ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΟΙ ΔΙΑΤΑΡΑΧΕΣ ΤΩΝ ΙΔΙΟΤΙΜΩΝ ΠΙΝΑΚΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΑΘΑΝΑΣΙΟΥ ΓΕΩΡΓΙΟΥ
[AM:ge14601]

ΕΠΙΒΛΕΠΩΝ ΚΑΘΗΓΗΤΗΣ : **ΨΑΡΡΑΚΟΣ ΠΑΝΑΓΙΩΤΗΣ**

ΑΘΗΝΑ   ΣΕΠ 2024

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ
ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ

ΟΙ ΔΙΑΤΑΡΑΧΕΣ ΤΩΝ ΙΔΙΟΤΙΜΩΝ ΠΙΝΑΚΩΝ

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

του

ΑΘΑΝΑΣΙΟΥ ΓΕΩΡΓΙΟΥ
[AM:ge14601]

ΤΡΙΜΕΛΗΣ ΕΠΙΤΡΟΠΗ:
ΨΑΡΡΑΚΟΣ ΠΑΝΑΓΙΩΤΗΣ (ΚΑΘΗΓΗΤΗΣ)
ΚΑΝΕΛΛΟΠΟΥΛΟΣ ΒΑΣΙΛΕΙΟΣ (ΚΑΘΗΓΗΤΗΣ)
ΣΤΕΦΑΝΕΑΣ ΠΕΤΡΟΣ (ΑΝΑΠΛΗΡΩΤΗΣ ΚΑΘΗΓΗΤΗΣ)

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΕΡΙΛΗΨΗ……………………………………………………..σελ. 5

ΕΙΣΑΓΩΓΗ……………………………….. ……………….σελ. 7

ΚΕΦΑΛΑΙΟ ΠΡΩΤΟ

ΙΔΙΟΤΙΜΕΣ-ΙΔΙΟΔΙΑΝΥΣΜΑΤΑ…………………………σελ. 9

ΟΙ ΔΙΑΤΑΡΑΧΕΣ ΤΩΝ ΙΔΙΟΤΙΜΩΝ……………………..σελ. 10

Η ΑΝΑΛΥΣΗ ΤΟΥ SCHUR………………………………σελ. 11

ΣΥΝΕΧΕΙΑ: ΘΕΩΡΗΜΑΤΑ OSTROWSKI-ELSNER….σελ. 13

ΘΕΩΡΗΜΑ HALL………………………………………σελ.21

ΘΕΩΡΗΜΑΤΑ BAUER-FIKE-HENRICI…………………σελ.23

ΥΠΟΛΕΙΜΜΑΤΙΚΑ ΟΡΙΑ…………………………………..σελ.30

ΚΕΦΑΛΑΙΟ ΔΕΥΤΕΡΟ

Η ΘΕΩΡΙΑ ΤΟΥ GERSCHGORIN………………………σελ.35

ΔΙΑΓΩΝΙΕΣ ΟΜΟΙΟΤΗΤΕΣ………………………………σελ. 37

ΚΕΦΑΛΑΙΟ ΤΡΙΤΟ

ΚΑΝΟΝΙΚΟΙ ΚΑΙ ΔΙΑΓΩΝΟΠΟΙΗΣΙΜΟΙ ΠΙΝΑΚΕΣ……σελ. 43

ΘΕΩΡΗΜΑ HOFFMAN-WIELANDT……………………..σελ.47

ΘΕΩΡΗΜΑ WEYL………………………………………σελ.51

ΚΕΦΑΛΑΙΟ ΤΕΤΑΡΤΟ

ΕΡΜΙΤΙΑΝΟΙ ΠΙΝΑΚΕΣ…………………………………σελ.55

ΑΔΡΑΝΕΙΑ-ΠΑΡΕΜΒΟΛΗ………………………………σελ.55

ΘΕΩΡΗΜΑ SYLVESTER-JACOBI………………………σελ.56

ΘΕΩΡΗΜΑ CAUCHY……………………………………σελ.56

ΘΕΩΡΗΜΑ WIELANDT…………………………………σελ.57

ΘΕΩΡΗΜΑ MIRSKY……………………………………σελ.62

ΥΠΟΛΕΙΜΜΑΤΙΚΑ ΟΡΙΑ………………………………………σελ.64

ΘΕΩΡΗΜΑ MIRSKY – SCHMIDT……………………………σελ. 69

ΚΕΦΑΛΑΙΟ ΠΕΜΠΤΟ

ΜΗ ΕΡΜΙΤΙΑΝΕΣ ΔΙΑΤΑΡΑΧΕΣ……………………………..σελ. 72

ΘΕΩΡΗΜΑ WILKINSON…………………………………… σελ. 72

ΟΡΙΑ ΟΜΟΙΟΤΗΤΑΣ…………………………………………σελ. 74

ΣΥΖΗΤΗΣΗ……………………………………………………….σελ.80

ΣΥΜΠΕΡΑΣΜΑΤΑ……………………………………………….σελ.81

ΒΙΒΛΙΟΓΡΑΦΙΑ…………………………………………………..σελ.82

ΠΕΡΙΛΗΨΗ

Η παρούσα εργασία αναφέρεται στις διαταραχές των ιδιοτιμών ενός πίνακα και προσφέρει μια εισαγωγή στις επιπτώσεις που έχουν οι διαταραχές αυτές στις ιδιοτιμές και γενικότερα στη συμπεριφορά του αντίστοιχου πίνακα.

Παρουσιάζονται τα βασικά θεωρήματα που αποτελούν τα κύρια εργαλεία ανάλυσης των ιδιοτιμών και των διαταραχών τους. Μέσω αυτών των θεωρημάτων αφενός εξετάζεται η επίδραση των διαταραχών στη σταθερότητα και στη συμπεριφορά των συστημάτων αφετέρου παρέχονται διάφορα μαθηματικά εργαλεία αξιολόγησης του βαθμού παραμόρφωσης και αστάθειας που προκαλούν οι διαταραχές στις ιδιοτιμές.

Η θεωρία των διαταραχών των ιδιοτιμών αναπτύσσει αποτελεσματικές μεθόδους διαχείρισης αυτών των καταστάσεων προκειμένου να εξασφαλιστεί η αξιοπιστία και η αποτελεσματικότητα των αντίστοιχων συστημάτων και εφαρμογών.

Λέξεις κλειδιά: Πίνακες-Διαταραχές ιδιοτιμών - Ιδιοτιμές

SUMMARY

This paper refers to the perturbations of the eigenvalues of a matrix and offers an introduction to the effects these perturbations have on the eigenvalues and more generally on the behavior of the corresponding matrix.

The basic theorems that constitute the main tools for the analysis of eigenvalues and their perturbations are presented. Through these theorems, on the one hand, the effect of disturbances on the stability and behavior of systems is examined, on the other hand, various mathematical tools are provided to evaluate the degree of distortion and instability caused by disturbances in the eigenvalues.

Eigenvalue perturbation theory develops effective methods of handling these situations in order to ensure the reliability and efficiency of the corresponding systems and applications.

Keywords: Tables-Eigenvalue perturbations – Eigenvalues

ΕΙΣΑΓΩΓΗ

Η θεωρία των διαταραχών των ιδιοτιμών ενός πίνακα αποτελεί ένα σημαντικό κεφάλαιο της γραμμικής άλγεβρας και της θεωρίας πινάκων με εφαρμογές στη φυσική, στη μηχανική, στη πληροφορική και στη στατιστική. Η συγκεκριμένη θεωρία εστιάζει στη μελέτη των αλλαγών που παρουσιάζονται στις ιδιοτιμές και κατ΄επέκταση και στα αντίστοιχα ιδιοδιανύσματα ενός πίνακα εξαιτίας μικρών μικρών διαταραχών των στοιχείων αυτού του πίνακα. Οι λόγοι της εμφάνισης τέτοιων διαταραχών είναι πολλοί με κυριότερους τις ατέλειες που παρουσιάζουν τα δεδομένα, τα λάθη στη συλλογή των δεδομένων καθώς και σε μικρές αλλοιώσεις της δομής του πίνακα.

Η παρούσα εργασία επικεντρώνεται στην ανάλυση των διαταραχών των ιδιοτιμών ενός πίνακα και στις συνέπειες που επέρχονται στους διάφορους επιστημονικούς κλάδους. Μελετάται η σημασία της ευαισθησίας των ιδιοτιμών σε σχέση με τις μικρές διαταραχές του πίνακα, καθώς και τους τρόπους με τους οποίους οι διαταραχές αυτές επηρεάζουν την αξιοπιστία και τη σταθερότητα των υπολογισμών και των μετρήσεων που στηρίζονται σε αυτό το πίνακα.

Μέσα από τη συγκεκριμένη εργασία εξετάζονται οι εξής βασικές πτυχές του προβλήματος των διαταραχών των ιδιοτιμών:

1. Η θεωρητική βάση της θεωρίας των διαταραχών των ιδιοτιμών. Συγκεκριμένα αναλύονται οι θεωρίες, τα θεωρήματα και τα μαθηματικά εργαλεία που χρησιμοποιούνται για τη μελέτη των διαταραχών των ιδιοτιμών ενός πίνακα.

2.Οι εφαρμογές σχετικά με πρακτικά προβλήματα. Εξετάζονται οι τρόποι με τους οποίους επηρεάζουν οι διαταραχές των ιδιοτιμών τομείς όπως η επεξεργασία σήματος, η ανάλυση δεδομένων και η μηχανική.

3. Οι αριθμητικές μέθοδοι και οι υπολογιστικές τεχνικές που χρησιμοποιούνται προκειμένου να αντιμετωπιστούν οι διαταραχές των ιδιοτιμών.

Η κατανόηση της συγκεκριμένης θεωρίας καθώς και οι τρόποι αντιμετώπισης των συγκεκριμένων διαταραχών θα βοηθήσει στη εξασφάλιση της αξιοπιστίας και της ακρίβειας των αποτελεσμάτων σε πραγματικές και σημαντικές εφαρμογές.

ΚΕΦΑΛΑΙΟ ΠΡΩΤΟ

ΙΔΙΟΤΙΜΕΣ-ΙΔΙΟΔΙΑΝΥΣΜΑΤΑ

Ιδιοδιάνυσμα ονομάζεται ένα διάνυσμα που δεν αλλάζει κατεύθυνση όταν πολλαπλασιάζεται με τον πίνακα A και η ιδιοτιμή του είναι το ποσό κατά το οποίο συρρικνώνεται ή διαστέλλεται κατά τη διαδικασία.

Το ζεύγος (x,λ) ονομάζεται ιδιοζεύγος του πίνακα A αν x είναι διάφορο του μηδενός και Ax=λx. Το διάνυσμα x ονομάζεται ιδιοδιάνυσμα και το λ είναι η ιδιοτιμή. Το σύνολο όλων των ιδιοτιμών του A γράφεται L(A). Η ισότητα Ax=λx μπορεί να γραφτεί στη μορφή (λI-A)x=0 από την οποία συνεπάγεται ότι το λ είναι μια ιδιοτιμή του A αν και μόνο αν η ποσότητα A-λI είναι τέτοια ώστε

Φ_A (λ) = det (A-λI)=0

Η συνάρτηση Φ_A (λ) είναι ένα πολυώνυμο n βαθμού ως προς λ και ονομάζεται χαρακτηριστικό πολυώνυμο του πίνακα A. Συνεπώς, ένας πίνακας έχει ακριβώς (n) ιδιοτιμές και κάθε διακριτή ιδιοτιμή μετριέται σύμφωνα με την πολλαπλότητα της ως ρίζα της χαρακτηριστικής εξίσωσης

Φ_A (λ) = 0. Μια ιδιοτιμή της οποίας η πολλαπλότητα είναι 1 ονομάζεται απλή ιδιοτιμή. Ο χαρακτηρισμός των ιδιοτιμών ως προς το χαρακτηριστικό πολυώνυμο έχει κάποιες σημαντικές συνέπειες.

Πρώτον, αφού Φ_A (λ) = 0 τότε και μόνο τότε Φ_{A^H} $(\bar{\lambda})$ = 0 και κάθε ιδιοτιμή λ του A αντιστοιχεί σε μια ιδιοτιμή $\bar{\lambda}$ του A^H. Επομένως υπάρχει ένα διάνυσμα y τέτοιο ώστε A^H y = $\bar{\lambda}$ y ή ισοδύναμα $y^H A$ = λy^H Το διάνυσμα y ονομάζεται αριστερό ιδιοδιάνυσ**μ**α του A .

Δεύτερον εάν ο πίνακας A είναι πραγματικός τότε το χαρακτηριστικό πολυώνυμο του είναι πραγματικό και οι πιθανές μιγαδικές ιδιοτιμές θα πρέπει να εμφανίζονται σε μιγαδικά συζυγή ζεύγη.

Τρίτον, αν το A είναι ένας πίνακας της μορφής

$$\begin{pmatrix} A_{11} & A_{12} & \dots & A_{1k} \\ 0 & A_{22} & \dots & A_{2k} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & A_{kk} \end{pmatrix}$$

τότε Φ_A (λ) = Φ_{A11} (λ) + Φ_{A22} (λ) + …+ Φ_{Akk} (λ). Επομένως οι ιδιοτιμές ενός μπλοκ τριγωνικού πίνακα είναι οι ιδιοτιμές των διαγώνιων μπλοκ του. Ειδικότερα οι ιδιοτιμές ενός τριγωνικού πίνακα είναι τα διαγώνια στοιχεία του.

Αν (x , λ) και (y , λ) είναι ιδιοζεύγη του πίνακα A τότε και το (α x + b y , λ) είναι επίσης ιδιοζεύγος υπό την προϋπόθεση ότι (α x + b y) δεν είναι μηδέν. Με αυτό τον τρόπο το σύνολο όλων των ιδιοδιανυσμάτων που αντιστοιχούν σε μια ιδιοτιμή (λ) μαζί με το μηδενικό διάνυσμα σχηματίζουν έναν υποχώρο που ισούται με N(λI-A). Η διάσταση αυτού του υποχώρου, η dim(N(λI-A)) = n – rank(λI-A), ονομάζεται γεωμετρική πολλαπλότητα του λ

Αν (x,λ) είναι ένα ιδιοζεύγος του πίνακα A και ο U είναι μη δυναμικός τότε (U^{-1} A U) U^{-1}x = λ U^{-1} x. Η τελευταία σχέση δείχνει ότι (U^{-1}x, λ) είναι ένα ιδιοζεύγος του U^{-1} A U. Με αυτό τον τρόπο ο μετασχηματισμός ομοιότητας A→U^{-1} A U αντιπροσωπεύει τις ιδιοτιμές του A και μετασχηματίζει τα ιδιοδιανύσματα μέσω του U^{-1}. Αφού rank(λI-A)= rank(λI-U^{-1} A U), οι γεωμετρικές πολλαπλότητες των ιδιοτιμών είναι αμετάβλητες κάτω από μετασχηματισμούς ομοιότητας. Θα ισχύει επομένως Φ_A (λ) = ΦU^{-1} A U(λ).

ΟΙ ΔΙΑΤΑΡΑΧΕΣ ΤΩΝ ΙΔΙΟΤΙΜΩΝ.

Από όλα τα προβλήματα στη θεωρία διαταραχών πινάκων, η διαταραχή των ιδιοτιμών ενός πίνακα παρουσιάζει τις πιο ποικίλες τεχνικές δυσκολίες. Το πρόβλημα αυτό εκφράζεται με τον παρακάτω τρόπο: με δεδομένο έναν πίνακα Α που ανήκει στο C^{nxn} και μιας διαταραχής Ε του Α, πώς σχετίζονται μεταξύ τους τα φάσματα L(A) και L(A+E);

Πρέπει να καταστεί σαφές ότι, ο όρος «σχετικός» εμπεριέχει πολύ περισσότερα από την φυσική έννοια όπως αυτή είναι γνωστή. Το πιο σημαντικό όμως είναι η διαφορετική συμπεριφορά υπό διαταραχή των διαφόρων κατηγοριών πινάκων που έχουν τις ίδιες ιδιοτιμές. Στο σημείο αυτό θα πρέπει να αναφερθούν κάποια σημαντικά θεωρήματα και ορισμοί.

Έστω μ, ν, ρ είναι νόρμες στους χώρους C^{mxn} , C^{nxk} $και$ C^{mxk}. Οι νόρμες μ, ν, ρ θα ονομάζονται συνεπείς αν ρ(ΑΒ)≤ μ(Α).ν(Β) όπου ο Α ανήκει στον C^{mxn} και ο Β ανήκει στον C^{nxk}. Ειδικότερα, μια μήτρα ν στο C^{nxn} είναι συνεπής εάν ισχύει ν (ΑΒ) ≤ ν(Α) ν(Β) για όλους τους πίνακες Α,Β που ανήκουν στον C^{nxn}. Εφόσον οι διανυσματικές νόρμες μπορούν να ταυτιστούν με τις νόρμες μήτρας, ο ορισμός περιλαμβάνει την έννοια της συνέπειας μιας διανυσματικής νόρμας και μιας νόρμας μήτρας.

Για παράδειγμα η νόρμα Frobenius και το διάνυσμα δεύτερης νόρμας είναι συνεπείς γιατί $\|Ax\|_F \leq \|A\|_F \|x\|_2$ δεδομένου ότι $\|x\|_2 = \|x\|_F$.

ΠΑΡΑΔΕΙΓΜΑΤΑ

1. Έστω ένας πίνακας A_{2X2} με ιδιοτιμές λ_1 και λ_2. Εισάγεται ένας μικρός πίνακας διαταραχών Ε και ζητείται να υπολογιστούν οι διαταραγμένες ιδιοτιμές του Α+Ε.

Λύση. Ο γενικός τύπος για την εύρεση των ιδιοτιμών ενός πίνακα 2Χ2 είναι ο εξής: $\lambda_{1,2} = (\,\mathrm{Tr(A)} \pm \sqrt{(\mathrm{Tr(A)}^2 - 4\mathrm{Det(A)})}\,)/2$.

Οι διαταραγμένες ιδιοτιμές του πίνακα Α+Ε θα δίνονται από τους παρακάτω τύπους:

$\lambda_1' \approx \lambda_1$ + Tr(E) - $\sqrt{(\mathrm{Tr(A)^2 - 4Det(A)})}$)* Tr(E) / (2 * (Tr(A)² - 4Det(A))) και

$\lambda_2' \approx \lambda_2$ + Tr(E) + $\sqrt{(\mathrm{Tr(A)^2 - 4Det(A)})}$)* / (2 * (Tr(A)² - 4*Det(A))).

2. Έστω ένας συμμετρικός πίνακας A με ιδιοτιμές λ_1, λ_2 και λ_3 και ένας συμμετρικός πίνακας διαταραχών E. Ζητείται να βρεθούν οι διαταραγμένες ιδιοτιμές του A + E.

Λύση. Είναι γνωστό ότι για τους συμμετρικούς πίνακες, οι ιδιοτιμές τους είναι πραγματικοί αριθμοί. Οι διαταραγμένες ιδιοτιμές του A + E θα δίνονται από τις σχέσεις

$\lambda_1' \approx \lambda_1 + E_{11} + E_{22} + E_{33}$,

$\lambda_2' \approx \lambda_2 + E_{11} + E_{22} + E_{33}$ και

$\lambda_3' \approx \approx \lambda_3 + .E_{11} + E_{22} + E_{33}$.

3. Έστω ένας πίνακας 3x3 A με ιδιοτιμές λ_1, λ_2 και λ_3. Εισάγεται ένας μικρός πίνακας διαταραχών E και ζητείται να βρεθούν τις διαταραγμένες ιδιοτιμές του A + E.

Λύση. Οι σχέσεις διαταραχής για μήτρες 3x3 είναι :

$\lambda_1' \approx \lambda_1 + (E_{11} + E_{22} + E_{33}) + 1/2 * (E_{12} + E_{21} + E_{23} + E_{32} + E_{31} + E_{13})$
$\lambda_2' \approx \lambda_2 + (E_{11} + E_{22} + E_{33}) - 1/2 * (E_{12} + E_{21} + E_{23} + E_{32} + E_{31} + E_{13})$
$\lambda_3' \approx \lambda_3 + (E_{11} + E_{22} + E_{33}) + 1/2 * (E_{12} + E_{21} + E_{23} + E_{32} + E_{31} + E_{13})$

ΘΕΩΡΗΜΑ. Έστω $\|.\|$ μία συνεπής νόρμα μήτρας στον C^{nxn}. Τότε θα υπάρχει μία νόρμα (v) στον C^n που θα είναι συνεπής με τη νόρμα $\|.\|$

Απόδειξη. Επιλέγεται ένα μη μηδενικό διάνυσμα (α) που να ανήκει στον C^n και ορίζουμε v(x) = $\|$x $a^T\|$.

Από τη παραπάνω σχέση προκύπτει

v(Ax) = $\|$Ax $a^T\| \leq \|$A$\|\,\|$ x $a^T\| = \|$A$\|$v(x).

με άμεσο επακόλουθο η (ν) να είναι συνεπής με την $\|.\|$

Οι νόρμες συνεπή πίνακα έχουν μια σημαντική σχέση με τις ιδιοτιμές του πίνακα. Ας ορίσουμε τη φασματική ακτίνα ενός πίνακα A να είναι ο αριθμός ρ(A) = max {|λ| : λ ε L (A)}.

ΘΕΩΡΗΜΑ. Έστω $\|.\|$ είναι μία συνεπής μήτρα. Θα ισχύει για οποιονδήποτε πίνακα A η σχέση ρ(A) ≤ ‖A‖.

Απόδειξη. Με βάση το προηγούμενο θεώρημα, θα υπάρχει ένα διάνυσμα νόρμας (ν) που θα είναι συνεπές με την $\|.\|$.

Έστω (x) ένα ιδιοδιάνυσμα του A που αντιστοιχεί στην ιδιοτιμή (λ). Αυτό σημαίνει ότι Ax=λx. Με χρήση νόρμας θα έχουμε

|λ| ν(x) = ν(λx) = ν(Ax) ≤ ‖**A**‖ν(x).

Θεωρώντας ότι ν(x) > 0, διαιρούμε με το ν(x) και έχουμε |λ| ≤ ‖**A**‖.

ΠΑΡΑΔΕΙΓΜΑ 1. Έστω $\mathrm{A}_{\mathbf{0}}$ =0. Τότε όλες οι ιδιοτιμές του $\mathrm{A}_{\mathbf{0}}$ θα είναι μηδενικές. Θεωρούμε E μια διαταραχή του $\mathrm{A}_{\mathbf{0}}$. Εφόσον

$\mathrm{A}_{\mathbf{0}}$ + E = E, θα ισχύει ότι αφού το (λ) ανήκει στο φάσμα ($\mathrm{A}_{\mathbf{0}}$ + E) τότε |λ| ≤ ‖E‖ για κάθε συνεπή μήτρα $\|.\|$

Έστω ο πίνακας

$$\mathrm{A1} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}$$

Ο πίνακας A1 έχει τις του τις ιδιοτιμές μηδενικές.

