εικόνα 1

Στην παραπάνω εικόνα, επαληθεύουμε τα αποτελέσματα του Παραδείγματος 4.0.1., ότι το φράγμα κατά Lee είναι αυστηρότερο από το φράγμα κατά Henrici. Πιο συγκεκριμένα, παρατηρούμε ότι η γραφική παράσταση του φράγματος κατά Lee είναι πολύ κοντά με την γραφική παράσταση της $\delta_{\|\cdot\|_F}(A)$ και μάλιστα, για $h \geq 2$ (περίπου), οι διαφορές μεταξύ τους είναι τόσο μικρές που οι γραφικές αυτές παραστάσεις συμπίπτουν.

Βιβλιογραφία

- 1) P. Henrici, Bounds for iterates, inverses, spectral variation and fields of values of non-normal matrices, *NumerischeMathematik*, 4, 24 - 40, 1962.
- 2) Steven L. Lee, A sharp upper bound for departure from normality, ORNL/TM-12426, August 1993.
- 3) Steven L. Lee, Bounds for departure from normality and the Frobenius norm of matrix eigenvalues, ORNL/TM-12853, December 1994.
- 4) Steven L. Lee, Best available bounds for departure from normality, *SIAM J. Matrix Anal. Appl.*, 17 (4), 984-991, 1996.
- 5) Bauer, F.L. and A.S. Householder, Moments and characteristic roots, *NumerischeMathematik*, 2, 42-53 (1960).
- 6) F.L. Bauer and C.T. Fike, Norms and exclusion theorems, *NumerischeMathematik*, 2, 137-141 (1960).
- 7) A.B. Farnell, Limits for the field of values of a matrix, *Amer. Math. Monthly*, 52, 488-493 (1945).
- 8) G.E. Forsythe, Singularity and near singularity in numerical analysis, *Amer. Math. Monthly*, 65, 229-240 (1958).
- 9) P.D. Lax and R.D. Richtmyer, Survey of the stability of linear finite difference equations, *Comm. Pure Appl. Math.*, 9, 267-293 (1956).
- 10) A. Ostrowski, On the continuity of characteristic roots in their dependence on the matrix elements, Technical Report no. 2, Project 225(37), Applied Mathematics and Statistics Laboratories, Stanford University, 1959.
- 11) A. Ostrowski, *Solution of Equations and Systems of Equations*, New York and London, Academic Press, 1960.
- 12) W.V. Parker, Characteristic roots and fields of values of a matrix, *Bull. Amer. Math. Soc.*, 57, 103-108 (1951).
- 13) B.J. Stone, Best possible ratios of certain matrix norms, *NumerischeMathematik*, 4 (1), 114-116 (1962).
- 14) H. Wielandt, Inclusion theorems for eigenvalues, *Nat. Bur. Standards Appl. Math. Ser.*, 29, 75-78 (1951).
- 15) P.J. Eberlein, On measures of non-normality for matrices, *Amer. Math. Monthly*, 72, 995-996 (1965).
- 16) G. Loizou, Nonnormality and Jordan condition numbers of matrices, *J. Assoc. Comp. Mach.*, 16(4), 580-584, 1969.
- 17) Π. Ψαρράκος, *Θέματα Ανάλυσης Πινάκων*, 2014.