Θεωρώντας ε > 0 δημιουργούμε τον πίνακα $\tilde{A}1$ που θα είναι ο

$$\begin{pmatrix} 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ \varepsilon & 0 & 0 & 0 \end{pmatrix}$$

Ο παραπάνω πίνακας έχει ως ιδιοτιμές τις ($\varepsilon^{1/4}$. Wi)

Έτσι, ενώ μια διαταραχή τάξης, ας πούμε, 10^{-8} θα προκαλέσει μια διαταραχή τάξης μόνο 10^{-8} στις ιδιοτιμές του A0, μπορεί να προκαλέσει μια διαταραχή , τόσο μεγάλη, 10^{-2} στις ιδιοτιμές του A1.

Αυτό το παράδειγμα δείχνει ότι μια γενική θεωρία διαταραχών για τις ιδιοτιμές πρέπει να είναι συντηρητική γιατί θα πρέπει να εξηγεί την ελαττωματική συμπεριφορά των ιδιοτιμών πινάκων όπως ο A1.

Θα χρησιμοποιηθούν συμβάσεις για να δηλωθούν οι διαταραχές. Ο A θα υποδηλώνει έναν (σύνθετο) πίνακα τάξης n και $\widetilde{A}$ = A + E θα υποδηλώνει μια διαταραχή του A. Οι ιδιοτιμές του A θα εκφράζονται ως L(A) δηλαδή {λ1,…,λn}, ενώ οι ιδιοτιμές του $\tilde{A}$ θα εκφράζονται ως L(A) = {$\tilde{\lambda}$1,…,$\tilde{\lambda}$n}.
Τα χαρακτηριστικά πολυώνυμα του A και του $\tilde{A}$ θα γράφονται αντίστοιχα φA (λ) και φ$\tilde{A}$ (λ).

Η ΑΝΑΛΥΣΗ ΤΟΥ SCHUR

Ένα βασικό πρόβλημα της θεωρίας πινάκων είναι η αναγωγή ενός πίνακα σε μια απλή μορφή μέσω μετασχηματισμών ομοιότητας. Για παράδειγμα, όπως θα προκύψει και παρακάτω ότι εάν ένας πίνακας A έχει τις κατάλληλες προδιαγραφές τότε θα υπάρχει ένας μη δυναμικός πίνακας X τέτοιος ώστε $X^{-1}\,A\,X = \Lambda =$ diag(λ_1 ,λ_2 ,…,λ_n) όπου οι λ_i είναι οι ιδιοτιμές του A. Τα πλεονεκτήματα μιας τέτοιας διαδικασίας είναι πάρα πολλά.

Έχει αποδειχθεί όμως ότι οι μετασχηματισμοί ομοιότητας μπορούν να δημιουργήσουν προβλήματα. Έστω, για παράδειγμα, ότι ο πίνακας X, που προαναφέρθηκε παραπάνω, περιέχει ένα σφάλμα E. Τότε θα δημιουργηθεί το διαγώνιο σφάλμα Λ + $X^{-1}\,A\,X$. Εάν ο πίνακας X και ο αντίστροφός του είναι μεγάλοι (ή αλλιώς είναι ασθενώς ρυθμισμένοι) τότε το αποτέλεσμα του μετασχηματισμού ομοιότητας θα μεγενθύνει το σφάλμα.

Στη πραγματικότητα οι μοναδιαίοι μετασχηματισμοί ομοιότητας της μορφής A→ $U^H \ A\,U$ είναι ιδιαίτερα επιθυμητοί αφού ούτε ο U ούτε ο ανάστροφός του U^H μπορούν να είναι μεγάλοι. Συγκεκριμένα, από το γεγονός ότι $U^H \ \mathrm{U} = \mathrm{I}$, προκύπτει ότι η j στήλη του U ικανοποιεί τη σχέση $\|\mathbf{u_j}\|_2$ = 1. Οπότε κανένα στοιχείο του U και του ανάστροφου του δεν μπορεί να είναι μεγαλύτερο από μια απόλυτη τιμή.

Γενικά δεν είναι δυνατό πάντα να αναχθεί ένας πίνακας σε διαγώνια μορφή με μοναδιαία ομοιότητα (η σχέση $U^H \ \mathrm{U} = \mathrm{I}$ υποδηλώνει ότι το U έχει περίπου n^2 συνθήκες όπου τα εκτός διαγωνίου στοιχεία του $U^H \ A\,U$ είναι μηδέν). Ωστόσο, το ακόλουθο θεώρημα του Schur δείχνει ότι ένας οποιοσδήποτε πίνακας ανάγεται σε πίνακα τριγωνικής μορφής με μοναδιαία ομοιότητα.

Το θεώρημα αυτό αναφέρει ότι υπάρχει ένας πίνακας U τέτοιος ώστε T= $U^H \ A\,U$ και είναι άνω τριγωνικός. Ο πίνακας U μπορεί να επιλεγεί έτσι ώστε οι ιδιοτιμές του A να εμφανίζονται στη διαγώνιο του T με οποιαδήποτε σειρά.

Υπολογιστικά, το παραπάνω θεώρημα εμφανίζει το στόχο του αλγόριθμου QR , του μοναδικού πιο επιτυχημένου αλγόριθμου για τον υπολογισμό ιδιοσυστημάτων γενικών πινάκων.

<u>Υπενθύμιση.</u> Ένας πίνακας A ονομάζεται κανονικός όταν ισχύει $A^H \ A =$ A A^H.

Υπάρχει ένα θεώρημα που διατυπώνεται ως εξής: αν ένας πίνακας A είναι κανονικός τότε κάθε ανάλυση κατά Schur του A είναι διαγώνια.

Θέτοντας T=$U^H \ A\ U$ και T= diag($\lambda_1, \lambda_2, \ldots, \lambda_n$) καθώς και U=($u_1, u_2, \ldots, u_n$) τότε AU = UΛ ή A$u_j$ =$\lambda_j u_j$. Επομένως τα u_j είναι ιδιοδιανύσματα του A. Εφόσον τα ιδιοδιανύσματα είναι κατά ζεύγη ορθογώνια, έχουμε το ακόλουθο συμπέρασμα: Ένας κανονικός πίνακας τάξης n έχει ένα σύστημα ορθοκανονικών ιδιοδιανυσμάτων στο C^n. Προκύπτει επομένως ότι ένας ενιαίος πίνακας είναι ένας κανονικός πίνακας με ιδιοτιμές στον μοναδιαίο κύκλο ενώ ένας ερμιτιανός πίνακας είναι ένας κανονικός πίνακας με πραγματικές ιδιοτιμές.

Το τελευταίο συμπέρασμα υπονοεί ότι οποιοσδήποτε ερμιτιανός πίνακας μπορεί να γραφτεί με τη μορφή A= UΛU^H όπου ο U είναι ενιαίος και ο Λ είναι διαγώνιος και πραγματικός. Ο συγκεκριμένος τύπος θα ονομάζεται φασματική ανάλυση του A. Μερικές φορές θα γράφεται με τη μορφή A = $\sum_i \lambda_i \ u_i \ u_i^H$ όπου u_i είναι η i στήλη του U. Αυτός ο τύπος επιτρέπει την επέκταση βαθμωτών συναρτήσεων σε ερμιτιανούς πίνακες, δηλαδή φ(A) = $\sum_i \varphi(\lambda_{i)} \ u_i \ u_i^H$. Συγκεκριμένα, αν ο A είναι θετικά ημιορισμένος, οι ιδιοτιμές του είναι μη αρνητικές και η τετραγωνική του ρίζα $A^{1/2} = \sum_i \lambda_i^{1/2} \ u_i \ u_i^H$ είναι καλά ορισμένη.

ΓΕΝΙΚΑ ΘΕΩΡΗΜΑΤΑ ΔΙΑΤΑΡΑΧΩΝ

Συνέχεια: Θεωρήματα Ostrowski – Elsner.

Το πρώτο πράγμα που γίνεται αντιληπτό σχετικά με τις ιδιοτιμές είναι ότι είναι συνεχείς, γεγονός που προκύπτει από το θεώρημα του Rouche. Η πραγματικότητα είναι ότι το χαρακτηριστικό πολυώνυμο ενός πίνακα, όντας το ίδιο πολυώνυμο των στοιχείων του πίνακα είναι μια συνεχής συνάρτηση.

Το θεώρημα του Rouché , δηλώνει ότι για οποιεσδήποτε δύο σύνθετες συναρτήσεις *f* και g ολομορφικές μέσα σε κάποια περιοχή με κλειστό περίγραμμα, εάν $|g(z)| < |f(z)|$ και επί, τότε *τα* f και $f+g$ περιέχουν τον ίδιο αριθμό μηδενικών όπου κάθε μηδενικό μετριέται τόσες φορές όση η πολλαπλότητά του.

ΘΕΩΡΗΜΑ. Έστω λ μια ιδιοτιμή αλγεβρικής πολλαπλότητας (m). Τότε για οποιαδήποτε νόρμα $\|.\|$ και όλα τα πολύ μικρά ε>0 υπάρχει δ>0 τέτοιο ώστε αν $\|E\| < \delta$ ο δίσκος D(λ, ε) που ισούται με $\{\zeta \in C: |\zeta - \lambda| \leq \varepsilon\}$ περιέχει ακριβώς m ιδιοτιμές του A.

Απόδειξη. Έστω ε τόσο μικρό ώστε το D(λ, ε) να περιέχει μόνο την ιδιοτιμή λ του A. Έστω η(ζ) = φ$\tilde{A}$ (ζ) - φA (ζ). Λόγω της συνέχειας του χαρακτηριστικού πολυωνύμου θα ισχύει $\tilde{A} \rightarrow$A και η συνάρτηση η(ζ) θα συγκλίνει στο μηδέν στο συμπαγές σύνολο ∂D. Εφόσον το φA(ζ) είναι μη μηδενικό στο ∂D υπάρχει δ>0 έτσι ώστε το απόλυτο του η(ζ)< | φA(ζ) | στο ∂D. Άρα $\|E\| < \delta$. Από το θεώρημα του Rouche φA και φ$\tilde{A}$ = (φA + η) έχουν τον ίδιο αριθμό μηδενικών στο D.

Το προηγούμενο θεώρημα είναι ένα παράδειγμα θεωρήματος ποιοτικής διαταραχής. Δηλώνει ότι μια διαταραχή πρέπει να είναι μικρή χωρίς να παρέχει όριο στο μέγεθος της διαταραχής.

Στη συνέχεια θα αναπτυχθεί το θεώρημα του Elsner, το οποίο παρέχει σαφή όρια. Πρώτα όμως θα πρέπει να εισαχθεί η αντίστοιχη σημειογραφία για να περιγραφεί ο τρόπος που οι ιδιοτιμές δύο πινάκων τοποθετούνται η μία ως προς την άλλη.

Ορισμός: ‘Εστω ο πίνακας A έχει ιδιοτιμές λ1, λ2,... λn και ο πίνακας $\tilde{A}$ έχει ιδιοτιμές $\tilde{\lambda}$1,…,$\tilde{\lambda}$n. Τότε η φασματική διακύμανση του $\tilde{A}$ ως προς τον A είναι $sv_A(\tilde{A}) = \max_i \min_j |\tilde{\lambda}_i - \lambda_j|$.

Η απόσταση Hausdorff μεταξύ των ιδιοτιμών του Α και του $\tilde{A}$ είναι

hd (Α, $\tilde{A}$) = max { svA($\tilde{A}$) , sv$\tilde{A}$ (A)}.

Η βέλτιστη απόσταση αντιστοίχισης μεταξύ των ιδιοτιμών του Α και του $\tilde{A}$ είναι md (Α, $\tilde{A}$) = min $_\pi$ {max $_i$ $|\tilde{\lambda}_{\pi(i)} - \lambda_i|$}, όπου (π) λαμβάνεται πάνω σε όλες τις μεταθέσεις των {1,2,...,n}.

Η συνάρτηση svA ($\tilde{A}$) δεν είναι μετρική. Μπορεί να είναι μηδέν, ακόμη και όταν οι ιδιοτιμές του Α και του $\tilde{A}$ είναι διαφορετικές (π.χ. όταν n=2 και $\lambda_1=\tilde{\lambda}_1=\tilde{\lambda}_2=0$ ενώ $\lambda_2=1$). Γεωμετρικά η συνάρτηση (sv) έχει την εξής ερμηνεία: αν D_i ={ζ : $|\zeta - \lambda_i| \leq$ svA($\tilde{A}$)}, i=1,2,….,n, τότε L($\tilde{A}$) ⊂ $\cup_{i=1}^{n} Di$.
Τα άλλα λόγια, οι ιδιοτιμές του $\tilde{A}$ βρίσκονται στην ένωση των δίσκων ακτίνας svA ($\tilde{A}$) με επίκεντρο τις ιδιοτιμές του Α.

Η απόσταση Hausdorff οριοθετεί τη φασματική διακύμανση και είναι στην πραγματικότητα μια μέτρηση. Η απόσταση αντιστοίχισης περιορίζει την απόσταση Haussdorff και είναι επίσης μια μέτρηση. Ο ισχυρισμός ότι η απόσταση αντιστοίχισης είναι μικρή είναι από τις καλύτερες ιδιότητες για τις ιδιοτιμές ενός πίνακα και τη διαταραχή του. Σημαίνει ότι μπορούν να ομαδοποιηθούν σε κοντινά ζεύγη.

ΘΕΩΡΗΜΑ ΤΟΥ ELSNER.

Για κάθε πίνακα Α και $\tilde{A}$ ισχύει

hd (Α, $\tilde{A}$) ≤ $(\|\mathrm{A}\|2 + \|\tilde{A}\|2)^{1-1/n} \|\mathrm{E}\|2^{1/n}$

Απόδειξη. Εφόσον η δεξιά πλευρά της ανισότητας είναι συμμετρική ως προς τον Α και τον $\tilde{A}$ τότε αρκεί να αποδειχθεί ότι οριοθετεί το svA ($\tilde{A}$). Έστω ότι το μέγιστο στη σχέση

επιτυγχάνεται για την ιδιοτιμή $\tilde{\lambda}$ του $\tilde{A}$ και έστω $\tilde{x}1,\ldots,\tilde{x}n$

ορθοκανονικά διανύσματα με $\tilde{A}\,x1 = \tilde{\lambda}$ x1. Τότε

Τότε svA($\tilde{A}$) = max $_i$ min $_j$ $|\tilde{\lambda}_i - \lambda_j|$,

svA $(\tilde{A})^n \leq \prod i$ $|\lambda i - \tilde{\lambda}|$ = det (A- $\tilde{\lambda}I) \leq \prod i \,\|(\text{A-}\,\tilde{\lambda}I)\,xi\|2$=

(ανισότητα Hadamard)

=$\|(\text{A} - \tilde{A})\text{x}1\|2 \quad \prod i > 1 \|(\text{A-}\,\tilde{\lambda}I)\,xi\|2 \leq \|\text{E}\|2 \;(\|\text{A}\|2 + \|\tilde{A}\|2)^{n-1}$ Το αποτέλεσμα συνεπάγεται από τη λήψη των ν-στων ριζών της ανωτέρω ανισότητας και από τη συμμετρία του ορίου που προκύπτει.

Όπως προαναφέρθηκε, το πιο επιθυμητό όριο είναι ένα όριο στην αντίστοιχη απόσταση. Σε ορισμένες περιπτώσεις, τα όρια της φασματικής διακύμανσης ή η απόσταση Haussdorff μπορούν να μετατραπούν σε ένα τέτοιο όριο.

Εφόσον η τεχνική, με τις κατάλληλες παραλλαγές, μπορεί να εφαρμοστεί και σε άλλα προβλήματα, θα γίνει ανάπτυξη αυτής της τεχνικής προκειμένου να συμπτυχθούν τα αποτελέσματα. Μπορεί να τεθεί θ svA($\tilde{A}$) ≤ μ $\|\text{E}\|_2^{1/n}$≡ ε όπου μ = $(\max\{2\|\text{A} + \tau\text{E}\|2 : 0 \leq \tau \leq 1\})^{1-1/n}$. Θέτωντας Di ={ζ : $|$ ζ - λi$|$ $\leq \varepsilon$, i=1,2,…,n, η προσαρμογή αυτή στοχεύει στη δεσμευμένη μονοτονία του (τE). Είναι δυνατόν ότι εάν κάποιος (m) από τους δίσκους Di, απομονωθεί από τους άλλους, τότε η ένωσή τους περιέχει ακριβώς (m) ιδιοτιμές του πίνακα $\tilde{A}$.. Για να γίνει κατανοητό το παραπάνω, υποτίθεται χωρίς βλάβη της γενικότητας ότι οι δίσκοι (m) που απομονώθηκαν από τους άλλους είναι οι D1,D2,...Dm. Για 0 ≤τ $\leq$ 1 τίθεται $D_i^{(\tau)}$ ={ζ : $|$ ζ - λi$|$ $\leq$ μ$\|\tau E\|_2^{1/n}$. Λόγω του γεγονότος ότι $\|\text{A}\|_2$ + $\|\widetilde{A_t \;\; au}\|2 \leq$ μ, από το θεώρημα του Elsner θα ισχύει svA $\widetilde{(A_\tau)}$ $\leq$ μ $\|\tau\text{E}\|_2^{\frac{1}{n}}$ = $\tau^{\frac{1}{n}}$ε. Αυτό σημαίνει ότι οι ιδιοτιμές του πίνακα $\tilde{A}$ θα βρίσκονται στην ένωση των δίσκων $D_i^{(\tau)}$. Τώρα $U_{i=1}^{m}$ = $D_i^{(0)}$περιέχει ακριβώς (m) ιδιοτιμές του $\tilde{A}_0$ = A, δηλαδή λ1(A), λ2(A),... λm(A). Δεδομένου ότι $(\tau^{1/n}$ ε)

είναι μια αύξουσα συνάρτηση του (τ) , καθώς το (τ) μεταβάλλεται από το 0 έως το 1, η περιοχή $U_{i=1}^{m}\ D_i^{(\tau)}$ παραμένει ανεξάρτητη από τους άλλους δίσκους. Λόγω προηγούμενου θεωρήματος οι ιδιοτιμές του $\tilde{A}_\tau$ είναι συνεχείς στο (τ) και δεν μπορούν να κάνουν άλμα από τη μια περιοχή στην άλλη. Επομένως το $U_{i=1}^{m}\ D_i^{(1)}$ πρέπει να περιέχει ακριβώς (m) ιδιοτιμές του πίνακα $\tilde{A}_1$ = A. Είναι πλέον εύκολο να αποκτηθεί ένα όριο στο md (A, $\tilde{A}$).
Έστω C_1,C_2,...C_k τα συνδεδεμένα συστατικά του $U_{i=1}^{n}$ Di. Αν C_i είναι η ένωση των m_i των δίσκων D_i τότε περιέχει ακριβώς m_i ιδιοτιμές του A και m_i ιδιοτιμές του $\tilde{A}$. Επιλέγεται επομένως η μετάθεση (π) για να συσχετιστούν οι ιδιοτιμές του A σε κάθε Cl με τις αντίστοιχες ιδιοτιμές του $\tilde{A}$. Εφόσον κάθε ιδιοτιμή του A στο Ci βρίσκεται εντός του ($2m_l$ - 1) δ(A, $\tilde{A}$) οποιουδήποτε σημείου του Ci κάθε ιδιοτιμή του $\tilde{A}$ στο Cl είναι εντός του
($2m_l$ - 1) δ (A, $\tilde{A}$)$\leq (2\mathrm{n}-1)$ δ (A, $\tilde{A}$) *τ*ων αντίστοιχων ιδιοτιμών του A.

Ορισμός.

Έστω T ανήκει στο C^{mxn} όπου m≤n. Αν (π) είναι οποιαδήποτε μετάθεση των {1,2,…,n} και 1 ≤ j_1 ≤ j_2≤ …≤ j_m ≤ n τότε το διάνυσμα ($\tau_{\pi(1)},j_1$,…, $\tau_{\pi(m)},j_m$) θα ονομάζεται διάνυσμα μετάθεσης του T.
Ένα πολύ χρήσιμο θεώρημα είναι το θεώρημα του Hall.

ΘΕΩΡΗΜΑ ΤΟΥ HALL.

Έστω T ανήκει στον C^{mxn} (m≤n). Στη συνέχεια, υπάρχουν οι πίνακες μετάθεσης P και Q τέτοιοι ώστε ο πίνακας PTQ να έχει μία μηδενική υπομήτρα pxq με p+q>n εάν και μόνο εάν κάθε διάνυσμα μετάθεσης του T περιέχει μια μηδενική συνιστώσα.

Απόδειξη.

Θα δείξουμε πρώτα ότι αν το T μπορεί να μετατεθεί στη μορφή

$$T = \begin{array}{c c} & \begin{array}{cc} q & n-q \end{array} \\ \begin{array}{c} p \\ m-p \end{array} & \begin{pmatrix} 0 & T_{12} \\ T_{21} & T_{22} \end{pmatrix} \end{array},$$

όπου p+q>n, τότε οποιοδήποτε διάνυσμα μετάθεσης πρέπει να έχει μηδενικά στοιχεία. Ας υποτεθεί ότι το αντίθετο ($\tau_{i1,j1}$,,....,,$\tau_{im,jm}$) είναι ένα διάνυσμα μετάθεσης με μη μηδενικά στοιχεία. Στη συνέχεια, από τον παραπάνω πίνακα τα στοιχεία πρώτα να είναι διακριτά και να βρίσκονται μεταξύ (p+1) και (m) (συμπεριλαμβανομένου). Από αυτό προκύπτει ότι m-p>=m-(n-q) ή n>=p+q, που είναι άτοπο.

Η απόδειξη του αντίστροφου είναι επαγωγική. Για m=1 ή n=1 στερείται ενδιαφέροντος. Επομένως, θα θεωρηθεί ότι m,n>1 και ότι κάθε διάνυσμα μετάθεσης του T έχει μηδενική συνιστώσα. Χωρίς βλάβη της γενικότητας είναι δυνατόν να υποτεθεί ότι $\tau_{mn} \neq 0$. Τότε κάθε διάνυσμα μετάθεσης του υποπίνακα (m-1)x(n-1) πρέπει να έχει μηδενική συνιστώσα. Για οποιοδήποτε διάνυσμα μετάθεσης που δεν έχει μηδενική συνιστώσα θα μπορούσε να συνδυαστεί με το (τmn) για να δώσει ένα διάνυσμα μετάθεσης του T που δεν έχει μηδενική συνιστώσα (άτοπο).

Από την επαγωγική υπόθεση, το T μπορεί να πάρει την παραπάνω μορφή όπου τώρα θα ισχύει ότι p+q=n. Από αυτό προκύπτει ότι ο T_{12} είναι τετραγωνικός και ο T_{21} έχει τουλάχιστον τόσες στήλες όσες και σειρές. Τώρα τουλάχιστον ένας από τους πίνακες T_{21} ή T_{12} πρέπει να έχει όλα τα διανύσματα μετάθεσης με μηδενικές συνιστώσες. Διότι διαφορετικά θα ήταν δυνατόν να συνδυαστούν διανύσματα μετάθεσης από τα T_{21} και T_{12} που έχουν μη μηδενικά στοιχεία για να σχηματιστεί ένα διάνυσμα μετάθεσης T με μη μηδενικά στοιχεία. Υποθέτοντας ότι όλα τα διανύσματα μετάθεσης του T_{21} έχουν μηδενικές συνιστώσες, από την υπόθεση της επαγωγής, μπορούν να μετατεθούν οι γραμμές και οι στήλες του T έτσι ώστε να έχει τη μορφή.

$$\begin{array}{cc} & \begin{array}{cc} s & n-\cancel{t}\ S \end{array} \\ \begin{array}{c} p \\ r \\ m-p-r \end{array} & \begin{pmatrix} 0 & \hat{T}_{12} \\ 0 & \hat{T}_{22} \\ T_{31} & \hat{T}_{22} \end{pmatrix}, \end{array}$$

όπου r+s>q. Τότε (p+r)+s>p+q=n, οπότε αποδεικνύεται ότι ο παραπάνω πίνακας είναι ο κατάλληλος.

ΘΕΩΡΗΜΑ OSTROWSKI – ELSNER.

Έστω λ ∈ C μία ιδιοτιμή ενός τυχαίου πίνακα A, αλγεβρικής πολλαπλότητας m. Για κάθε νόρμα πινάκων ‖.‖ και για κάθε "αρκετά μικρό"

ε > 0, υπάρχει ένας δ > 0 τέτοιος ώστε για κάθε πίνακα E με μέτρο το E μικρότερο του δ, ο δίσκος Δ(λ, ε) = {z ∈ C : |z − λ| ≤ ε} να περιέχει ακριβώς m ιδιοτιμές του πίνακα $\tilde{A}$ = A + E, λαμβάνοντας υπόψη και τις πολλαπλότητες.

Απόδειξη. Έστω ένας αριθμός ε > 0 αρκετά μικρός ώστε η ιδιοτιμή λ να είναι η μοναδική ιδιοτιμή του πίνακα A εντός του δίσκου Δ(λ, ε). Το ενδιαφέρον εστιάζεται Ε στις (ν) διαταραχές της μορφής $\tilde{A}$ = A + E με ‖E‖ ≤ δ, όπου ο αριθμός δ > 0 είναι "αρκετά μικρός". Έστω τα χαρακτηριστικά πολυώνυμα των A και $\tilde{A}$, φA(z) και φ$\tilde{A}$ (z), καθώς και τη διαφορά τους η(z) = φ$\tilde{A}$ (z) − φA(z). Από τη συνέχεια των συντελεστών του χαρακτηριστικού πολυώνυμου ως προς τα στοιχεία του πίνακα, προκύπτει ότι για $\tilde{A}$ → A, η συνάρτηση η(z) τείνει στο 0. Αφού λοιπόν το πολυώνυμο φA(z) δεν μηδενίζεται πουθενά στο σύνορο ∂Δ(λ, ε), υπάρχει δ > 0 τέτοιος ώστε για κάθε E με ‖E‖ ≤ δ, ισχύει ότι |η(z)| ≤ |φ(z)| για κάθε z ∈ ∂Δ. Άρα από το θεώρημα του Rouche, τα πολυώνυμα φA(z) και $\varphi\tilde{A}$ (z) = φA(z) + η(z) έχουν τον ίδιο αριθμό ριζών εντός του Δ. Το τελικό συμπέρασμα είναι ότι :

md (A, $\tilde{A}$) ≤ (2n-1) $(\|\mathrm{A}\|_2 + \|\widetilde{\mathrm{A}}\|_2)^{1-1/n} \|\mathrm{E}\|_2^{\frac{1}{n}}$

Στην πραγματικότητα, χρησιμοποιήθηκε μόνο το γεγονός ότι το θεώρημα του Elsner δίνει ένα όριο του svA ($\tilde{A}$). Ο Elsner χρησιμοποιώντας και μια εφαρμογή του θεωρήματος του Hall έδειξε ότι ο παράγοντας (2n-1) μπορεί να αντικατασταθεί από 2[n/2]. Συνοψίζοντας ισχύει το επόμενο θεώρημα.

ΘΕΩΡΗΜΑ. Έστω τ ≥ 0. Αν β(τ) είναι ένα μη φθίνον όριο στο sv_{A} (A +τE), τότε md (A, $\tilde{A}$) ≤ (2n-1) β(1). Αν β(τ) είναι ένα μη φθίνον όριο στο hd (A, A +τE) τότε md (A, $\tilde{A}$) ≤ 2[n/2] β(1).

ΤΑ ΘΕΩΡΗΜΑΤΑ ΤΩΝ BAUER-FIKE ΚΑΙ HENRICI.

Έχει ήδη επισημανθεί το γεγονός ότι οποιαδήποτε γενική διαταραχή που σχετίζεται με τις ιδιοτιμές ενός πίνακα θα πρέπει να είναι αρκετά συντηρητική. Στα θεωρήματα των Elsner και Ostrowski - Elsner, η συντηρητικότητα της οποιασδήποτε διαταραχής φαίνεται από το γεγονός ότι το δ(A, $\tilde{A}$) είναι ανάλογο με την νιοστή ρίζα του σφάλματος ν(A, $\tilde{A}$). Το Παράδειγμα 1, ή μάλλον μια μικρή προέκτασή του, δείχνει ότι αυτή η νιοστή ρίζα είναι απαραίτητη. Ωστόσο, στις περισσότερες περιπτώσεις μια τέτοια αναζήτηση δεν είναι ρεαλιστική.

Για να γίνει σαφής ο τρόπος με τον οποίο μπορεί να συμβεί αυτό, ας εξεταστεί εκ νέου το παράδειγμα 1 και ας οριστεί A_n = nA1, όπου το (n) θεωρείται μικρό. Έστω ότι το E δίνεται με τέτοιο τρόπο ώστε $\|\mathrm{E}\|_2$= ε .Αν το (ε) είναι πολύ μικρότερο από το (n), το όριο του Elsner θα είναι γενικά της σωστής τάξης, εκτός εάν το E έχει ειδική δομή. Από την άλλη, αν ε ≥ n τότε

$\|A_n + E\|_2 \leq$ n + ε ≤ 2 ε και καμία ιδιοτιμή του A_nδεν μπορεί να μεταβάλλεται περισσότερο από (2^{ε}). Ο λόγος που η τέταρτη ρίζα του σφάλματος είναι μη ρεαλιστική στη δεύτερη περίπτωση είναι ότι το (A_n +E) μπορεί να θεωρηθεί ως διαταραχή του μηδενικού πίνακα, ο οποίος έχει καλή συμπεριφορά. Στη συνέχεια θα γίνει προσπάθεια να εξαχθεί ένα όριο που θα λαμβάνει υπόψη αυτό το φαινόμενο. Η βασική ιδέα στηρίζεται στο γενικό θεώρημα των Bauer και Fike ενώ απαιτείται και η χρήση ενός ακόμα θεωρήματος.

Στο θεώρημα αυτό θεωρείται μια νόρμα πινάκων $\|.\|$ στον C^{nxn}. Τότε για κάθε Α $\in C^{nxn}$ θα ισχύει ρ(Α) ≤$\|A\|$ **(**όπου ρ(Α) = max {│λ│: λ ∈ σ(Α)}= φασματική ακτίνα του πίνακα Α δηλ. το μέγιστο των απολύτων τιμών των ιδιοτιμών του πίνακα Α).

Απόδειξη. Για κάθε ιδιοτιμή του πίνακα Α, θα ισχύει λ ≤ ρ(Α) καθώς επίσης θα υπάρχει μια τουλάχιστον ιδιοτιμή $\lambda_0 \in$ σ(Α) τέτοια ώστε λ_0 = ρ(Α). Έστω ένα ιδιοδιάνυσμα $x_0 \in C^n$ του Α που αντιστοιχεί στην ιδιοτιμή (λ_0) και ο nxn πίνακας X_0 = [x_0, $x_0, \ldots, x_0$] (πίνακας με όλες τις στήλες του ίσες με x_0). Τότε θα ισχύει Α X_0 = $\lambda_0 X_0$ και $|\lambda_0| \, \|X_0\| = \|\lambda_0 X_0\| = \|A X_0\| \leq \|A\| \, \|X_0\|$**.** Άρα ρ(Α) = $|\lambda_0| \leq \|A\|$**.**

ΘΕΩΡΗΜΑ BAUER – FIKE. Έστω Α ανήκει στο C^{nxn} με φάσμα L(A) και $\|.\|$ μια νόρμα πινάκων. Αν Q είναι ένας αντιστρέψιμος πίνακας nxn τότε για κάθε ιδιοτιμή $\tilde{\lambda}$ του διαταραγμένου πίνακα $\tilde{A}$ = A + E (με το $\tilde{\lambda}$ να μην είναι ιδιοτιμή του Α) θα ισχύει ότι $\|Q^{-1} (A - \tilde{\lambda} I_n)^{-1} Q\|^{-1} \leq \|Q^{-1} E Q\|$**.**

Απόδειξη. Ισχύει ότι

Q^{-1} ($\tilde{A}$ - $\tilde{\lambda} I_n$) Q = Q^{-1} [(A - $\tilde{\lambda} I_n$) + E] Q =

= Q^{-1} (A - $\tilde{\lambda} I_n$) Q { I_n + [Q^{-1} $(A - \tilde{\lambda} I_n)^{-1}$ Q] [Q^{-1} EQ]}.

Ο πίνακας $\widetilde{(A}$ - $\tilde{\lambda} I_{n)}$ είναι μη αντιστρέψιμος οπότε

0 = det ($\tilde{A}$ - $\tilde{\lambda}$ I_n) Q = det (Q^{-1} (A - $\tilde{\lambda}$ I_n) Q [$I_n + Q^{-1}$ (A $-$ $\tilde{\lambda}$ $I_n)^{-1}$ Q Q^{-1} E Q]}

= [det (Q^{-1} (A - $\tilde{\lambda}$ I_n) Q)] [det (I_n $+ Q^{-1}$ (A $-$ $\tilde{\lambda}$ $I_n)^{-1}$ Q Q^{-1} E Q]. Παρατηρείται ότι η πρώτη ορίζουσα του τελευταίου γινομένου δεν είναι μηδενική. Αυτό συνεπάγεται ότι ο πίνακας (Q^{-1} (A $-$ $\tilde{\lambda}$ $I_n)^{-1}$ Q Q^{-1} E Q) έχει ιδιοτιμή το -1. Είναι γνωστό ότι κάθε ιδιοτιμή ενός τυχαίου πίνακα είναι μικρότερη ή ίση από τη νόρμα του πίνακα αυτού (από παραπάνω θεώρημα). Επομένως

1 ≤ $\| Q^{-1}$ (A $-$ $\tilde{\lambda}$ $I_n)^{-1}$ Q Q^{-1} E Q$\|$ ≤ $\| Q^{-1}$ (A $-$ $\tilde{\lambda}$ $I_n)^{-1}$ Q$\| \|Q^{-1}$ E Q$\|$**.** Από τη τελευταία ανισότητα προκύπτει το ζητούμενο. Πρέπει να επισημανθεί ότι αν η αριστερή πλευρά της ανίσωσης θεωρείται μηδέν όταν $\tilde{\lambda}$ ∈ L(A) τότε η ανισότητα ισχύει για όλες τις ιδιοτιμές του πίνακα $\tilde{A}$**.**

Το θεώρημα Bauer-Fike είναι ένα θεμελιώδες αποτέλεσμα της θεωρίας διαταραχής ιδιοτιμών για ερμιτιανούς πίνακες. Παρέχει ένα όριο για το πόσο μπορούν να αλλάξουν οι ιδιοτιμές ενός ερμιτιανού πίνακα όταν υποβάλλεται σε διαταραχή.

Παράδειγμα. Έστω ένας ερμιτιανός πίνακας A_{nxn} με ιδιοτιμές λ_1, λ_2, ..., λ_n και ο πίνακας διαταραχών Ε. Με τη χρήση του θεωρήματος Bauer-Fike να δοθεί ένα άνω όριο της διαταραχής των ιδιοτιμών του A που προκαλείται από τη διαταραχή Ε.

Λύση. Το θεώρημα Bauer-Fike δηλώνει ότι για έναν ερμιτιανό πίνακα A και έναν πίνακα διαταραχών Ε, το όριο της διαταραχής των ιδιοτιμών οριοθετείται από: $\Delta\lambda_i$ ≤ $\| E \|_2$, όπου $\Delta\lambda_i$ είναι η διαταραχή στην i-η ιδιοτιμή, $\| E \|_2$ είναι η φασματική νόρμα του πίνακα διαταραχών Ε. Άρα, για κάθε ιδιοτιμή λ_i, η μέγιστη διαταραχή οριοθετείται από το $\| E \|_2$.

Παράδειγμα. Έστω ένας ερμιτιανός πίνακας A με ιδιοτιμές λ_1, λ_2, λ_3 και ένας πίνακας διαταραχών E με γνωστή φασματική νόρμα $\| E \|_2$. Να υπολογιστεί η μέγιστη δυνατή διαταραχή σε καθεμία από τις παραπάνω ιδιοτιμές χρησιμοποιώντας το θεώρημα Bauer-Fike.

Λύση. Από το θεώρημα Bauer-Fike η μέγιστη διαταραχή για κάθε ιδιοτιμή θα είναι $\Delta\lambda_1 \leq\| E \|_2$, $\Delta\lambda_2 \leq\| E \|_2$ και $\Delta\lambda_3 \leq\| E \|_2$. Η διαταραχή σε κάθε ιδιοτιμή οριοθετείται από τη φασματική νόρμα του πίνακα διαταραχών E και για κάθε ιδιοτιμή λ_i, η μέγιστη διαταραχή είναι $\| E \|_2$. Αυτό σημαίνει ότι οι ιδιοτιμές μπορούν να μεταβληθούν το πολύ κατά $\| E \|_2$ λόγω της υφιστάμενης διαταραχής.

Το θεώρημα Bauer-Fike παρέχει ένα χρήσιμο αναλυτικό εργαλείο για την βαθύτερη κατανόηση των επιδράσεων των διαταραχών στις ιδιοτιμές των ερμιτιανών πινάκων , ιδιαίτερα σε εφαρμογές όπως η κβαντομηχανική.

Η πρώτη εφαρμογή του θεωρήματος Bauer-Fike είναι να αποδειχθεί το θεώρημα της διαταραχής του Henrici. Είναι θεώρημα απόκλισης από την κανονικότητα. Είναι γνωστό ότι εάν ένας πίνακας είναι κανονικός, η μορφή του Schur είναι διαγώνια. Συνεπώς, το μέγεθος των εκτός διαγώνιων όρων στη μορφή Schur μπορεί να χρησιμοποιηθεί για τη μέτρηση της απόκλισης ενός πίνακα από την κανονικότητα.

Έστω το ν είναι μια νόρμα στο C^{nxn} και έστω U το σύνολο του ενιαίου V έτσι ώστε (V^H AV) να είναι άνω τριγωνικός. Για κάθε V∈ U έχουμε V^H AV = Λ_V + R_V όπου το R_V είναι αυστηρά άνω τριγωνικός. Τότε η ν-απόκλιση από την κανονικότητα του A είναι ο αριθμός δ_V(A) =min [ν(R_V)] για V ∈ U. Η απόκλιση από την κανονικότητα δεν είναι εύκολο να υπολογιστεί, αφού η μορφή Schur δεν είναι μοναδική. Ωστόσο, αν το A έχει ιδιοτιμές λ_i, τότε με τη μοναδιαία αναλλοίωτη του κανόνα Frobenius θα ισχύει για οποιαδήποτε μορφή Schur

$\|\mathrm{A}\|_F^2 = \sum |\lambda_i|^2 + \|\mathrm{R}\|_F^2$.

ΘΕΩΡΗΜΑ. Για κάθε πίνακα A με ιδιοτιμές (λ_i) ισχύει ότι

δ_F (A) = $(\|\mathrm{A}\|_{\mathbf{F}}^{\mathbf{2}} - \sum |\lambda_i|^2)^{1/2}$.

ΘΕΩΡΗΜΑ (Henrici). Έστω ότι (ν) είναι μια νόρμα στον C^{nxn} έτσι ώστε ν(C) ≥ $\|\mathrm{C}\|_2$ για όλα τα C ∈ C^{nxn}. Τότε για κάθε ιδιοτιμή $\tilde{\lambda}$ του πίνακα $\tilde{A}$ υπάρχει μια ιδιοτιμή λ του πίνακα A τέτοια ώστε

$(|\tilde{\lambda} - \lambda|/\delta_\nu(\mathrm{A}))^n/[1+(|\tilde{\lambda} - \lambda|/\delta_\nu(\mathrm{A})) + \cdots (|\tilde{\lambda} - \lambda|/\delta_\nu(\mathrm{A}))^{n-1}] \leq$

$(\|\mathrm{E}\|_2/\delta_\nu(\mathrm{A}))$.

Απόδειξη. Έστω $\tilde{\lambda}$ μια ιδιοτιμή του πίνακα $\tilde{A}$ και έστω

U^H A U =Λ + R μια μορφή του Schur αναφορικά με τον πίνακα A. Τότε από το θεώρημα Bauer – Fike θα ισχύει $\|(\Lambda - \tilde{\lambda}\mathrm{I} + \mathrm{R})^{-1}\|_{\mathbf{2}}^{-\mathbf{1}} \leq \|\mathrm{E}\|_2$. Με δεδομένο ότι ο R είναι αυστηρά άνω τριγωνικός πίνακας θα έχουμε

$(\Lambda - \tilde{\lambda}\mathrm{I} + \mathrm{R})^{-1} = \{ \mathrm{I} - (\Lambda - \tilde{\lambda}\mathrm{I})^{-1}\mathrm{R} + \ldots + (-1)^{n-1}[(\Lambda - \tilde{\lambda}\mathrm{I})^{-1}\mathrm{R}]^{n-1}\}(\Lambda - \tilde{\lambda}\mathrm{I})^{-1}$.

Έτσι αν δ = min { λ ∈ L(A) : $|\tilde{\lambda} - \lambda|$} τότε

$\|(\Lambda - \tilde{\lambda}\mathrm{I} + \mathrm{R})^{-1}\|_2 \leq \delta^{-1}\{1 + \delta^{-1}\delta_\nu(\mathrm{A}) + \ldots[\delta^{-1}\delta_\nu(\mathrm{A})]^{n-1}\}$. Ως εκ τούτου

$\|(\Lambda - \tilde{\lambda}\mathrm{I} + \mathrm{R})^{-1}\|_2^{-1} \geq \delta / [1 + \frac{\delta_\nu(\mathrm{A})}{\delta} + \ldots [\frac{\delta_\nu(\mathrm{A})}{\delta}]^{n-1}$.

Με τη χρήση των παραπάνω ανισοτήτων και διαιρώντας με (δ) προκύπτει το ζητούμενο.

Το αξιοσημείωτο με το θεώρημα του Henrici είναι ότι παρέχει μια συνεχή μετάβαση μεταξύ των δύο περιπτώσεων που αναφέρονται στην αρχή αυτής της υποενότητας: δηλαδή, η περίπτωση στην οποία το όριο διαταραχής είναι ανάλογο με την νιοστή ρίζα του σφάλματος και την περίπτωση στην οποία είναι ανάλογο με το ίδιο το σφάλμα. Για να γίνει αυτό πλήρως κατανοητό, έστω ψ(η) = η^n / (1 + η +…+ η^{n-1}), έτσι ώστε η αριστερή πλευρά του τύπου $(\left|\tilde{\lambda} - \lambda\right|/\delta_\nu(\mathrm{A}))^n/[1 + (\left|\tilde{\lambda} - \lambda\right|/\delta_\nu(\mathrm{A})) + \cdots (\left|\tilde{\lambda} - \lambda\right|/\delta_\nu(\mathrm{A}))^{n-1}] \leq (\left\|\mathrm{E}\right\|_2/\delta_\nu(\mathrm{A}))$ να έχει τη μορφή : $\psi[\left|\tilde{\lambda} - \lambda\right|/\delta_\nu(\mathrm{A})]$. Για (η) μικρό τότε το ψ(η) ≈ η^n και ο τύπος παίρνει την ασυμπτωτική μορφή $[\mathrm{sv}_\mathrm{A}(\tilde{A})/\ \delta_\nu(\mathrm{A})] \leq (\frac{\|\mathrm{E}\|_2}{\delta_\nu(\mathrm{A})})^{1/n}$. Στη περίπτωση που το (η) είναι μεγάλο, τότε το ψ(η) ≈ η και ο τύπος παίρνει την ασυμπτωτική μορφή $\mathrm{sv}_\mathrm{A}(\tilde{A}) \leq \left\|\mathrm{E}\right\|_2$. Επομένως εξάγεται το παρακάτω συμπέρασμα: αν $\left\|\mathrm{E}\right\|_2/\delta_\nu(\mathrm{A}) < n^{-1}$, τότε $[\mathrm{sv}_\mathrm{A}(\tilde{A})/\ \delta_\nu(\mathrm{A})] \leq n^{1/n} (\frac{\|\mathrm{E}\|_2}{\delta_\nu(\mathrm{A})})^{1/n}$, ενώ αν $\left\|\mathrm{E}\right\|_2/\delta_\nu(\mathrm{A}) > \eta^{-1}$ τότε

$\mathrm{sv}_\mathrm{A}(\tilde{A}) \leq \left\|\mathrm{E}\right\|_2 + \delta_\nu(\mathrm{A})$.

Η απόδειξη του παραπάνω συμπεράσματος είναι η ακόλουθη:

Αν ψ(η) < 1/n τότε η < 1. Αν $\left\|\mathrm{E}\right\|_2/\delta_\nu(\mathrm{A}) < n^{-1}$,

$n^{-1}\ [\mathrm{sv}_\mathrm{A}(\tilde{A})/\ \delta_\nu(\mathrm{A})]^n \leq \psi\ [\mathrm{sv}_\mathrm{A}(\tilde{A})/\delta_\nu(\mathrm{A})] \leq (\frac{\|\mathrm{E}\|_2}{\delta_\nu(\mathrm{A})})$ που αποδεικνύει το πρώτο σκέλος του συμπεράσματος. Στη συνέχεια, αν ψ(η) > 1 τότε η > 1 και

ψ(η) = η / (1 + η^{-1} + …+ $\eta^{-(n-1)}$ ≥ η (1 - η^{-1}) = η – 1.

Αν $\left\|\mathrm{E}\right\|_2/\delta_\nu(\mathrm{A})$ > 1 θα ισχύει ότι

$[\mathrm{sv}_\mathrm{A}(\tilde{A})/\ \delta_\nu(\mathrm{A})] \leq \psi^{-1}\ (\frac{\|\mathrm{E}\|_2}{\delta_\nu(\mathrm{A})}) \leq (\frac{\|\mathrm{E}\|_2}{\delta_\nu(\mathrm{A})}) + 1$.

Χρησιμοποιώντας το θεώρημα: έστω τ ≥ 0. Αν β(τ) είναι ένα μη φθίνον όριο στο sv_A (A +τE), τότε md (A, $\tilde{A}$) ≤ (2n-1) β(1) και αν β(τ) είναι ένα μη φθίνον όριο στο hd (A, A +τE) τότε md (A, $\tilde{A}$) ≤ 2[n/2] β(1). Από τη μονοτονία της (ψ) προκύπτει το ακόλουθο συμπέρασμα αναφορικά με το όριο του md (A, $\tilde{A}$): έστω ψ που ορίζεται από τη σχέση ψ(η) = η^n / (1 + η +…+ η^{n-1}). Τότε md (A, $\tilde{A}$) ≤ (2n-1) $\delta_\nu(\mathrm{A})\ \psi^{-1}\ (\frac{\|\mathrm{E}\|_2}{\delta_\nu(\mathrm{A})})$.

Έχει ειπωθεί πολλές φορές ότι μια από τις πιο σημαντικές πτυχές του θεωρήματος του Henrici είναι αυτή που μαζί με τα θεωρήματα του Elsner και των Ostrowsi – Elsner αναδεικνύει τη σημαντικότητα της νιοστής ρίζας του σφάλματος στο όριο του. Τα παραδείγματα που δείχνουν ότι η παρουσία του είναι απαραίτητη, εξαρτώνται όλα από τον πίνακα που έχει ένα Jordan μπλοκ ίσο με τη τάξη του. Το παρακάτω θεώρημα δείχνει ότι για έναν πίνακα με τα μικρότερα Jordan blocks η ρίζα είναι μικρότερη.

Για λόγους πληρότητας αναφέρουμε ότι η κανονική μορφή Jordan ενός πίνακα A είναι ένας άλλος πίνακας που έχει στη κύρια διαγώνιο τις ιδιοτιμές του A, στη δευτερεύουσα διαγώνιο τα στοιχεία θα είναι 0 ή 1 και τα υπόλοιπα στοιχεία είναι μηδενικά.

ΘΕΩΡΗΜΑ. Έστω Q^{-1} AQ = J είναι η κανονική μορφή Jordan του πίνακα A και έστω (m) το μέγεθος του μεγαλύτερου Jordan block στον J. Τότε για κάθε ιδιοτιμή $\tilde{\lambda} \in \lambda(\tilde{A})$ υπάρχει μια ιδιοτιμή λ του A έτσι ώστε :

$$(\left|\tilde{\lambda} - \lambda\right|^{m}) / (1 + \left|\tilde{\lambda} - \lambda\right| + \ldots + \left|\tilde{\lambda} - \lambda\right|^{m-1}) \leq \left\|Q^{-1}\,\mathrm{E}\,Q\right\|_2.$$

Απόδειξη. Έστω $\tilde{\lambda}$ μια ιδιοτιμή του J. Η $\tilde{\lambda}$ είναι και ιδιοτιμή του A. Η Jordan μορφή J του A περιέχει Jordan block J_k το καθένα από τα αποία έχει τη μορφή $J_k = \begin{pmatrix} \lambda_\kappa & 1 & 0 \ldots & 0 \\ 0 & \lambda_\kappa & 1 \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 \ldots & \lambda_\kappa \end{pmatrix}$ όπου λ_κ είναι η ιδιοτιμή που αντιστοιχεί στο Jordan block J_k . Η απόσταση $\left|\tilde{\lambda} - \lambda\right|$ όπου λ είναι μια άλλη ιδιοτιμή του A, θα έχει επίδραση στις ιδιοτιμές του J_k. Εφόσον m θεωρηθεί το μέγεθος του μεγαλύτερου Jordan block , τότε η ανισότητα θα αφορά τη διαφορά $\tilde{\lambda}$ και λ. Έστω ο πίνακας E= J - Q^{-1}AQ που είναι μηδενικός λόγω της ισότητας Q^{-1}AQ = J. Άρα $\|\mathrm{E}\|_2$ = 0. Έστω τώρα ότι E είναι ένας μη μηδενικός πίνακας που αναπαριστά τις μικρές διαταραχές. Τότε η ανισότητα

$(\left|\tilde{\lambda} - \lambda\right|^{m}) / (1 + \left|\tilde{\lambda} - \lambda\right| + \ldots + \left|\tilde{\lambda} - \lambda\right|^{m-1}) \leq \left\|Q^{-1}\, \mathrm{E}\, Q\right\|_2$ θα συνδέει την απόσταση των ιδιοτιμών με τη νόρμα του παραγόμενου πίνακα των διαταραχών. Για E=0 η ανισότητα ισχύει ενώ για μικρές διαταραχές η ανισότητα προσεγγίζεται.

ΥΠΟΛΕΙΜΜΑΤΙΚΑ ΟΡΙΑ.

Έστω οι στήλες του X αποτελούν τη βάση για έναν αμετάβλητο υπόχωρο του A. Είναι γνωστό ότι υπάρχει ένας μοναδικός πίνακας M (στη δεδομένη στιγμή είναι $X^T\, AX$) τέτοιος ώστε AX-XM = 0. Ο πίνακας M είναι η αναπαράσταση του A στο R (X) σε σχέση με τη βάση X και επομένως η ιδιοδομή του M είναι ένα υποσύνολο της ιδιοδομής του A. Τώρα έστω ότι οι στήλες του X εκτείνονται σε έναν υπόχωρο που είναι κατά προσέγγιση αμετάβλητος. Για παράδειγμα, το X μπορεί να προέρχεται από έναν αριθμητικό αλγόριθμο για την προσέγγιση αμετάβλητων υπόχωρων. Τότε για οποιοδήποτε M το υπόλοιπο R=AX-XM είναι μη μηδενικό, αν και πιθανώς με μια σωστή επιλογή του M μπορεί να γίνει αρκετά μικρό. Ένα σημαντικό πρόβλημα στη θεωρία των διαταραχών είναι το εξής: δεδομένου κάποιας νόρμας στο R, να προσδιοριστεί πόσο κοντά είναι το R(X) σε έναν αμετάβλητο υποχώρο του A και πώς οι ιδιοτιμές του M σχετίζονται με αυτές του A. Το βασικό εργαλείο για να μελετηθεί το πρόβλημα αυτό έρευνά είναι το ακόλουθο θεώρημα της αντίστροφης διαταραχής.

ΘΕΩΡΗΜΑ. Έστω A $\in C^{nxn}$, X $\in C^{nxp}$ και

M $\in C^{pxp}$. Έστω R=AX-XM. Αν Y^H είναι ένας πίνακας που ικανοποιεί τη σχέση Y^H X = I και $\tilde{A}$ = A – R Y^H τότε $\tilde{A}$ X – X M = 0.

Απόδειξη. Από τη σχέση $\tilde{A}$ = A – R Y^H => $\tilde{A}$ X = AX - R Y^HX

=>$\tilde{A}$ X = AX – RI => $\tilde{A}$ X – X M = AX – AX + XM – XM =>

=> $\tilde{A}$ X – X M = 0.

Από το θεώρημα προκύπτει ότι αν το R είναι μικρό τότε ο R(X) είναι ένας ακριβής αμετάβλητος υποχώρος ενός πίνακα $\tilde{A}$ που είναι κοντά στον A, στην πραγματικότητα εντός $\|R\ Y^H\|$ του A σε οποιαδήποτε νόρμα $\|.\|$. Επιπλέον, το M είναι η αναπαράσταση του $\tilde{A}$ στον R(X) και οι ιδιοτιμές του είναι επομένως και ιδιοτιμές του $\tilde{A}$. Εφόσον είναι γνωστή η νόρμα $\|\mathrm{E}\|$, μπορεί να χρησιμοποιηθεί το κατάλληλο θεώρημα διαταραχών για ιδιοτιμές για να είναι δυνατή η πρόσβαση στην ακρίβεια των ιδιοτιμών του M. Σαν παράδειγμα μπορεί να αναφερθεί η σχέση md (A, $\tilde{A}$) ≤ (2n-1) $\delta_\nu(\mathrm{A})\ \psi^{-1}\ (\frac{\|\mathrm{E}\|_2}{\delta_\nu(\mathrm{A})})$, όπου

ψ(η) = η^n / (1 + η +…+ η^{n-1}). Με αφετηρία την παραπάνω ανίσωση, αν θεωρηθούν $\mu_1, \ldots, \mu_p$ οι ιδιοτιμές του πίνακα M, τότε θα υπάρχουν οι ιδιοτιμές $\lambda_{j1}, \ldots, \lambda_{jp}$ του πίνακα A έτσι ώστε :

$|\mu_i - \lambda_{ji}| \leq$ (2n-1) $\delta_\nu(\mathrm{A})\ \psi^{-1}\ (\frac{\|\mathrm{R}\,Y^H\|_2}{\delta_\nu(\mathrm{A})})$.

Το πρόβλημα της σωστής επιλογής των πινάκων M και Y εξακολουθεί να παραμένει. Γενικά το πρόβλημα αυτό είναι δυσεπίλυτο. Ωστόσο για την περίπτωση των ενιαίων αμετάβλητων προτύπων υπάρχει μια ικανοποιητική λύση.

ΘΕΩΡΗΜΑ. Αν στο παραπάνω θεώρημα υποτεθεί ότι X^H X = I και έστω $\|.\|$ μια μοναδική αμετάβλητη νόρμα, τότε η $\|\mathrm{R}\|$ ελαχιστοποιείται από το M = $X^H\ A$ X. Αλλά και το R Y^H ελαχιστοποιείται από το M = $X^H\ A$ X και Y = X.

Απόδειξη. Έστω (X $X_\perp$) μοναδιαίο και από τη σχέση R=AX-XM θα ισχύει

$\|\mathrm{R}\| = \|(\mathrm{X}\ X_\perp)^{\boldsymbol{H}}\mathrm{R}\| = \left\|\begin{matrix} X^H\ A\,X - M \\ X_\perp\ AX \end{matrix}\right\|$.

Επομένως η $\|R\|$ ελαχιστοποιείται όταν $X^H\, A\, X - M$ = 0. Για να ελαχιστοποιηθεί το $\|R\, Y^H\|$ θα πρέπει να ληφθεί υπόψη ότι η σχέση Y^H X = I υπονοεί ότι Y = X + $X_\perp$ S για ορισμένα S. Τότε

$$\|R\, Y^H\| = \|(X\, X_\perp)^{\boldsymbol{H}} R Y^H\, (X\, X_\perp)\| =$$

$$= \left\| \begin{matrix} (X^H\, A\, X - M) & (X^H\, A\, X - M) S^H \\ (X_\perp A X) & (X_\perp A X S^H) \end{matrix} \right\|.$$ Οπότε η ποσότητα $\|R\, Y^H\|$ ελαχιστοποιείται όταν ($X^H\, A\, X - M$ = 0) και S=0.

Παρατηρήσεις.

Η θεωρία των διαταραχών των ιδιοτιμών παρουσιάζεται με δύο μορφές. Αρχικά θεωρούνται κάποια συγκριτικά αδόμητα σφάλματα και στη συνέχεια γίνεται προσπάθεια να δεσμευτούν οι διαταραχές με κάποιο κανόνα σφαλμάτων. Άλλες προσεγγίσεις επιβάλλουν κάποια δομή στα σφάλματα. Το πρόβλημα είναι τότε να προσδιοριστεί πώς αυτή η δομή επηρεάζει τις διαταραγμένες ιδιοτιμές. Για παράδειγμα πότε είναι αναλυτικές συναρτήσεις της μεταβλητής και ποιες διαδρομές ακολουθούν στο μιγαδικό επίπεδο. Το θεώρημα του Rouche είναι δυνατόν να δώσει εξηγήσεις και να βοηθήσει στην εξαγωγή συμπερασμάτων αναφορικά με την εφαρμογή πιο σύνθετων αναλύσεων στη μελέτη των διαταραγμένων ιδιοτιμών. Σε ορισμένες εφαρμογές μπορεί να έχουμε εκτός από ένα υπόλοιπο για έναν κατά προσέγγιση υποχώρο και ένα υπόλοιπο για τον αντίστοιχο αμετάβλητο υποχώρο.

ΑΣΚΗΣΗ

Η άσκηση που θα αναλυθεί παρακάτω είναι μια εφαρμογή των Kahan, Parlett και Jiang.

Έστω οι πίνακες A ∈ C^{nxn} και X,Y ∈ C^{nxp} έχουν ορθοκανονικές στήλες και υποτίθεται ότι ο πίνακας Y^H X είναι μη κανονικός. Για κάθε M ∈ C^{pxp} έστω

N = $(Y^H X)^{-1}$ M$(Y^H X)$. Τίθεται R= AX-XM και $S^H = Y^H$ A - NY^H. Τότε θα υπάρχει το λιγότερο ένας πίνακας Ε έτσι ώστε: (A+E)X = XM και

Y^H(A+E) = NY^H. Επιπλέον η μικρότερη λύση της νόρμας Frobenius ικανοποιεί τη σχέση $\|E\|_F = (\|R\|_F^2 + \|S\|_F^2 - \|F_{11}\|)^{1/2}$ όπου

$F_{11} = Y^H R = S^H X$. Η μικρότερη λύση της φασματικής νόρμας ικανοποιεί τη σχέση $\|E\|_2 = \max\{\|R\|_2, \|S\|_2\}$. Ως υπόδειξη δίνεται ότι αν έστω (X X⊥) και (Y Y⊥) είναι ορθοκανονικοί και τεθεί ότι $\begin{pmatrix} Y^H \\ Y_\perp^H \end{pmatrix} E(X\ X_\perp) = \begin{pmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{pmatrix}$.

Τότε να δειχθεί μόνο ότι ο F_{22} είναι ελεύθερος.

Λύση. Η έννοια του ελεύθερου πίνακα συνάδει με την έννοια του μη περιορισμένου ή του μη αυθαίρετα δεσμευμένου πίνακα. Το τελευταίο πρακτικά αποδεικνύεται εφόσον ο εν λόγω πίνακας είναι δυνατόν να εκφραστεί ως γραμμικός συνδυασμός άλλων πινάκων.

Από την υπόθεση ισχύει R= AX-XM και $S^H = Y^H$ A - NY^H. Εφόσον πάλι από την υπόθεση ισχύει $Y^H R = S^H X \Rightarrow Y^H$ (AX-XM) = (Y^H A - NY^H)X. Από την εκφώνηση ισχύει ότι $(Y^H X)^{-1} = N^{-1}$ M γιατί ουσιαστικά ο N^{-1} είναι ο αντίστροφος του $(Y^H X)$. Επομένως AX-XM = $(Y^H X)(Y^H AX - NY^H X)(Y^H X)^{-1}$. Με αντικατάσταση των εκφράσεων των N και N^{-1} συνεπάγεται ότι:

AX-XM = $(Y^H X)$ M $((Y^H X)^{-1}$ οπότε

$F_{22} = Y^H X = Y^H AX - NY^H X$. Άρα αποδείχθηκε ότι ο πίνακας F_{22} γράφεται ως γραμμικός συνδυασμός των ορθοκανονικών πινάκων X,Y και των πινάκων A,N. Οπότε αποδείχθηκε και το ζητούμενο.

ΚΕΦΑΛΑΙΟ ΔΕΥΤΕΡΟ

Η ΘΕΩΡΙΑ ΤΟΥ GERSCHGORIN

Τα αποτελέσματα της προηγούμενης ενότητας δεν προτείνουν έναν τρόπο εκχώρησης ενός αριθμού υπό συνθήκη σε μια ιδιοτιμή. Το πρόβλημα είναι ότι οι ιδιοτιμές που σχετίζονται με ένα μη τετριμμένο μπλοκ Jordan δεν είναι διαφοροποιήσιμες συναρτήσεις των στοιχείων του πίνακα.

Ωστόσο, αυτό δεν σημαίνει ότι οι μεμονωμένες ιδιοτιμές δεν μπορούν να συμπεριφέρονται με τοπικά γραμμικό τρόπο και επομένως να εμπεριέχουν κάποια συνθήκη. Στη συνέχεια θα μελετηθεί ένα από τα πιο ισχυρά εργαλεία για την ανίχνευση της ευαισθησίας μιας μεμονωμένης ιδιοτιμής, το θεώρημα Gerschgorin.

Με αυστηρά μαθηματικούς όρους, το θεώρημα του Gerschgorin δεν μπορεί να θεωρηθεί ως θεώρημα διαταραχής. Δηλώνει ότι οι ιδιοτιμές ενός πίνακα βρίσκονται στην ένωση ορισμένων δίσκων στο μιγαδικό επίπεδο. Ωστόσο, όπως θα ειπωθεί και παρακάτω, μπορεί να χρησιμοποιηθεί για τον καθορισμό ορίων διαταραχών μεγάλης ακρίβειας. Υπάρχουν διάφοροι τρόποι για να θεμελιωθεί το θεώρημα του Gerschgorin. Εδώ θα προσεγγιστεί μέσω του θεωρήματος Bauer-Fike.

ΘΕΩΡΗΜΑ (Gerschgorin). Για $\mathrm{A} \in C^{nxn}$ θέτουμε

$a_i = \sum_{j \neq i} | a_{ij}$ και Gi(A) = { z ∈ C : $| z - a_{ii} | \leq a_i$ }. Τότε $\mathrm{L(A)} \subset \cup_{i=1}^{n}$ Gi(A).

Επιπλέον, εάν m από τους Gerschgorin δίσκους Gi(A) απομονωθούν από τους άλλους n-m δίσκους, τότε υπάρχουν ακριβώς m ιδιοτιμές του A στην ένωσή τους.

Απόδειξη.

Έστω D=diag(a_{11}, a_{22}, a_{nn}). Στο θεώρημα Bauer-Fike πραγματοποιούνται οι παρακάτω αντικαταστάσεις:

Q ⇽ I,

A ⇽ D,

$\tilde{A}$ ⇽ A,

$\|.\|$ **⇽** $\|.\|_{\infty}$**.** Τότε από τη πρώτη ανισότητα της

$1 \leq \| [Q^{-1} (A - \tilde{\lambda} I_n)^{-1} Q] [Q^{-1} E Q] \| \leq \| Q^{-1} (A - \tilde{\lambda} I_n)^{-1} Q \| \| Q^{-1} E Q \|$

γίνεται σαφές ότι κάθε ιδιοτιμή του A βρίσκεται στο δίσκο Gerschgorin.

Το επόμενο παράδειγμα καταδεικνύει ότι το θεώρημα του Gerschgorin είναι μια πιο βελτιωμένη έκδοση του θεωρήματος του Elsner ενώ ταυτόγχρονα απεικονίζει και μια ανεπάρκεια όσον αφορά την απλή χρήση του θεωρήματος Gerschgorin.

Παράδειγμα. Έστω ο πίνακας A = $\begin{pmatrix} 1 & 10^{-4} \\ 10^{-4} & 1 \end{pmatrix}$. Όσον αφορά τον A ως διαταραχή του πίνακα diag(1,2) βρίσκουμε από το θεώρημα Elsner ότι η μια ιδιοτιμή πρέπει να βρίσκεται στο διάστημα [1- 0,021, 1+0,021] και η άλλη στο διάστημα [2-0,021, 2+0,021] (στην πραγματικότητα το θεώρημα αποδίδει διαστήματα που είναι λίγο μεγαλύτερα σε μήκος από 0.04).

Από την άλλη πλευρά, σύμφωνα με το θεώρημα του Gerschgorin, καθένα από τα διαστήματα [1- 10^{-4}, 1+10^{-4}] και [2-10^{-4}, 2+10^{-4}] πρέπει να περιέχει μια ιδιοτιμή του A. Έτσι το θεώρημα του Gerschgorin είναι καλύτερο από αυτό του Έλσνερ για περισσότερες από δύο τάξεις μεγέθους. Ωστόσο, οι ιδιοτιμές του Α είναι περίπου (1-10^{-8}) και (2+10^{-8}). Άρα, το θεώρημα του Gerschgorin εξακολουθεί να είναι εκτός κατά τέσσερις τάξεις μεγέθους.

Πρέπει να σημειωθεί ότι στο παραπάνω παράδειγμα έχουν αντικατασταθεί οι δίσκοι στο μιγαδικό επίπεδο με διαστήματα στον άξονα των πραγματικών αριθμών. Ο λόγος για αυτήν την αντικατάσταση γιατί οι δύο δίσκοι, είτε αυτοί που παράγονταιι από το θεώρημα του Elsner είτε αυτοί που παράγονται από το θεώρημα του Gerschgorin, περιέχουν ο καθένας από μία ιδιοτιμή. Εφόσον ο Α είναι πραγματικός, οι μιγαδικές ιδιοτιμές του πρέπει να εμφανίζονται ως μιγαδικά συζυγή ζεύγη και να περιέχονται στην τομή των δίσκων με τον άξονα των πραγματικών αριθμών.

ΔΙΑΓΩΝΙΕΣ ΟΜΟΙΟΤΗΤΕΣ

Το παράδειγμα που παρατέθηκε παραπάνω δείχνει ότι τα όρια που παρέχονται από το θεώρημα του Gerschgorin δεν χρειάζεται να είναι πολύ ευκρινή. Άλλωστε δεν υπάρχει κάποια ξεκάθαρη εξήγηση του λόγου που θα έπρεπε το θεώρημα του Gerschgorin να παρέχει ευκρινή όρια. Ο πίνακας Α έχει ειδική δομή. Είναι σχεδόν διαγώνιος, και αποδεικνύεται ότι αυτή η δομή μπορεί να οδηγήσει στην κατασκευή περισσότερων ευκρινών ορίων.

Η γενική τεχνική που ακολουθείται φαίνεται στην απλούστερή της μορφή με τον πίνακα του προηγούμενου παραδείγματος. Έστω D_a=diag(α,1) και θέτουμε

$A_\alpha = D_a \, A \, D_a{}^{-1} = \begin{pmatrix} 1 & 10^{-4}\,\alpha \\ 10^{-4}\,\alpha^{-1} & 2 \end{pmatrix}$.

Εφόσον ο A_α είναι όμοιος με τον Α, τότε θα έχουν και τις ίδιες ιδιοτιμές. Ωστόσο οι δύο αυτοί πίνακες έχουν διαφορετικούς Gerschgorin δίσκους. Καθώς το α ελαττώνεται, ο πρώτος δίσκος συρρικνώνεται, ενώ ο άλλος επεκτείνεται. Τελικά, ο δεύτερος δίσκος επεκτείνεται κατά πολύ και τείνει να ενσωματώσει τον πρώτο, αλλά μέχρι αυτό να γίνει, ο πρώτος παρέχει ένα ολοένα και πιο βελτιωμένο όριο στην ιδιοτιμή.

Συγκεκριμένα, όσο ισχύει 10^{-4} α + 10^{-4} α^{-1} < 1, οι δύο δίσκοι από το θεώρημα Gerschgorin θα παραμείνουν απομονωμένοι. Είναι εύκολο να διαπιστωθεί ότι αυτό θα ισχύει εφόσον το α είναι λίγο μεγαλύτερο από 10^{-4}, έστω α= (1.01) . 10^{-4}. Αυτό απομονώνει μια ιδιοτιμή του A στο διάστημα [1-(1.01) . 10^{-8}, 1+(1.01). 10^{-8}] που είναι ένα πολύ καλό όριο. Το παραπάνω παράδειγμα είναι σχετικά απλό. Ωστόσο, η τεχνική αυτή της μείωσης ενός δίσκου Gerschgorin θα φανεί στην απόδειξη του παρακάτω θεωρήματος.

ΘΕΩΡΗΜΑ. Έστω λ μια απλή ιδιοτιμή του πίνακα A, με δεξιο και αριστερό ιδιοδιάνυσμα x και y αντίστοιχα. Έστω επίσης μια διαταραχή του πίνακα A, η $\tilde{A}$ = A + E . Τότε θα υπάρχει μια μοναδική ιδιοτιμή $\tilde{\lambda}$ του $\tilde{A}$ τέτοια ώστε

$\tilde{\lambda}$ = λ + $(y^H E_x)$/($(y^H$ x) + O($\|E\|^2$**).**

Απόδειξη Έστω δ>0 η απόσταση μεταξύ του λ και των άλλων ιδιοτιμών του A.Έστω J= Y^H A X η κανονική μορφή Jordan του A, στην οποία οι υπερδιαγώνιοι είναι ίσες με δ/3 ή μηδέν. Να σημειωθεί ότι οι πρώτες στήλες x και y των X και Y είναι τα δεξιά και αριστερά ιδιοδιανύσματα που αντιστοιχούν στο λ και αφού Y^H =X^{-1} έχουμε y^Hy = 1. ο παρονομαστής στη ζητούμενη σχέση είναι μη μηδενικός. Έστω τώρα ο πίνακας $\tilde{J}$ = Y^H (A + E) X. Αυτός ο πίνακας έχει τη μορφή που απεικονίζεται παρακάτω για n=5:

$$\tilde{J} = \begin{pmatrix} \lambda + y^H E x & \varepsilon & \varepsilon & \varepsilon & \varepsilon \\ \varepsilon & \mu & \tau & \varepsilon & \varepsilon \\ \varepsilon & \varepsilon & \mu & \tau & \varepsilon \\ \varepsilon & \varepsilon & \varepsilon & \mu & \tau \\ \varepsilon & \varepsilon & \varepsilon & \varepsilon & \mu \end{pmatrix}$$

Εδώ χρησιμοποιήθηκε το ε για να δείξει γενικά μια ποσότητα που οριοθετείται από $\|Y\|\,\|\mathrm{E}\|\,\|\mathrm{X}\|$, (μ) για μια ιδιοτιμή του A εκτός του (λ + ε) και (τα) για μια ποσότητα που περιορίζεται από (ε+δ/3). Με έναν μετασχηματισμό διαγώνιας ομοιότητας, μπορούμε να αντικαταστήσουμε τον πίνακα $\tilde{J}$ από έναν πίνακα της μορφής

$$\tilde{\boldsymbol{J}}_{\boldsymbol{a}} = \begin{pmatrix} \lambda + y^H E x & \alpha\varepsilon & \alpha\varepsilon & \alpha\varepsilon & \alpha\varepsilon \\ \alpha^{-1}\varepsilon & \mu & \tau & \varepsilon & \varepsilon \\ \alpha^{-1}\varepsilon & \varepsilon & \mu & \tau & \varepsilon \\ \alpha^{-1}\varepsilon & \varepsilon & \varepsilon & \mu & \tau \\ \alpha^{-1}\varepsilon & \varepsilon & \varepsilon & \varepsilon & \mu \end{pmatrix}$$

Τώρα ο πρώτος δίσκος Gerscchgorin του $\tilde{\boldsymbol{J}}_{\boldsymbol{a}}$ έχει κέντρο $(\lambda + y^H E x)$ και ακτίνα που οριοθετείται από (n-1)αε. Οι άλλοι δίσκοι έχουν κέντρο (μ) και ακτίνα οριοθετείται από α^{-1} ε + τ + (n-1)ε. Ως εκ τούτου, αν

$[\alpha^{-1}$ ε + δ/3 + nε + (n-1)αε]< δ,

ο πρώτος δίσκος Gerschgorin θα διαχωρίζεται από τους άλλους.

Τώρα έστω το Ε αρκετά μικρό προκειμένου να ισχύει ότι (2/3)δ - nε > δ/2. Τότε αν [nεα^2 – (δ/2) α+ε]<0 θα ικανοποιείται η ανισότητα

$[\alpha^{-1}$ ε + δ/3 + nε + (n-1)αε]< δ. Αυτή η τελευταία συνθήκη θα ικανοποιηθεί εάν α = 4ε/δ υπό τον όρο ότι απαιτείται να είναι το Ε τόσο μικρό ώστε

(16nε^2 < 1). Σε αυτή την περίπτωση, η ακτίνα του πρώτου δίσκου Gerschgorin οριοθετείται από (4nε^2 / δ) = O(ε^2). Εφόσον αυτός ο δίσκος έχει κέντρο στο $(\lambda + y^H E x)$, η μοναδική ιδιοτιμή που περιέχει είναι η $\tilde{\lambda}$.

Μια άμεση συνέπεια του παραπάνω θεωρήματος είναι ότι οι απλές ιδιοτιμές ενός πίνακα είναι διαφοροποιήσιμες συναρτήσεις των στοιχείων του συγκεκριμένου πίνακα.

ΣΥΜΠΕΡΑΣΜΑ. .Εφόσον ισχύουν οι υποθέσεις του παραπάνω θεωρήματος η ιδιοτιμή λ είναι διαφοροποιήσιμη συνάρτηση του Α. Επιπλέον $\frac{\partial\lambda}{\partial\alpha_{ij}} = \frac{\overline{\eta}_i\,\xi_j}{y^H x}$.

Απόδειξη. Εξ ορισμού μια συνάρτηση f(A) είναι διαφοροποιήσιμη εάν υπάρχει ένας γραμμικός τελεστής $f_A^{'}$ τέτοιος ώστε f(A + E) = f(A) + $f_A^{'}$ (A) + O($\|\mathrm{E}\|$).

Η εξίσωση $\tilde{\lambda}$ = λ + $(y^H Ex)$/($(y^H$ x) + O($\|\mathrm{E}\|^{\mathbf{2}}$**).**

εμφανίζει έναν τέτοιο τελεστή για την ιδιοτιμή λ: δηλαδή

Ε → $\frac{y^H Ex}{y^H \mathrm{x}}$. Για να αποδειχθεί η σχέση $\frac{\partial\lambda}{\partial\alpha_{ij}} = \frac{\overline{\eta}_i\,\xi_j}{y^H x}$. πρέπει να ληφθεί υπόψη ότι $\frac{\partial\lambda}{\partial\alpha_{ij}} = \lim_{\tau\to 0}\frac{\lambda\left(A+\tau\,\mathbf{1}_i\ \ \mathbf{1}_j^T\right)-\lambda(A)}{\tau}$. Αλλά από την εξίσωση

$\tilde{\lambda}$ = λ + $(y^H\ E_x)$/($(y^H$ x) + O($\|\mathrm{E}\|^{\mathbf{2}}$**)** συνεπάγεται ότι

$\lambda\left(A+\tau\,\mathbf{1}_i\ \ \mathbf{1}_j^T\right)-\lambda(A)$ = τ$\frac{y^H\mathbf{1}_i\ \ \mathbf{1}_j^T x}{y^H x}$ + O(τ^2). Από τη τελευταία σχέση προκύπτει το ζητούμενο.

Η απόδειξη του παραπάνω θεωρήματος είναι σχεδόν εξίσου ενδιαφέρουσα με το ίδιο το θεώρημα, αφού δίνει μια εικόνα για τους παράγοντες που καθιστούν σημαντικούς τους όρους ανώτερης τάξης. Συγκεκριμένα υπάρχει η απαίτηση οι όροι που περιλαμβάνουν (ε/δ) να είναι αρκετά μικροί. Ο παρονομαστής (δ) δείχνει ότι εάν μια απλή ιδιοτιμή βρίσκεται κοντά στις γειτονικές τιμές της, το εύρος των διαταραχών για το οποίο η παράγωγος παρέχει επαρκή προσέγγιση θα περιοριστεί. Το μέγεθος του αριθμητή δεν εξαρτάται μόνο από το Ε, αλλά από τα μεγέθη των αναγωγικών μετασχηματισμών Χ και Υ. Αν οι μετασχηματισμοί αυτοί είναι μεγάλοι, αναμένεται ότι οι όροι υψηλότερης τάξης θα γίνουν σημαντικοί. Αξίζει να σημειωθεί ότι μια μικρή τιμή του (δ) θα τείνει να επιδεινώσει το αποτέλεσμα. Η εξίσωση $\tilde{\lambda}$ = λ + $(y^H Ex)$/($(y^H$ x) + O($\|\mathrm{E}\|^{\mathbf{2}}$**)**

μπορεί να γραφτεί $\tilde{\lambda}$ = $(y^H(A+E)x)$/($(y^H$ x) + O($\|\mathrm{E}\|^{\mathbf{2}}$**).**

Η ποσότητα $[(A+E)x]$ ονομάζεται "Πηλίκο Rayleigh" και ένας τρόπος να αποσαφηνιστεί το θεώρημα είναι να ειπωθεί ότι το πηλίκο Rayleigh παρέχει μια προσέγγιση πρώτης τάξης στη διαταραγμένη ιδιοτιμή. Θα γενικευθεί η έννοια του πηλίκου Rayleigh δίνοντας ακριβή όρια για τους όρους δεύτερης τάξης.

Το θεώρημα παρέχει επίσης μια τιμή υπό συνθήκη για μια απλή ιδιοτιμή. Βλέπουμε από τη σχέση $\tilde{\lambda}$ = λ + $(y^H Ex)$/($(y^H$ x) + O($\|\mathrm{E}\|^2$**)** ότι $|\tilde{\lambda} - \lambda| \leq \approx \frac{\|\mathrm{y}\|.\|\mathrm{x}\|}{|y^H \mathrm{x}|}$ $\|\mathrm{E}\|$**,** για οποιοδήποτε σταθερό ζεύγος μήτρας και διανύσματος. Επομένως η ποσότητα
v=$\frac{\|\mathrm{y}\|.\|\mathrm{x}\|}{|y^H \mathrm{x}|}$ είναι μια τιμή υπό συνθήκη για το (λ). Όταν$\|.\|$ = $\|.\|_2$, ο αριθμός (v) είναι η τομή της γωνίας μεταξύ των x και y. Ισούται με μονάδα όταν τα (x) και (y) βρίσκονται στην ίδια κατεύθυνση και αυξάνεται απεριόριστα καθώς τα (x) και (y) τείνουν να γίνουν κάθετα μεταξύ τους. Είναι απαραίτητο να επισημανθούν δύο ενδιαφέροντα σημεία. Το πρώτο είναι ότι αν το (λ) είναι απλή ιδιοτιμή, το αριστερό και το δεξιό ιδιοδιάνυσμά της δεν μπορούν να είναι ορθογώνια, αν και είναι εύκολο να κατασκευαστούν παραδείγματα όπου τα δύο αυτά ιδιοδιανύσματα είναι τόσο κοντά στην ορθογωνικότητα όσο είναι αυτό επιθυμητό. Το δεύτερο είναι ότι το αριστερό και το δεξιό ιδιοδιάνυσμα που αντιστοιχούν σε ένα μη τετριμμένο μπλοκ Jordan πρέπει να είναι ορθογώνια.

Παρατηρήσεις.

Ο Gerschgorin καθιέρωσε το θεώρημά του για να επισημάνει ότι η ένωση των Gerschgorin δίσκων του A είναι το συμπλήρωμα του συνόλου όλων των (ζ) για τα οποία κυριαρχεί διαγώνια το (ζI – A). Στην πραγματικότητα, το θεώρημα ο Gerschgorin δεν ισχύει εκτός αν ο πίνακας είναι μη αναγώγιμος.

Γενικότερα, αν (π) είναι οποιαδήποτε πρόταση τέτοια ώστε η π(Α) να είναι αληθής αν και μόνο εάν το Α δεν είναι πίνακας με τις ιδιοτιμές στη κύρια διαγώνια και τα υπόλοιπα στοιχεία του να είναι μηδενικά, τότε το συμπλήρωμα του συνόλου {ζ:π(ζI -A) είναι αληθές} περιέχει όλες τις ιδιοτιμές του Α. Μεταβάλλοντας τη (π) μπορεί κάποιος να πάρει διαφορετικές περιοχές, μερικές από τις οποίες αντιμετωπίζονται σε ασκήσεις. Ο Gerschgorin επισήμανε ότι η ένωση k απομονωμένων δίσκων περιέχει ακριβώς k ιδιοτιμές.

ΑΣΚΗΣΗ

Έστω Ax=λx και ας υποθέσουμε │ξj│ ⪀ │ξi│(i=1,…,n). Να δειχθεί ότι το λ βρίσκεται στο δίσκο Gerschgorin με κέντρο το α_{jj}.

Απόδειξη. Για να δειχθεί ότι το λ βρίσκεται στο δίσκο Gerschgorin με κέντρο το α_{jj} θα πρέπει να γίνει υπενθύμιση του ορισμού των δίσκων του Gerschgorin κάτω από την υπόθεση ότι │ξj│ ⪀ │ξi│(i=1,…,n). Οι δίσκοι του Gerschgorin ορίζονται με βάση τα διαγώνια στοιχεία του πίνακα Α. Για ένα τετραγωνικό πίνακα A = [α_{ij}] οι δίσκοι του Gerschgorin ορίζονται ως εξής: έστω α_{jj} το διαγώνιο στοιχείο του Α στη j γραμμή και στη j στήλη. Ο δίσκος του Gerschgorin με κέντρο το α_{jj} είναι το σύνολο των μιγαδικών αριθμών z που ικανοποιούν τη σχέση │z - α_{jj}│ ≤ R_j όπου $R_j = \sum i \neq j$ │α_{ij}│.

Τώρα, ας εξετάσουμε την εξίσωση του ιδιοδιανύσματος Ax = λx. Από τη σχέση αυτή συνεπάγεται ότι $\sum_{i=1}^{n} a_{ij}\ \xi_i$ = λξ_j.

Έστω τώρα ότι │$a_{ij}\ \xi_i$│ = │λξ_j│. Από την υπόθεση ισχύει ότι │ξj│ ⪀ │ξi│(i=1,…,n). Επομένως συνεπάγεται η ανίσωση │a_{ij}││ξ_i│≤ │λ││ξ_j│. Με χρήση αθροίσματος και στα δύο μέλη συνεπάγεται $\sum_{i=1}^{n}$ │a_{ij}│ │ξ_i│ ≤ $\sum_{i=1}^{n}$ │λ│ │ξ_j│=>│λ│⪀$\sum i \neq j$│α_{ij}│.j │α_{ij}│. Συγκρίνοντας με τον ορισμό του R_j στο δίσκο του Gerschgorin γίνεται άμεσα αντιληπτό ότι θα πρέπει $|\lambda| \geqq R_j$ που σημαίνει ότι $|\lambda - \alpha_{jj}| \leq R_j$. Η τελευταία ανίσωση αποδεικνύει το ζητούμενο.

ΑΣΚΗΣΗ

Ένας τετραγωνικός πίνακας Α είναι αναγώγιμος εάν υπάρχει ένας πίνακας μετάθεσης Ρ τέτοιος ώστε $P^T AP = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$

όπου τα A_{11}, A_{22} είναι τετραγωνικοί πίνακες. Δείξτε ότι ένας μη αναγώγιμος διαγώνια κυρίαρχος πίνακας για τον οποίο τουλάχιστον μία από τις διαγώνιες είναι αυστηρά κυρίαρχη είναι μη μοναδιαίος.

Απόδειξη

Ένας τετραγωνικός πίνακας ονομάζεται «αυστηρά κυρίαρχος» ως προς μια διαγώνιό του όταν τα στοιχεία της διαγωνίου αυτής είναι μεγαλύτερα από όλα τα στοιχεία εκτός της διαγωνίου αυτής.

Έστω τώρα $A = \begin{pmatrix} A_{11} & A_{12} \\ 0 & A_{22} \end{pmatrix}$ μη αναγώγιμος διαγώνια κυρίαρχος πίνακας, όπου A_{11}, A_{22} είναι τετραγωνικοί πίνακες, A_{11} είναι μη αναγώγιμος και τουλάχιστον μία από τις διαγώνιες του είναι αυστηρά κυρίαρχη. Θα δειχθεί ότι ο A_{11} είναι μη μοναδιαίος.

Έστω (λ) μια ιδιοτιμή του A_{11} και (v) το αντίστοιχο ιδιοδιάνυσμα. Εξ ορισμού τότε A_{11} v = λ v. Αν υποτεθεί ότι λ=1 τότε και ο Α θα έχει σαν ιδιοτιμή το (λ) λόγω του γεγονότος ότι οι πίνακες Α και A_{11} έχουν την ίδια δομή. Αυτό όμως είναι άτοπο γιατί αντιτίθεται στην υπόθεση ότι ο Α είναι μη αναγώγιμος. Άρα ο A_{11} είναι μη μοναδιαίος.

ΑΣΚΗΣΗ

Έστω x,y το αριστερό και δεξιό αντίστοιχα ιδιοδιάνυσμα που αντιστοιχεί στην απλή ιδιοτιμή λ. Έστω θ = γωνία (x,y). Να δειχθεί ότι υπάρχει ένας πίνακας Ε που ικανοποιεί τη σχέση $\frac{\|\mathbf{E}\|_2}{\|\mathbf{A}\|_2} \leq \cot\theta$ όπου λ είναι μια πολλαπλή ιδιοτιμή του πίνακα. $A + E$.

Λύση. Ο πίνακας Ε, έστω ότι, ικανοποιεί τη συνθήκη $\|A + E\|_2$ = λ όπου λ μια πολλαπλή ιδιοτιμή του πίνακα $A + E$. Έστω επίσης ότι το διάνυσμα z =x+y. Τότε ο πίνακας Α+Ε επιδρά στο διάνυσμα z ως εξής:

$(A + E)$ z = $(A + E)$(x + Ey) = A x +λ y + E x +λ E y . Το E(x +λy) << λ(x+y) οπότε το διάνυσμα z έχει τη κατεύθυνση του x+y, δηλαδή είναι πολύ κοντά στο χώρο που παράγουν τα ιδιοδιανύσματα x και y. Από αυτό βγαίνει το συμπέρασμα ότι ο πίνακας $A + E$ δεν έχει σημαντική επίδραση στη κατεύθυνση που παίρνει το z λόγω των x ,y. Τότε συνάγεται ότι ο πίνακας Ε μπορεί να έχει μικρή δεύτερη ορθογώνια νόρμα σε σχέση με τον πίνακα Α, δηλαδή $\frac{\|\mathbf{E}\|_2}{\|\mathbf{A}\|_2} \leq \cot\theta$.

ΚΕΦΑΛΑΙΟ ΤΡΙΤΟ

ΚΑΝΟΝΙΚΟΙ ΚΑΙ ΔΙΑΓΩΝΟΠΟΙΗΣΙΜΟΙ ΠΙΝΑΚΕΣ

Κανονικός πίνακας ονομάζεται κάθε πίνακας που ικανοποιεί τη σχέση A^HA=AA^H. Από αυτό προκύπτει ότι οι ερμιτιανοί πίνακες και οι μοναδιαίοι ή ταυτοτικοί πίνακες είναι κανονικοί. Με δεδομένη τη μεγάλη σημασία αυτής της κατηγορίας των πινάκων, είναι φυσικό να αναζητείται μια ειδική θεωρία διαταραχών για τις ιδιοτιμές τους. Η κατάσταση περιπλέκεται περισσότερο από το γεγονός ότι οι κανονικοί πίνακες, αντίθετα με τους Ερμιτιανούς πίνακες, υπάρχει η πιθανότητα να έχουν μιγαδικές ιδιοτιμές που δεν μπορούν να ταξινομηθούν κατά μέγεθος. Ωστόσο, οι κανονικοί πίνακες έχουν αρκετά καλή δομή προκειμένου να επιτρέψουν την απόδειξη του εξαιρετικά χρήσιμου θεωρήματος Hoffman-Wielandt. Δεδομένου ότι οποιοσδήποτε κανονικός πίνακας μπορεί να διαγωνοποιηθεί με μοναδιαίο μετασχηματισμό, οι κανονικοί πίνακες είναι ειδικές περιπτώσεις διαγωνοποιήσιμων πινάκων. Δηλαδή, πίνακες που μπορούν να διαγωνοποιηθούν με μετασχηματισμό ομοιότητας (αυτοί οι πίνακες μερικές φορές ονομάζονται και κανονικοποιήσιμοι).

ΤΟ ΘΕΩΡΗΜΑ HOFFMAN-WIELANDT.

Είναι σχετικά εύκολο να ληφθούν όρια στις φασματικές μεταβλητές $sv_A(\widetilde{A})$ ενός πίνακα $\tilde{A}$ σε σχέση με τον πίνακα A. Υπάρχει η δυνατότητα ένα τέτοιο όριο να μετατραπεί σε όριο εντός του md(A, $\tilde{A}$) με τίμημα όμως ένα συντελεστή (2n-1) στο όριο αυτό. Το θεώρημα Hoffman-Wielandt βασίζεται στο ότι αν οι πίνακες A και $\tilde{A}$ είναι κανονικοί, δεν είναι απαραίτητο να υπάρξει το προαναφερθέν τίμημα για να παρθεί ένα όριο στο

md_2 (A, $\tilde{A}$) = min $_\pi$ $\sum_i \sqrt{|\tilde{\lambda}_{\pi(i)} - \lambda_{(i)}|^2}$

Όταν το π παίρνει τιμές στο σύνολο των ακεραίων 1,2,…n. Ο δείκτης 2 αναφέρεται στη νόρμα-2. Σε αυτό το συμβολισμό η πιο συνηθισμένη απόσταση αντιστοίχισης είναι md∞.

ΘΕΩΡΗΜΑ (Hoffman-Wielandt)

Έστω A και $\tilde{A}$ κανονικοί πίνακες. . Τότε

md_2 (A, $\tilde{A}$) ≤ $\|\tilde{A} - \mathrm{A}\|_F$ όπου η md_2 (A, $\tilde{A}$) ορίζεται από τη σχέση

md_2 (A, $\tilde{A}$) =min $_\pi$ $\sum_i \sqrt{|\tilde{\lambda}_{\pi(i)} - \lambda_{(i)}|^2}$.

Απόδειξη. Εφόσον η $\|.\|_F$ είναι μοναδιαία αμετάβλητη, μπορεί να υποτεθεί ότι A=Λ diag(λ1,...,λn). Έστω τώρα ότι

$\widetilde{\mathrm{A}} = \mathrm{W}\widetilde{\Lambda}W^H$, όπου ο πίνακας W είναι μοναδιαίος και $\tilde{\Lambda}$ =diag(λ1,...,λn).

Το θεώρημα μπορεί να αποδειχθεί αν δειχθεί ότι $\|\Lambda - \mathrm{V}\tilde{\Lambda}\mathrm{V}^{\mathbf{H}}\|_F$ θεωρηθεί ως συνάρτηση του ενιαίου πίνακα, V που ελαχιστοποιείται όταν ο $\mathrm{V} = \mathrm{P}_\pi$ που είναι ένας πίνακας

μετάθεσης που αντιστοιχεί σε κάποια μετάθεση π. Σε αυτή την περίπτωση

$$\begin{aligned}\mathrm{md}_2^2(A,\tilde{A}) = \mathrm{md}_2^2(\Lambda,\tilde{\Lambda}) &\leq \sum_i |\lambda_i - \tilde{\lambda}_{\pi(i)}|^2 \\ &\leq \|\Lambda - W\tilde{\Lambda}W^{\mathrm{H}}\|_{\mathrm{F}}^2 = \|A - \tilde{A}\|_{\mathrm{F}}^2.\end{aligned}$$

Δηλώνοντας τα στοιχεία του V με vij, εμφανίζεται ο άμεσος υπολογισμός $\|\Lambda - \mathrm{V}\ \tilde{\Lambda}\ \ V^H\|_F^2 = \sum_i |\lambda_i|^2 + \sum_i |\tilde{\lambda}_i|^2$ - φ(V)

όπου φ(V) = $\sum_{i,j}(\lambda_i\ \bar{\tilde{\lambda}}_j + \bar{\lambda}_i\ \tilde{\lambda}_j)\ |v_{ij}|^2$. Με αυτό τον τρόπο το συγκεκριμένο πρόβλημα μετατοπίζεται στο να δειχθεί ότι το φ(V) μεγιστοποιείται όταν το V είναι κάποιος πίνακας μετάθεσης. Εφόσον ο V είναι ενιαίος, ο πίνακας του οποίου τα στοιχεία είναι $|v_{ij}|^2$ είναι διπλά στοχαστικός. Για κάθε διπλά στοχαστικό πίνακα ορίζεται ψ(S) = $\sum_{i,j}(\lambda_i\ \bar{\tilde{\lambda}}_j + \bar{\lambda}_i\ \tilde{\lambda}_j)\ \sigma_{ij}$.

Είναι σαφές ότι max φ(V) <= max ψ(S), αφού δεν έχει κάθε διπλά στοχαστικός πίνακας στοιχεία της μορφής $|v_{ij}|^2$, όπου ο V=(v_{ij}) είναι ενιαίος πίνακας. Επομένως, αν αποδειχθεί ότι το ψ μεγιστοποιείται όταν το S είναι ένας πίνακας μετάθεσης P_π, τότε εφόσον ο P_π είναι μοναδιαίος, μεγιστοποιείται και το φ. Με βάση το θεώρημα του Birkhoff, μπορεί να ισχυριστεί κάποιος ότι οποιοσδήποτε διπλά στοχαστικός πίνακας S μπορεί να γραφτεί ως ένας κυρτός συνδυασμός των πινάκων μεταθέσεωνP_π: δηλαδή S=$\sum_\pi \alpha_\pi \ P_\pi$όπου τα α_π είναι μη αρνητικά και αθροίζονται στο 1. Εφόσον το ψ είναι γραμμικό στο S τότε θα ισχύει ψ(S)= $\sum_\pi \alpha_\pi$ψ(P_π). Συνεπάγεται ότι αν π είναι η μετάθεση για την οποία το ψ(P_π) είναι μέγιστο, τότε ψ(S)<=ψ(P_π).). Ως εκ τούτου, το φ(P_π).) είναι επίσης μέγιστο και το π είναι η μετάθεση που απαιτείται από το θεώρημα.

Είναι απαραίτητο να ισχύει η υπόθεση ότι τόσο ο πίνακας Α όσο και ο πίνακας $\tilde{A}$ είναι κανονικοί. Ως υπενθύμιση, ένας πίνακας Α ονομάζεται κανονικός αν A^HA=AA^H.

Έστω ότι A= $\begin{pmatrix} 0 & 0 \\ 0 & 4 \end{pmatrix}$ και $\widetilde{A}$ = $\begin{pmatrix} -1 & -1 \\ 1 & 1 \end{pmatrix}$. Γίνεται άμεσα αντιληπτό ότι ο Α είναι κανονικός ενώ ο $\tilde{A}$ δεν είναι κανονικός. Οι ιδιοτιμές του Α είναι 0 και 4 ενώ του $\tilde{A}$ είναι και οι δύο 0. Οπότε $md_2{}^2$ (A, $\widetilde{A}$) = 16 > 12 = $\left\|\widetilde{A} - A\right\|_F^{\,2}$.

Αυτή η απαίτηση περιπλέκει την πρακτική εφαρμογή του θεωρήματος Hoffman-Wielandt, γιατί το άθροισμα των δύο κανονικών πινάκων μπορεί να μην είναι κανονικό. Επιπλέον και το άθροισμα μιας κανονικής και μιας Ερμιτιανής μήτρας μπορεί να μην είναι κανονικό. Οπότε το εύρος των διαταραχών που μπορεί να χειριστεί αυτό το θεώρημα είναι πάρα πολύ αυστηρά περιορισμένο.

Ένα χαρακτηριστικό παράδειγμα είναι η προσπάθεια εξαγωγής υπολειμματικών ορίων από ένα αποτέλεσμα οπισθοδρομικής διαταραχής. Η δυσκολία είναι ότι ο πίνακας $\tilde{A}$ δεν χρειάζεται να είναι κανονικός. Ωστόσο, το παρακάτω αποτέλεσμα δίνει ένα υπολειμματικό όριο για μια μεμονωμένη ιδιοτιμή.

ΘΕΩΡΗΜΑ : Έστω A ένας κανονικός πίνακας. Αν $\|x\|_2 = 1$ τότε $min_{1\le i\le n} |\lambda_i - x^H A\, x| \le \|Ax - (x^H Ax)x\|_2$.

Απόδειξη. Εφόσον ο πίνακας A είναι κανονικός, υπάρχει ένας ενιαίος πίνακας U τέτοιος ώστε A=UΛU^H, όπου Λ=diag (λ1,...λn). Ως εκ τούτου θα ισχύει ότι

$\|[Ax - (x^H Ax)I]x\|_2 = \|U(\Lambda - (x^H Ax)I)U^H x\|_2 \ge min_{1\le i\le n} |\lambda_i - x^H A\, x|$ από το οποίο προκύπτει το ζητούμενο.

ΔΙΑΓΩΝΟΠΟΙΗΣΙΜΟΙ ΠΙΝΑΚΕΣ.

Το κύριο γενικό συμπέρασμα για τους διαγωνοποιήσιμους πίνακες προκύπτει από το θεώρημα Bauer-Fike.

ΘΕΩΡΗΜΑ. Ας υποθέσουμε ότι ο πίνακας A είναι διαγωνοποιήσιμος, δηλαδή ισχύει X^{-1}AX=Λ, όπου το Λ είναι διαγώνιος πίνακας. Έστω τώρα μία συνεπής νόρμα πίνακα $\|.\|$ τέτοια ώστε $\|$diag(δ1,...,δn$\|$ = $max_i |\delta_i|$. Τότε $sv_A(\widetilde{A}) \le \|X^{-1}$ E X$\|$ και

$sv_A(\widetilde{A}) \le$ κ(X) $\|$ E$\|$ όπου κ(X) = $\|X\| \, \|X^{-1}\|$. Εξάλλου

md(A, $\widetilde{A}$) ≤ (2n-1) $\|X^{-1}$ E X$\|$ ≤(2n-1) κ(X) $\|$ E$\|$.

Διευκρίνιση: Συνεπής νόρμα πίνακα_ονομάζεται η νόρμα που ικανοποιεί τις παρακάτω ιδιότητες για κάθε πίνακα A και για κάθε σταθερά c.

1. Η νόρμα είναι μη αρνητική και μηδενίζεται όταν είναι ο πίνακας μηδενικός ($\|A\| \ge 0$).

2. Για κάθε σταθερά c ισχύει $\| cA \| = |c| . \| A \|$.

3. $||A+B|| \leq ||A||+||B||$ για δύο πίνακες A,B.

Απόδειξη. Έστω $\tilde{\lambda}$ είναι μια ιδιοτιμή του πίνακα $\widetilde{A}$. Με βάση τις υποθέσεις του θεωρήματος, η ανισότητα του θεωρήματος Bauer-Fike παίρνει τη μορφή $\|(\Lambda - \tilde{\lambda}\mathrm{I})^{-1}\|^{-1} \leq \|X^{-1} \text{ E X}\|$ από την οποία προκύπτει η σχέση $\|(\Lambda - \tilde{\lambda}\mathrm{I})^{-1}\|^{-1} \leq \|X^{-1} \text{ E X}\|$ από την οποία συνεπάγεται ότι $\|(\Lambda - \tilde{\lambda}\mathrm{I})^{-1}\|^{-1} \leq \|X^{-1} \text{ E X}\|$

από την οποία προκύπτει η σχέση $sv_A(\widetilde{A}) \leq \|X^{-1} \text{ E X}\|$ με άμεση συνέπεια την $sv_A(\widetilde{A}) \leq$ κ(X) $\|$ E$\|$. Οπότε προκύπτει και η τελευταία διπλή ανισότητα.

Αυτά τα όρια ισχύουν για τις ευρέως χρησιμοποιούμενες μορφές νορμών $\|.\|_p$ (p=1,2,...,∞) (και μάλιστα για όλες τις Holder νόρμες). Εφόσον ο πίνακας A είναι κανονικός τότε $sv_A(\widetilde{A}) \leq \| \mathrm{E} \|_2$. Αν και η σχέση $sv_A(\widetilde{A}) \leq \|X^{-1} \text{ E X}\|$ είναι ισχυρότερη από την $sv_A(\widetilde{A}) \leq$ κ(X) $\|$ E$\|$, συνήθως δεν θα υπάρξει τίποτα περισσότερο από μια εκτίμηση του $\|$ E$\|$ οπότε εξ ανάγκης θα χρησιμοποιηθεί το ασθενέστερο όριο. Αν αντικατασταθεί το X με το XD, όπου ο D είναι διαγώνιος, το k(X) αλλάζει, παρόλο που το X συνεχίζει να διαγωνοποιεί τον πίνακα A. Επιπλέον, κάνοντας μια στήλη του X πολύ μεγάλη ή πολύ μικρή, το k(X) γίνεται αυθαίρετα μεγάλο. Αυτές οι σκέψεις οδηγούν στο ερώτημα: ποια είναι η βέλτιστη κλιμάκωση του X; Γενικά, αυτή είναι μια πολύ δύσκολη ερώτηση. ωστόσο, για τη νόρμα Frobenius μπορεί να δοθεί μια απάντηση.

ΘΕΩΡΗΜΑ. Έστω X∈ C^{nxn} είναι μη μοναδιαίος και έστω ότι Y^H X = I. Τότε κ_F (X)≥ $\sum \|\mathrm{y_i}\|_2 \ \|\mathbf{x}_i\|_2$, με την ισότητα να ισχύει αν και μόνο αν υπάρχει α ≠ 0 τέτοιο ώστε

$\|\mathrm{y_i}\|_2$= α $\|\mathbf{x}_i\|_2$, i=1,2,…,n.

Απόδειξη. Από την ανισότητα Gauchy θα ισχύει ότι

$$\begin{aligned}\kappa_F^2(X) &= (\|x_1\|_2^2 + \cdots + \|x_n\|_2^2)(\|y_1\|_2^2 + \cdots + \|y_n\|_2^2) \\ &\geq (\|x_1\|_2\|y_1\|_2 + \cdots + \|x_n\|_2\|y_n\|_2)^2.\end{aligned}$$

Η ισότητα ισχύει εάν και μόνο εάν τα $\|\mathrm{x}_i\|_2$ και τα $\|\mathrm{y_i}\|_2$ είναι ανάλογα, δηλαδή $\|\mathrm{y_i}\|_2$= α $\|\mathbf{x}_i\|_2$, i=1,2,…,n.

Υπάρχουν δύο παρατηρήσεις που μπορούν να γίνουν σχετικά με αυτό το θεώρημα. Πρώτον, η προαναφερθείσα αναλογική κλιμάκωση δεν είναι, πιθανώς, μια κακή στρατηγική για άλλες νόρμες όπως η $\|.\|_{\mathrm{p}}$, (p=1,2,... ∞,). Δεύτερον, εάν οι ιδιοτιμές του A είναι απλές, η βέλτιστη κ_F(X) είναι το άθροισμα των επιμέρους ιδιοτιμών. Αυτό δείχνει ότι τα όρια που προκύπτουν από τη σχέση md(A, $\widetilde{A}$) ≤ (2n-1) $\|X^{-1}$ E X$\|$ ≤ (2n-1) κ(X)$\|$E$\|$ είναι ρεαλιστικά με την έννοια ότι εάν η βέλτιστη κ_F(X) είναι μεγάλη, τότε θα πρέπει να υπάρχει τουλάχιστον μία ιδιοτιμή που δεν είναι ορθά ρυθμισμένη.

Παρατηρήσεις.

Το θεώρημα Hoffman-Wielandt δίνει μια στοιχειώδη προσέγγιση που είναι ισχυρότερη του θεωρήματος του Birkhoff. Το θεώρημα Hoffman-Wielandt μπορεί να ξαναγραφτεί και με διαφορετικό τρόπο. Έστω Φ μια συμμετρική μετρική συνάρτηση και έστω $\|.\|_\Phi$ είναι μία συσχετισμένη ενιαία αμετάβλητη νόρμα. Έστω

$$\mathrm{md}_\Phi(A, \tilde{A}) = \min_\pi \Phi(|\tilde{\lambda}_{\pi(1)} - \lambda_1|^2, \ldots, |\tilde{\lambda}_{\pi(n)} - \lambda_n|^2),$$

όπου ως συνήθως το (π) κυμαίνεται σε όλες τις μεταθέσεις των ακεραίων 1,..,n. Τότε για Φ(x) ≡ $\|x\|_2$, το θεώρημα Hoffman-Wielandt δηλώνει ότι md_{Φ} (A, $\widetilde{A}$) ≤ $\|\widetilde{A} - A\|_{\Phi}$. Είναι λογικό να υποτεθεί ότι η σχέση

md_{Φ} (A, $\widetilde{A}$) ≤ $\|\widetilde{A} - A\|_{\Phi}$ αληθεύει και για κανονικούς πίνακες και για αμετάβλητες νόρμες. Η υπόθεση δεν είναι αληθής ακόμη και για ορθογώνιους πίνακες.

ΘΕΩΡΗΜΑ Weyl.

Πρόκειται για βασικό θεώρημα της θεωρίας των διαταραχών των ιδιοτιμών γιατί μέσω αυτού μπορεί να κατανοηθεί η συμπεριφορά των ιδιοτιμών ενός πίνακα όταν αυτός υποστεί μικρές διακυμάνσεις.

Η βασική διατύπωση του θεωρήματος Weyl είναι: έστω A_{nxn} ένας πίνακας με ιδιοτιμές $\lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ και A+AΔ είναι ο πίνακας που προκύπτει από μια ,μικρή διακύμανση του A και ο ΔA είναι ένας πίνακας μικρής νόρμας. Τότε για κάθε i, j θα ισχύει $|\lambda_i$ (A+AΔ) - λ_j(A)$|$ ≤ $\|\Delta A\|_2$. Αυτό σημαίνει ότι η απόλυτη διαφορά των ιδιοτιμών του A+ΔA και του A είναι περίπου μικρότερη από τη δεύτερη νόρμα του πίνακα της διακύμανσης ΔA.

Απόδειξη. Για την απόδειξη του παραπάνω θεωρήματος απαιτείται η χρήση του θεμελιώδους λήμματος της προβολής και του θεωρήματος της μετατόπισης των ιδιοτιμών. Το λήμμα της προβολής αναφέρει ότι για οποιοδήποτε υπόχωρο V του R^n και για οποιοδήποτε διανυσματικό υπόχωρο W που είναι ορθογώνιος στο V, η προβολή του x στο W μειώνει την ευκλείδεια νόρμα του x. Στη συνέχεια γίνεται χρήση του θεωρήματος Taylor για τη συνάρτηση ιδιοτιμών στο A+ΔA και ακολουθεί το λήμμα της για να εκτιμηθεί η διαφορά των ιδιοτιμών. Έστω λ_i (A+AΔ) = λ_i (A) + δ_i όπου δ_i είναι η μικρή διαφορά στις ιδιοτιμές λώτης μικρής διακύμανσης ΔA. Θεωρώντας έναν

υπόχωρο W_i που σχηματίζεται από τα ιδιοδιανύσματα που αντιστοιχούν στις i- οστές ιδιοτιμές του Α, από το λήμμα της προβολής είναι γνωστό ότι η προβολή της ΔΑ στον υποχώρο W_i μειώνει την ευκλείδεια νόρμα της. Επομένως θα ισχύει $|\delta_{\mathrm{i}}| = |\lambda_{\mathrm{i}}(A+A\Delta) - \lambda_{\mathrm{j}}(A)| \leq \|pr_{W_i}(\Delta A)\|_2$. Αφού ΔΑ είναι ένας πίνακας με μικρή νόρμα τότε η προβολή του σε κάθε υπόχωρο W_i θα είναι ομοίως πίνακας μικρής νόρμας. Άρα για κάθε $|\delta_{\mathrm{i}}|$ θα είναι περίπου μικρότερο από τη νόρμα του Δ. Αυτό έχει σαν επακόλουθο ότι
$|\delta_{\mathrm{i}}| = |\lambda_{\mathrm{i}}(A+A\Delta) - \lambda_{\mathrm{j}}(A)| \leq \|\Delta A\|_2$.

ΑΣΚΗΣΗ.

Έστω Α και $\tilde{A}$ κανονικοί πίνακες.. Αν υπάρχουν κυρτά σύνολα Β και $\tilde{B}$ τέτοια ώστε:

1. Το Β περιέχει k ιδιοτιμές του Α.
2. Το $\tilde{B}$ περιέχει τουλάχιστον (n-k+1) ιδιοτιμές του $\tilde{A}$.
3. Η απόσταση από το Β στο $\tilde{B}$ είναι (δ).

Τότε να δειχθεί ότι δ $\leq \|\widetilde{A} - A\|_2$.

Λύση.

Για να αποδειχθεί αυτή η σχέση, πρέπει να χρησιμοποιηθούν οι ιδιότητες της 2-νόρμας και οι ιδιότητες των κυρτών συνόλων. Αρχικά, ορίζεται ο κανονικός πίνακας Α= UΛU^T, όπου Λ είναι ο διαγώνιος πίνακας με τις ιδιοτιμές του Α και U είναι ο πίνακας με τα αντίστοιχα ιδιοδιανύσματα. Ομοίως $\widetilde{A} = U^t\,\Lambda'\,(U^t)^T$, όπου Λ' είναι ο διαγώνιος πίνακας με τις ιδιοτιμές του $\widetilde{A}$ και U^t είναι ο πίνακας με τα αντίστοιχα ιδιοδιανύσματα. Με βάση την εκφώνηση, το Β περιέχει k ιδιοτιμές του Α, δηλαδή αν λ_i είναι μια ιδιοτιμή του Α τότε θα ανήκει στο Β. Από την εκφώνηση πάλι, (n-k+1) ιδιοτιμές του $\tilde{A}$ περιέχονται στο $\tilde{B}$. Άρα αν λ_j' είναι μια από αυτές τις ιδιοτιμές, τότε θα ανήκει στο $\tilde{B}$.

Στη συνέχεια θα πρέπει να εκφραστεί η νόρμα $\|\widetilde{A} - \mathrm{A}\|_2$ συναρτήσει των παραπάνω ιδιοτιμών και ιδιοδιανυσμάτων. Προκειμένου να πραγματοποιηθεί αυτός υπολογισμός , θα χρησιμοποιηθεί το θεώρημα Weyl. Το θεώρημα αυτό αναδιατυπώνεται ως εξής: αν έστω λ_i η i-οστή ιδιοτιμή του πίνακα A και λ_j' η j-οστή ιδιοτιμή του πίνακα $\tilde{A}$ τότε για κάθε i και j θα ισχύει ότι

$|\lambda_j' - \lambda_i| \leq \|\widetilde{A} - \mathrm{A}\|_2$.

Τώρα αν λ_j' ανήκει στο $\tilde{B}$ τότε θα υπάρχει τουλάχιστον μια ιδιοτιμή λ_i του A που θα ανήκει στο B, τέτοια ώστε η απόλυτη τιμή της διαφοράς τους να είναι μικρότερη ή ίση με τη 2- νόρμα της διαφοράς των πινάκων $\tilde{A}$ και A, δηλαδή $|\lambda_j' - \lambda_i| \leq \|\widetilde{A} - \mathrm{A}\|_2$ = δ. Συνεπάγεται ότι αν λ_j' ανήκει στο $\tilde{B}$, θα υπάρχει τουλάχιστον μια ιδιοτιμή λ_i του A που θα ανήκει στο B που απέχει το πολύ απόσταση δ από τη λ_j'. Άρα το B θα περιέχει τουλάχιστον μια ιδιοτιμή του $\widetilde{A}$.Αυτό θα ισχύει για κάθε λ_j' που ανήκει στο $\tilde{B}$ επομένως αφού υπάρχουν τουλάχιστον (n-k+1) ιδιοτιμές του $\widetilde{A}$ στο $\tilde{B}$, το B θα περιέχει τουλάχιστον (n-k+1) ιδιοτιμές του A. Το συμπέρασμα που εξάγεται είναι ότι η απόσταση του B και του $\tilde{B}$ είναι το πολύ (δ≤ $\|\widetilde{A} - \mathrm{A}\|_2$).

ΕΦΑΡΜΟΓΗ (Bhatia and Davis).

Έστω A και $\widetilde{A}$ ορθογώνιοι πίνακες με τις ιδιοτιμές τους να βρίσκονται σε ένα ημικύκλιο του μοναδιαίου κύκλου. Οι ιδιοτιμές ταξινομούνται με τη σειρά με την οποία εμφανίζονται στο ημικύκλιο, έστω αριστερόστροφα. Να δειχθεί ότι

$max_i\,|\widetilde{\lambda_\iota} - \lambda_i| \leq \|\widetilde{A} - \mathrm{A}\|_2$.

Απόδειξη. Έστω *λi* η *i*-οστή ιδιοτιμή του πίνακα A και $\widetilde{\lambda_\iota}$ η *i*-οστή ιδιοτιμή του πίνακα $\widetilde{A}$ με τις αντίστοιχες ιδιοτιμές να είναι ταξινομημένες. Από το θεώρημα Weyl έχουμε ότι $|\widetilde{\lambda_\iota} - \lambda_i| \leq \|\widetilde{A} - \mathrm{A}\|_2$.

Επομένως $max_i\,|\widetilde{\lambda_\iota} - \lambda_i| \leq \|\widetilde{A} - \mathrm{A}\|_2$

ΚΕΦΑΛΑΙΟ ΤΕΤΑΡΤΟ

ΕΡΜΙΤΙΑΝΟΙ ΠΙΝΑΚΕΣ

Σε αυτή την ενότητα θα μελετηθεί η διαταραχή των ιδιοτιμών των Ερμιτιανών πινάκων. Πρόκειται για ένα πεδίο με πάρα πολλά αποτελεσματικά θεωρήματα εκ των οποίων θα αναφερθούν δειγματοληπτικά τα πιο σημαντικά. Η αρχή θα γίνει με δύο ενδιαφέροντα θεωρήματα, το θεώρημα αδράνειας του Sylvester και το θεώρημα παρεμβολής του Cauchy. Στη συνέχεια θα δοθεί μια σειρά ισχυρών ορίων για τις διαταραχές.

Ο πίνακας A θα υποδηλώνει έναν Ερμιτιανό πίνακα με ιδιοτιμές $\lambda_1 \geq \lambda_2 \geq \ldots \geq \lambda_n$ και το $\tilde{A}$=A+E θα δηλώνει μια ερμιτιανή διαταραχή του A με ιδιοτιμές $\tilde{\lambda}_1 \geq \tilde{\lambda}_2 \geq \ldots \geq \tilde{\lambda}_n$.

ΑΔΡΑΝΕΙΑ ΚΑΙ ΠΑΡΕΜΒΟΛΗ (Inertia and Interference).

Ένα θεμελιώδες πρόβλημα της θεωρίας πινάκων είναι να προσδιοριστεί το τι παραμένει αμετάβλητο κάτω από κάποια κατηγορία μετασχηματισμών. Για παράδειγμα, οι ιδιοτιμές και η δομή Jordan ενός πίνακα δεν μεταβάλλονται από τους μετασχηματισμούς ομοιότητας. Για τους Ερμιτιανούς πίνακες είναι φυσικό να εξεταστούν μετασχηματισμοί που αφήνουν τους πίνακες Ερμιτιανούς και οι οποίοι μετασχηματισμοί οδηγούν στην κατηγορία των μετασχηματισμών συνάφειας. Πρόκειται για μετασχηματισμούς της μορφής X^H A X, όπου ο X είναι μη μοναδιαίος. Με εξαίρεση τη περίπτωση που ο X είναι μοναδιαίος, οι ιδιοτιμές του A δεν χρειάζεται να παραμένουν αμετάβλητες σε αυτόν τον μετασχηματισμό. Ωστόσο, ο αριθμός των θετικών, αρνητικών και μηδενικών ιδιοτιμών δεν αλλάζει.

ΘΕΩΡΗΜΑ (Sylvester – Jacobi)

Έστω A ένας Ερμιτιανός πίνακας. Η έννοια της αδράνειας (inertia) ορίζεται ως η διατεταγμένη τριάδα inertia (A)=[π(A), v(A), ζ(A)] όπου π(A), v(A), ζ (A) είναι αντίστοιχα ο αριθμός των θετικών, αρνητικών και μηδενικών ιδιοτιμών του A. Τότε για κάθε μη μοναδιαίο πίνακα X θα ισχύει ότι

inertia (X^H A X) = inertia (A).

Απόδειξη. Η απόδειξη βασίζεται στη μέθοδο της ατόπου απαγωγής.Έστω, επί παραδείγματι, ότι ο πίνακας A έχει περισσότερες θετικές ιδιοτιμές από το X^H A X. Έστω Y ο χώρος που περιλαμβάνει τα ιδιοδιανύσματα που αντιστοιχούν σε θετικές ιδιοτιμές του A. Τότε y∈Y==> $y^H Ay$>0. Έστω Z το διάστημα που περιλαμβάνει όλα τα διανύσματα της μορφής Xz όπου z είναι ιδιοδιανύσματα που αντιστοιχούν σε μια αρνητική ή μηδενική ιδιοτιμή του X^H A X. Τότε z∈Z ==> $z^H Az$ >0. Ως εκ τούτου, Y ∩ Z={0}. Αλλά από την υπόθεση, dim(Y) + dim(Z) >n όπου n είναι η τάξη του A. Άρα τα X και Y έχουν ένα κοινό διάνυσμα, που είναι άτοπο.

Μια σημαντική συνέπεια του θεωρήματος αδράνειας είναι το θεώρημα του Cauchy που συσχετίζει τις ιδιοτιμές μιας κύριας υπομήτρας με τις ιδιοτιμές του αρχικού πίνακα.

ΘΕΩΡΗΜΑ (Cauchy)

Έστω B μία κύρια υπομήτρα του ερμιτιανού πίνακα A τάξης (n-1) με ιδιοτιμές $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_{n-1}$. Τότε $\lambda_1 \geq \mu_1 \geq \lambda_2 \geq \mu_2 \geq \cdots \geq \mu_{n-1} \geq \lambda_n$.

Απόδειξη. Χωρίς βλάβη της γενικότητας, έστω ότι ο πίνακας B είναι η αρχική κύρια υπομήτρα του πίνακα A έτσι ώστε να μπορεί να γραφεί ως A=$\begin{pmatrix} B & \alpha \\ \alpha^H & \alpha \end{pmatrix}$. Έστω ότι δεν ισχύει το αποδεικτέο. Τότε για κάποια i είτε θα ισχύει (μ_i>λ_i) είτε θα ισχύει (λ_{i+1}> μ_i). Έστω ότι ο i είναι ο πρώτος τέτοιος δείκτης. Θα εξεταστεί η περίπτωση με (μ_i>λ_i), γιατί η άλλη περίπτωση είναι παρόμοια. Έστω μ_i >r>λ_i. Τότε ο πίνακας (B – rI) δεν είναι μοναδιαίος και ο πίνακας

$$H = \begin{pmatrix} B - \tau I & 0 \\ 0 & \alpha - \tau - a^{\mathrm{H}}(B - \tau I)^{-1} a \end{pmatrix} =$$

$$\begin{pmatrix} I & 0 \\ -a^{\mathrm{H}}(B - \tau I)^{-1} & 1 \end{pmatrix} \begin{pmatrix} B - \tau I & a \\ a^{\mathrm{H}} & \alpha - \tau I \end{pmatrix} \begin{pmatrix} I & -(B - \tau I)^{-1} a \\ 0 & 1 \end{pmatrix}$$

ταυτίζεται με τον (A – rI). Ως εκ τούτου, από το θεώρημα αδράνειας, ο H έχει τον ίδιο αριθμό θετικών ιδιοτιμών με τον (A-rI), δηλαδή (i-1). Αλλά ο H έχει τουλάχιστον τόσες θετικές ιδιοτιμές όσες και ο (B-rI) , δηλαδή i. Προκύπτει επομένως ένα άτοπο, οπότε θα ισχύει και το αποδεικτέο.

Εάν, στο θεώρημα, ο C είναι μία κύρια υπομήτρα του A τάξης (n-2) τότε οι ιδιοτιμές $\nu_1 \geq \nu_2 \geq \cdots \geq \nu_{n-2}$ του C ικανοποιούν την $\mu_1 \geq \nu_1 \geq \mu_2 \geq \nu_2 \geq \ldots \geq \nu_{n-1} \geq \mu_{n-1}$. Άρα θα ισχύει ότι

$\lambda_i \geq \nu_i \geq \lambda_{i+2}$, με i = 1, 2,…, n-2. Συνεχίζοντας με αυτό τον τρόπο την ίδια διαδικασία μέσα από υποπίνακες προκύπτει το ακόλουθο συμπέρασμα.

Συμπέρασμα. Έστω B ένας κύριος υποπίνακας τάξης (n-k) του A με ιδιοτιμές $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_{n-k}$. . Τότε $\lambda_i \geq \mu_i \geq \lambda_{i+k}$, με i=1,2,...,n-k.

Τελικά, το θεώρημα της σύμπλεξης ισχύει για περισσότερους απλούς κύριους υποπίνακες. Έστω ότι ο U∈ $\mathrm{C}^{\mathrm{mx(n-k)}}$ έχει ορθοκανονικές στήλες και ο V επιλέγεται έτσι ώστε ο (UV) να είναι μοναδιαίος. Στη συνέχεια, εφαρμόζοντας το παραπάνω συμπέρασμα στον πίνακα $(\mathrm{UV})^H$ A (UV) συνεπάγεται ότι:

Έστω ότι ο U∈ $\mathbb{C}^{m\times(n-k)}$ έχει ορθοκανονικές στήλες και έστω επίσης ότι οι ιδιοτιμές του U^H AU είναι $\mu_1 \geq \mu_2 \geq ... \geq\geq \mu_{n-k}$. Τότε θα ισχύει $\lambda_i \geq \mu_i \geq \lambda_{i+k}$, με i=1,2,...,n-k.

ΤΟ ΘΕΩΡΗΜΑ WIELANDT ΚΑΙ ΟΙ ΣΥΝΕΠΕΙΕΣ ΤΟΥ.

ΘΕΩΡΗΜΑ (WIELANDT).

Έστω $1 \leq i_1 \leq i_2 \leq \ldots \leq i_k \leq$ n. Τότε

$$\lambda_{i_1} + \lambda_{i_2} + \cdots + \lambda_{i_k} = \max_{\substack{\mathcal{X}_{i_1} \subset \mathcal{X}_{i_2} \subset \cdots \subset \mathcal{X}_{i_k} \\ \dim(\mathcal{X}_{i_j}) = i_j}} \; \min_{\substack{X = (x_{i_1}\ x_{i_2}\ \cdots\ x_{i_k}),\, x_{i_j} \in \mathcal{X}_{i_j} \\ X^H X = I}} \operatorname{trace}(X^H A X),$$

και

$$\lambda_{i_1} + \lambda_{i_2} + \cdots + \lambda_{i_k} = \min_{\substack{\mathcal{X}_{i_1} \supset \mathcal{X}_{i_2} \supset \cdots \supset \mathcal{X}_{i_k} \\ \dim(\mathcal{X}_{i_j}) = n - i_j + 1}} \; \max_{\substack{X = (x_{i_1}\ x_{i_2}\ \cdots\ x_{i_k}),\, x_{i_j} \in \mathcal{X}_{i_j} \\ X^H X = I}} \operatorname{trace}(X^H A X).$$

Σημειωτέον ότι οι λέξεις max και min (αντί για sup και inf) υποδηλώνουν ότι υπάρχει τόσο η μεγιστοποίηση όσο και η ελαχιστοποίηση των συγκεκριμένων μεγεθών..

Απόδειξη. Αρχικά θα δειχθεί ότι αν υπάρχει μια συγκεκριμένη ακολουθία $X_{i1} \subset X_{i2} \subset ... \subset X_{ik}$ υποχώρων με dim (X_{ij})=ij όπου ο X=(x_{i1} ,x_{i2} ,... x_{ik}), με $x_{ij} \in X_{ij}$ έχει ορθοκανονικές στήλες, τότε το (ίχνος) trace (X^HA X) ≥ $\Sigma_{ij}\ \lambda_{ij}$. Θεωρώντας ότι ο X_{ij} είναι ο χώρος που περιλαμβάνει τα ιδιοδιανύσματα του A που αντιστοιχούν στα λ_1, λ_2, …, $\lambda_{ij,}$ τότε το x_{ij} θα είναι ένας γραμμικός συνδυασμός αυτών των ιδιοδιανυσμάτων και εφόσον $x_{ij}{}^H\ x_{ij}$ =1, θα ισχύει $x_{ij}{}^H$ A$x_{ij} \geq \lambda_{ij}$. Συνεπάγεται επομένως ότι trace (X^HA X) = $\sum_{ij} x_{ij}{}^H$ A$x_{ij} \geq \Sigma_{ij}$ λ_{ij}. Άρα αρκεί να αποδειχθεί ότι :

$$\max_{\substack{\mathcal{X}_{i_1} \subset \mathcal{X}_{i_2} \subset \cdots \subset \mathcal{X}_{i_k} \\ \dim(\mathcal{X}_{i_j}) = i_j}} \min_{\substack{X=(x_{i_1}\ x_{i_2}\ \cdots\ x_{i_k}),\, x_{i_j} \in \mathcal{X}_{i_j} \\ X^H X = I}} \operatorname{trace}(X^{\mathrm{H}} A X) \leq \lambda_{i_1} + \lambda_{i_2} + \cdots + \lambda_{i_k}.$$

Η απόδειξη θα γίνει με απαγωγή στο n. Να σημειωθεί ότι το θεώρημα δεν παρουσιάζει σημαντικότητα για k=n, αφού στη περίπτωση αυτή ο X^HA X είναι όμοιος με τον A. (Υπενθύμιση: στη Γραμμική Άλγεβρα δύο πίνακες A, B θα ονομάζονται όμοιοι αν υπάρχει ένας αντιστρέψιμος πίνακας P τέτοιος ώστε A = P^{-1} B P). Άρα το θεώρημα ισχύει για n=1. Έστω τώρα ότι n>1 και k<n και οι X_{i1} ⊂ X_{i2} ⊂... ⊂ X_{ik} με dim (X_{ij})=ij έχουν δοθεί. Πρέπει να δειχθεί ότι υπάρχει ένας πίνακας X=(x_{i1} , x_{i2} ,... x_{ik}), με ορθοκανονικές στήλες τέτοιες ώστε trace (X^HA X) ≤ λ_1+ λ_2+ …+ $\lambda_{i\kappa}$.

Αρχικά υποτίθεται ότι i_k <n. Έστω $\check{X}_{n-1}$ ένας υπόχωρος με διάσταση (n-1) που περιέχει τον X_{ik}. Έστω Z=(z_1, z_2,…, z_{n-1}) είναι ένας πίνακας με ορθοκανονικές στήλες τέτοιος ώστε R[(z_1,…, z_{ij})] = X_{ij} και R(Z) = $\check{X}_{n-1}$. Έστω τώρα ένας πίνακας B=Z^HAZ. Σύμφωνα με το προαναφερθέν συμπέρασμα, οι ιδιοτιμές μ_i του πίνακα B ικανοποιούν την ανίσωση $\mu_i \leq \lambda_i$, όπου $i = 1,2,\ldots,$n-1.

Στη συνέχεια, έστω Y_{ij} = {Z^Hx : x∈X_{ij}}. Παρατηρείται ότι αφού X_{ij} ⊂ R(Z), αν y ∈ Y_{ij}, τότε x = Zy ∈ X_{ij}. Επιπλέον y^HBy = x^HAx. Από τις υποθέσεις, υπάρχουν ορθοκανονικά διανύσματα y_{ij} ∈ Y_{ij}, τέτοια ώστε $\sum_j y_{ij}^H$ B $y_{ij} \leq \sum_j \mu_{ij}$. Όμως αν x_{ij} = Z y_{ij}, τότε x $_{ij}$ ∈ X_{ij} και $\sum_j y_{ij}^H$ B $y_{ij} = \sum_j x_{ij}^H$ A x_{ij}.

Επομένως θα ισχύει trace (X^HA X) = $\sum_{ij} {x_{ij}}^H$ A$x_{ij} \leq \sum_j \lambda_{ij}$.

Τώρα ας υποτεθεί ότι i_k = n και έστω l είναι ι μεγαλύτερος δείκτης τέτοιος ώστε i_l + 1 < i_{l+1}. Για λόγους ευκολίας θα συμβολιστούν το (i_l) με (p) και το (i_{l+1}) με (q). Έστω $\check{X}_{n-1}$ ένας (n-1)-διάστατος υποχώρος που περιέχει το X_p και τα ιδιοδιανύσματα που αντιστοιχούν στις ιδιοτιμές λ_1, λ_2, … λ_n. Εφόσον τα q,q+1,...,n-1 είναι μεταξύ των δεικτών i_j, θα ισχύει ότι X_p ⊂ X_q ∩ $\check{X}_{n-1}$ ⊂…⊂

X_{n-1} $\cap \check{X}_{n-1} \subset$ $\check{X}_{n-1}$. Εφόσον για i=q,...,n-1, dim{X_i $\cap$ $\check{X}_{n-1}$} $\geq$ n-1 είναι δυνατόν να βρεθούν υπόχωροι $\check{X}_{q-1}, \ldots, \check{X}_{n-2}$ έτσι ώστε $\check{X}_{q-1} \subset X_q, \ldots, \check{X}_{n-2} \subset X_{n-1}$ και $X_{i1} \subset \ldots \subset X_p \subset$ $\check{X}_{q-1} \subset \ldots \subset \check{X}_{n-1}$.

Τώρα θα εφαρμοστεί η παραπάνω κατασκευή για να δοθεί ένας πίνακας B με ιδιοτιμές $\mu_1 \geq \mu_2 \geq \ldots \geq \mu_{n-1}$ που να ικανοποιούν τη σχέση $\mu_i \leq \lambda_i$, όπου $i = 1,2,\ldots,$n-1.

και ένας πίνακας X=($\mathrm{x}_{i1} \ldots \mathrm{x}_p$ $\mathrm{x}_q \ldots$ x_{n-1}) έτσι ώστε $\mathrm{x}_{ij} \in X_{ij}$ με j=1,…,l και $\mathrm{x}_i \in \check{X}_i \subset X_{i+1}$ όπου i=q-1,…,n-1. Επιπλέον

trace (X^HA X) $\leq \sum_{j=1}^{l} \mu_{ij}$ + $\sum_{i=q-1}^{n-1} \mu_i \leq \sum_{j=1}^{l} \lambda_{ij}$ +$\sum_{i=q-1}^{n-1} \mu_i$.

Η κατασκευή $\check{X}_{n-1}$ περιέχει τα ιδιοδιανύσματα του A που αντιστοιχούν στις ιδιοτιμές $\lambda_q, \ldots,$ λ_n. Επομένως και αυτές οι ιδιοτιμές είναι ιδιοτιμές του B. Εφόσον τα μ_{q-1},...,μ_{n-1} είναι οι μικρότερες ιδιοτιμές του B, θα ισχύει ότι $\sum_{i=q-1}^{n-1} \mu_i \leq \sum_{i=q}^{n} \lambda_i$. Οπότε με αντικατάσταση επαληθεύεται η ανισότητα

trace (X^HA X) $\leq \sum_{j=1}^{l} \mu_{ij}$ + $\sum_{i=q-1}^{n-1} \mu_i \leq \sum_{j=1}^{l} \lambda_{ij}$ +$\sum_{i=q-1}^{n-1} \mu_i$.

Όταν στο θεώρημα του Wielandt, το k=1 τότε δίνεται ο χαρακτηρισμός (συμπέρασμα) κατά Fischer για τις ιδιοτιμές ενός Ερμιτιανού πίνακα. Το εν λόγω συμπέρασμα αναφέρει ότι οι ιδιοτιμές του πίνακα A δίνονται από τις σχέσεις λ_i = $max_{\dim(X)=i}$ $min_{x \in \mathrm{X}_{x^H x=1}}$ x^HAx και

λ_i = $max_{\dim(X)=i}$ $min_{x \in \mathrm{X}_{x^H x=1}}$ x^HAx και

λ_i = $min_{\dim(X)=n-i+1}$ $max_{x \in \mathrm{X}_{x^H x=1}}$ x^HAx. Για i =1 η δεύτερη από τις παραπάνω σχέσεις μετασχηματίζεται σε λ_1 = $max_{x^H x=1}$ x^HAx.

Η τελευταία σχέση έχει σημαντικές επιπτώσεις στη θεωρία των διαταραχών. Έστω ότι ο πίνακας $\tilde{A}$=A+E, όπου ο E είναι επίσης ερμιτιανός πίνακας. Θέτοντας με $\widetilde{\lambda_1}$ και ε_1 αντίστοιχα τις μεγαλύτερες ιδιοτιμές των $\tilde{A}$ και E, θα ισχύει ότι $\widetilde{\lambda_1}$ = $max_{x^H x=1}$ $\mathrm{x}^H \tilde{A}$x $\leq$ $max_{x^H x=1}$ x^HAx +$max_{x^H x=1}$ x^HEx$\leq$ $\leq \lambda_1$ +

ε_1. Με άλλα λόγια , όσο $|\varepsilon_1| \le \|E\|_2$ η διαταραχή Ε είναι σε θέση να αυξήσει τη μεγαλύτερη ιδιοτιμή του πίνακα Α αλλά όχι περισσότερο από τη τιμή της νόρμας $\|E\|_2$.

Στη συνέχεια θα γίνει μια προσπάθεια να γενικευθεί αυτό το συμπέρασμα. Αρχικά θα καθοριστεί ένα αποτέλεσμα για αθροίσματα ιδιοτιμών και στη συνέχεια θα γίνει μια εξατομίκευση σε μια μεμονωμένη ιδιοτιμή.

ΘΕΩΡΗΜΑ. Έστω ότι οι ιδιοτιμές του ερμιτιανού πίνακα E_{nxn} είναι

$\varepsilon_1 \ge \varepsilon_2 \ge \ldots \ge \varepsilon_n$ και έστω ακόμα ότι i_1, i_2,…, i_k είναι διακριτοί ακέραιοι μεταξύ του 1 και του n συμπεριλαμβανομένων και των δύο αυτών άκρων. Τότε θα ισχύει

$\lambda_{i1} + \ldots + \lambda_{ik} + \varepsilon_{n-k+1} + \ldots + \varepsilon_n \le \tilde{\lambda}_{i1} + \ldots + \tilde{\lambda}_{i\kappa} \le \lambda_{i1} + \ldots + \lambda_{ik} + \varepsilon_{1+} + \ldots + \varepsilon_{\kappa}$.

Απόδειξη. Χωρίς βλάβη της γενικότητας μπορεί να υποτεθεί ότι $i_1 < \ldots < i_n$ και αρχικά θα αποδειχθεί τη δεύτερη ανισότητα.

Με βάση προαναφερθέντα θεωρήματα θα υπάρχουν υπόχωροι

$X_{i1} \subset X_{i2} \subset \ldots \subset X_{ik}$ έτσι ώστε

$$\tilde{\lambda}_{i_1} + \cdots + \tilde{\lambda}_{i_k} = \min_{\substack{X=(x_{i_1}\ x_{i_2}\ \cdots\ x_{i_k}),\, x_{i_j} \in X_{i_j} \\ X^H X = I}} \operatorname{trace}(X^H \tilde{A} X).$$

Επιπλέον υπάρχουν διανύσματα $x_{ij} \in X_{ij}$ έτσι ώστε ο πίνακας X=($x_{i1} \ldots x_{i\kappa}$) unitary matrix και $\lambda_{i1} + \ldots + \lambda_{ik} \ge$ trace (X^HA X). Αυτό συνεπάγεται ότι

$\tilde{\lambda}_{i1} + \ldots + \tilde{\lambda}_{i\kappa} \le$ trace [X^H(A+E) X] $\le \lambda_{i1} + \ldots + \lambda_{ik}$ +

+ trace (X^HA X). Επιώθηκε παραπάνω ότι αν Β ένας κύριος υποπίνακας τάξης (n-k) του Α με ιδιοτιμές $\mu_1 \ge \mu_2 \ge \ldots \ge \mu_{n-k}$. τότε $\lambda_i \ge \mu_i \ge \lambda_{i+k}$, με i=1,2,...,n-k. Άρα trace (X^HA X) $\le \varepsilon_{1+} + \ldots + \varepsilon_{\kappa}$, που αποδεικνύει τη δεύτερη ανισότητα. Η πρώτη ανισότητα μπορεί να ληφθεί από τη δεύτερη θέτοντας

A=$\tilde{A}$ - E, από την οποία προκύπτει ότι λ_{i1} +…+λ_{ik} $\leq$ $\tilde{\lambda}_{i1}$ +…+$\tilde{\lambda}_{i\kappa}$ -ε_{n-k+1} -…-ε_n. Όταν k=1, το θεώρημα παρέχει ένα όριο διαταραχής.

Ο Weyl συμπέρανε ότι για i=1,2,…,n θα ισχύει ότι $\tilde{\lambda}_i$ ∈ [λ_i+ε_n, λ_i+ε_1]. Σχετικά με το συμπέρασμα αυτό θα πρέπει να επισημανθούν τρία σημεία. Πρώτον, το συμπέρασμα είναι παρόμοιο με το θεώρημα Gerschgorin γιατί παρέχει ένα σύνολο (n) διαστημάτων (δίσκους) των οποίων η ένωση Ωστόσο, είναι γνωστή ποια ιδιοτιμή ακριβώς αναζητείται σε κάθε διάστημα. Επιπλέον, είναι αδύνατο για μια ιδιοτιμή που περιλαμβάνει τις ιδιοτιμές του πίνακα $\tilde{A}$. αντιστοιχεί σε ένα σύμπλεγμα αλληλοκαλυπτόμενων διαστημάτων να μεταφερθεί εκτός του διαστήματος που ανήκει. Δεύτερον, τα διαστήματα δεν είναι συμμετρικά ως προς τις ιδιοτιμές λ_i. Στην πραγματικότητα, εάν το ε_n είναι θετικό, το i-στο διάστημα δεν θα περιέχει την ιδιοτιμή λ_i. Αυτό παρατηρείται όταν ο πίνακας E είναι θετικά ορισμένος. Με άλλα λόγια, εάν ο Ερμιτιανός πίνακας διαταράσσεται από θετικά ορισμένο πίνακα, οι ιδιοτιμές του θα πρέπει να αυξηθούν. Τρίτον, υπάρχει μια πιο αδύναμη και πιο συμβατική μορφή του θεωρήματος που εκφράζεται μέσω του ακόλουθου συμπεράσματος.

Συμπέρασμα. Θα ισχύει ότι max{$|\tilde{\lambda}_i$-$\lambda_i|$}≤ $\|\mathrm{E}\|_2$.

Αυτό το αποτέλεσμα προκύπτει άμεσα από το προηγούμενο συμπέρασμα και την παρατήρηση ότι $\|\mathrm{E}\|_2$=max {$|\varepsilon_1|$,$|\varepsilon_n|$}. Στη συνέχεια αυτό το συμπέρασμα θα γενικευθεί.

ΤΟ ΘΕΩΡΗΜΑ Mirsky.

Η ανίσωση ότι max{$|\tilde{\lambda}_i$-$\lambda_i|$}≤ $\|\mathrm{E}\|_2$ δύναται να ξαναγραφεί στη πιο συμμετρική μορφή $\|\mathrm{diag}\,(\tilde{\lambda}_i - \lambda_i)\|_2 \leq \|\mathrm{E}\|_2$.Η σχέση αυτή υποδηλώνει ότι γίνεται προσπάθεια αντικατάστασης της νόρμας $\|.\|_2$ με άλλους κανόνες για να ληφθούν νέα όρια διαταραχής. Στην πραγματικότητα η σχέση

max{$|\tilde{\lambda}_i$-$\lambda_i|$}≤ $\|\mathrm{E}\|_2$ ισχύει για κάθε μοναδιαία αμετάβλητη νόρμα.

ΘΕΏΡΗΜΑ (Mirsky). Έστω X και $\tilde{X}$ δύο ερμιτιανοί πίνακες ίσων διαστάσεων με μοναδικές τιμές $\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_p$ και

$\widetilde{\sigma_1} \geq \widetilde{\sigma_2} \geq \ldots \geq \widetilde{\sigma_p}$ αντίστοιχα. Τότε για κάθε μοναδιαία αμετάβλητη νόρμα $\|.\|$ θα ισχύει ότι

$\|\mathrm{diag}\,(\tilde{\sigma}_i - \sigma_i)\| \leq \|\tilde{X}$ - X$\|$.

Απόδειξη. Χωρίς βλάβη της γενικότητας, μπορεί να υποτεθεί ότι οι πίνακες X και $\tilde{X}$ είναι τετραγωνικοί (διαφορετικά αν δεν είναι, συμπληρώνοντας με μηδενικά σειρές και στήλες μπορούν να μετασχηματιστούν σε τετραγωνικούς). Εξετάζεται ο πίνακας

$\begin{pmatrix} 0 & X \\ X^H & 0 \end{pmatrix}$. Οι ιδιοτιμές του είναι σ_1, -σ_1, σ_2, -σ_2,…, σ_p, -σ_p. Ομοίως και στη περίπτωση που ο πίνακας περιέχει τον $\tilde{X}$. Τελικά αν $\varepsilon_1 \geq \ldots \geq \varepsilon_n$ είναι οι απλές τιμές του ($\tilde{X}$ – X), τότε οι ιδιοτιμές του πίνακα $\begin{pmatrix} 0 & \tilde{X} - \mathrm{X} \\ (\tilde{X} - \mathrm{X})^H & 0 \end{pmatrix}$ θα είναι

ε_1, -ε_1,…, ε_n, -ε_n. Υπενθυμίζοντας το θεώρημα με τη διατύπωση: έστω ότι οι ιδιοτιμές του πίνακα E είναι $\varepsilon_1 \geq \varepsilon_2 \geq \ldots \geq \varepsilon_n$ και έστω ακόμα ότι i_1, i_2,…, i_k είναι διακριτοί ακέραιοι μεταξύ του 1 και του n συμπεριλαμβανομένων και των δύο αυτών άκρων. Τότε θα ισχύει ότι

λ_{i1} +…+λ_{ik} +ε_{n-k+1} +…+$\varepsilon_n \leq \tilde{\lambda}_{i1}$ +…+$\tilde{\lambda}_{i\kappa} \leq \lambda_{i1}$ +…+λ_{ik} +ε_{1+} +…+ε_κ, αν θα τεθεί ότι για $\widetilde{\sigma_\kappa} \geq \sigma_\kappa$ το i_κ = κ και για

$\widetilde{\sigma_\kappa} < \sigma_\kappa$ το i_κ = n + κ, θα συνεπάγεται ότι

$$|\tilde{\sigma}_1 - \sigma_1| + \cdots + |\tilde{\sigma}_k - \sigma_k| \leq \epsilon_1 + \cdots + \epsilon_k, \qquad k = 1, \ldots, n.$$

Θεωρώντας την ανισότητα Φ $(\tilde{\sigma}_1 - \sigma_{1,\ldots,}\tilde{\sigma}_p - \sigma_p) \leq$ Φ $(\varepsilon_1,\ldots,\ \varepsilon_p)$, αυτή θα ισχύει για κάθε συμμετρική συνάρτηση Φ. Το αποτέλεσμα τώρα θα προκύψει από τον χαρακτηρισμό του von Neumann για τις μοναδιαίες αμετάβλητες νόρμες.

Άμεση συνέπεια του θεωρήματος του Mirsky είναι η γενίκευση του συμπεράσματος του Fischer που ειπώθηκε παραπάνω. Συγκεκριμένα, εξάγεται το εξής συμπέρασμα: έστω Φ μια συμμετρική συνάρτηση και $\|.\|_{\Phi}$ η αντίστοιχη αμετάβλητη νόρμα. Τότε $\|\mathrm{diag}\,(\tilde{\lambda}_i - \lambda_i)\|_{\Phi} \leq \|\mathrm{E}\|_{\Phi}$.

Απόδειξη. Έστω ρ = min { λ_n, $\overline{\lambda_n}$ }. Τότε οι ιδιοτιμές των πινάκων (A – ρI) και ($\tilde{A}$ – ρI) είναι μη αρνητικές. Δηλαδή, οι μοναδικές τους τιμές και οι ιδιοτιμές τους είναι ίδιες. Άρα ισχύει το θεώρημα του Mirsky και δίνει την ανισότητα
$\|\mathrm{diag}\,(\tilde{\lambda}_i - \lambda_i)\|_{\Phi} \leq \|\mathrm{E}\|_{\Phi}$.

Όταν η Φ δημιουργεί την νόρμα Frobenius, τότε λαμβάνεται ένα αντίστοιχο ερμιτιανό ανάλογο του θεωρήματος Hoffman-Wielandt.

ΣΥΜΠΕΡΑΣΜΑ. $\sqrt{\sum_{i=1}^{n}(\tilde{\lambda}_i - \lambda_i)^2} \leq \|\mathrm{E}\|_F$

Πρέπει να σημειωθεί ότι αυτό το αποτέλεσμα είναι ισχυρότερο από το θεώρημα Hoffman-Wielandt, καθώς καθορίζει μια σειρά των ιδιοτιμών που ικανοποιούν την ανισότητα, ενώ το θεώρημα Hoffman-Wielandt απλώς επιβεβαιώνει ότι υπάρχει μια τέτοια διάταξη.

ΥΠΟΛΕΙΜΜΑΤΙΚΑ ΟΡΙΑ.

Στη συνέχεια θα εξεταστούν διάφορες εφαρμογές του θεωρήματος του Mirsky. Αρχικά θα μελετηθούν τα υπολειμματικά όρια.

ΘΕΩΡΗΜΑ. Αν ο ερμιτιανός πίνακας X∈R^{nxk} έχει ορθοκανονικές στήλες, τότε κάθε μοναδιαία αμετάβλητη νόρμα του R ελαχιστοποιείται όταν M=X^H A X.

Απόδειξη. Έστω μια στήλη x_i του πίνακα X. Η αντίστοιχη μοναδιαία αμετάβλητη νόρμα του x_i θα είναι $\|x_i\|$ = $(x_i{}^H\ x_i)^{1/2}$. Έστω τώρα ένας οποιοσδήποτε ερμιτιανός πίνακας R και M οποιοσδήποτε πίνακας παράγεται από τη σχέση

X^H A X. Η μοναδιαία αμετάβλητη νόρμα του R θα είναι $\|$ R$\|$ = $max_{\|\mathrm{v}\|=1}|v^H$ R$v|$. Αν τεθεί v = X w, όπου w είναι ένα διάνυσμα της μοναδιαίας νόρμας τότε: v^H Rv = $(\mathrm{X\ w})^H$ R (X w) = X^H w^H R X w. Εφόσον τεθεί M=X^H A X τότε θα ισχύει w^H M w = w^H X^H A X w = v^H A v. Καθότι η επιλογή του w ή του v είναι οποιοδήποτε διάνυσμα της μοναδιαίας νόρμας, οι εκφράσεις (v^H Rv) και

(v^H A v) αντιπροσωπεύουν τη μέγιστη απόλυτη τιμή των αντίστοιχων συναρτήσεων για κάθε δυνατή επιλογή του (v). Άρα η μοναδιαία αμετάβλητη νόρμα του R ελαχιστοποιείται όταν M=X^H A X.

Έστω ότι δίνεται ένας πίνακας A (στη συγκεκριμένη περίπτωση ο A είναι ερμιτιανός) και ένας πίνακας X του οποίου ο χώρος στηλών προσεγγίζει έναν αμετάβλητο υποχώρο του πίνακα A. Αυτό σημαίνει ότι για κάποια επιλογή του M, το (υπολειπόμενο) αποτέλεσμα R = AX – XM θα είναι μικρό.

Δεδομένου ότι ο A είναι ερμιτιανός πίνακας , είναι δυνατόν να χρησιμοποιηθεί το θεώρημα του Mirsky προκειμένου να εξαχθεί ένα όριο για τις ιδιοτιμές του M ως προσέγγιση των αντίστοιχων ιδιοτιμών του πίνακα A.

ΘΕΩΡΗΜΑ. Έστω ότι ο ερμιτιανός πίνακας X ∈ R^{nxk} έχει ορθοκανονικές στήλες. Έστω M=X^H A X και R = AX – XM. Έστω ότι η Φ είναι μια συμμετρική συνάρτηση στο R^n και έστω ότι μέσω της $\|.\|_{\Phi}$. δηλώνεται η αντίστοιχη κλάση των ενιαία αμετάβλητων νορμών. Αν οι ιδιοτιμές του A είναι $\lambda_1 \geq \ldots \geq \lambda_n$

και οι ιδιοτιμές του M είναι $\mu_1 \geq \ldots \geq \mu_\kappa$, τότε θα υπάρχουν ακέραιοι $i_1 < \ldots < i_\kappa$ έτσι ώστε:

$\left\| \mathrm{diag}\,(\mu_j - \lambda_{ij}) \right\|_\Phi \leq \left\| \mathrm{XR^H} + \mathrm{R\,X^H} \right\|_\Phi = \Phi(\rho_1, \rho_2, \rho_3, \ldots)$,

όπου $\rho_1 \geq \rho_2 \geq \ldots$είναι οι μοναδικές τιμές του R.

Απόδειξη. Η πιο ενδιαφέρουσα περίπτωση είναι όταν ισχύει 2κ≤n. Για M=X^H A X, έστω E = - (XR^H + RX^H). Τότε ο πίνακας E είναι ερμιτιανός και επαληθεύεται εύκολα ότι

(A+E) X = XM. Έτσι το R(X) είναι ένας αμετάβλητος υποχώρος του A+E και σε κάθε ιδιοτιμή μ_jτου M αντιστοιχεί μια ιδιοτιμή $\widetilde{\lambda_{ij}}$του A+E.

Από το θεώρημα του Mirsky θα ισχύει ότι

$\left\| \mathrm{diag}\,(\tilde{\lambda}_i - \lambda_i) \right\|_\Phi \leq \left\| \mathrm{E} \right\|_\Phi$. Εντούτοις συνεπάγεται ότι

$\left\| \mathrm{diag}\,(\mu_i - \lambda_{ij}) \right\|_\Phi \leq \left\| \mathrm{E} \right\|_\Phi = \left\| \mathrm{X}R^H + RX^H \right\|_\Phi$

και απομένει μόνο η επαλήθευση της σχέσης

$\left\| \mathrm{diag}\,(\mu_j - \lambda_{ij}) \right\|_\Phi \leq \left\| \mathrm{XR^H} + \mathrm{R\,X^H} \right\|_\Phi = \Phi(\rho_1, \rho_2, \rho_3, \ldots)$ ή

ισοδύναμα ότι οι απλές τιμές του πίνακα E είναι ρ1,ρ1,ρ2,ρ2.... Αλλά αν επιλεγεί ο $X_\perp$ έτσι ώστε ο (X $X_\perp$) να είναι μοναδιαίος , τότε θα ισχύει το εξής:

$(\mathrm{X}\,X_\perp)^H$ E $(\mathrm{X}\,X_\perp)$ = $\begin{pmatrix} 0 & R^H X_\perp \\ {X_\perp}^H R & 0 \end{pmatrix}$. Από R(R) ⊂ R($X_\perp$) και από το γεγονός ότι οι στήλες του πίνακα $X_\perp$ είναι ορθοκανονικές, συμπεραίνεται ότι οι απλές τιμές του πίνακα ${X_\perp}^H R$ είναι ίδιες με εκείνες του πίνακα R καθώς και εκείνες του πίνακα E είναι ίδιες με τις αντίστοιχες του πίνακα R επαναλαμβανόμενες.

Για λόγους πληρότητας είναι χρήσιμο να παρατεθούν τα όρια για την φασματική νόρμα και για τη νόρμα Frobenius. Ειδικότερα για τη φασματική νόρμα θα ισχύει

$max_j \ \{ | \ \mu_j - \lambda_{ij} | \} \ \leq \ \left\| R \right\|_2$, ενώ για τη νόρμα Frobenius θα ισχύει $\sqrt{\sum_j (\mu_j - \lambda_{ij})^2} \leq \sqrt{2}\ \left\| R \right\|_F$.

Με την εφαρμογή των συνθηκών του θεωρήματος Hoffman-Wielandt, ο Kahan κατάφερε να αφαιρέσει τον παράγοντα $\sqrt{2}$ από τη τελευταία ανισότητα όμως τα υπολειμματικά όρια που προκύπτουν από τη νέα σχέση ενδέχεται να είναι ή πολύ καλά ή πολύ άσχημα.

Παρακάτω θα παρατεθεί ο τρόπος με τον οποίο ο Kahan κατάφερε να αφαιρέσει τον όρο $\sqrt{2}$ από τη σχέση $\sqrt{\sum_j (\mu_j - \lambda_{ij})^2} \leq \sqrt{2}\ \|R\|_F$.

Έστω ο ερμιτιανός πίνακας A και ο πίνακας M=X^H A X είναι διαγώνιοι ενώ ο R = AX – XM είναι συνάρτηση του X ή γενικότερα U= (X $\check{X}$) όπου ο U είναι μοναδιαίος. Τότε W=‖U‖ και δ_{ij} = (λ_i - μ_j), όπου λ_i οι ιδιοτιμές του πίνακα A και μ_j οι ιδιοτιμές του πίνακα M. Αν j≤k τότε η δ_{ij} είναι διαφορά ιδιοτιμών, διαφορετικά είναι μηδέν. Στη συνέχεια θα βρεθεί η νόρμα Frobenius του R.
$\|R\|_F{}^2 = \sum_i \sum_j |R_{ij}|^2$. Γνωρίζοντας ότι ο R είναι διαγώνιος πίνακας, θα ισχύει ότι R_{ij} = 0 όταν i ≠ j. Άρα το ενδιαφέρον εστιάζεται στις μη μηδενικές θέσεις, δηλαδή όταν i = j. Σε αυτή τη περίπτωση θα ισχύει
$\|R\|_F{}^2 = \sum_i |R_{ii}|^2$. Όμως ισχύει και η σχέση R = AX – XM = (X – X) Λ, όπου Λ η διαγώνια μορφή του πίνακα A. Επομένως R_{ii} = (λ_i - μ_i) και έπεται ο μετασχηματισμός $\|R\|_F{}^2 = \sum_i (\lambda_i - \mu_i)^2$. Βάσει του θεωρήματος του Birkhoff η ευκλείδεια νόρμα του R ελαχιστοποιείται όταν οι στήλες του πίνακα X είναι ορθοκανονικές (δηλαδή όταν ο πίνακας U είναι πίνακας μετάθεσης).

Με αυτό τον τρόπο ο Kahan αφαίρεσε τον όρο $\sqrt{2}$ από τη σχέση $\sqrt{\sum_j (\mu_j - \lambda_{ij})^2} \leq \sqrt{2}\ \|R\|_F$ αποδεικνύοντας ότι η η ευκλείδεια νόρμα του R ελαχιστοποιείται όταν οι στήλες του πίνακα X είναι ορθοκανονικές.

Παραπάνω ειπώθηκε ότι τα υπολειμματικά όρια που προκύπτουν από αυτή τη νέα σχέση ενδέχεται να είναι ή πολύ καλά ή πολύ άσχημα.

Παράδειγμα. Έστω ο πίνακας A= $\begin{pmatrix} 0 & \varepsilon \\ \varepsilon & 0 \end{pmatrix}$ και X= 1_1. Τότε M=0 και $\|R\|_2$ = ε και επιτυγχάνεται το όριο (οι ιδιοτιμές του πίνακα A είναι $\pm$ ε). Από την άλλη πλευρά , αν A= $\begin{pmatrix} 0 & \varepsilon \\ \varepsilon & 1 \end{pmatrix}$ τότε για M=0 το υπολειμματικό όριο θα είναι το ίδιο αλλά η μικρότερη ιδιοτιμή του πίνακα A θα είναι (-ε^2).

Σε αυτό το σημείο είναι χρήσιμο να γίνουν δύο επισημάνσεις. Η πρώτη επισήμανση αφορά το γεγονός ότι οι ιδιοτιμές του πίνακα M=X^H A X ονομάζονται «προσεγγίσεις Rayleigh-Ritz» των ιδιοτιμών του πίνακα A. Η δεύτερη επισήμανση αναφέρεται στο γεγονός ότι, παρόλο που ο πίνακας X περιέχει κατά προσέγγιση ιδιοδιανύσματα, η μόνη απαίτηση είναι να ληφθεί μια ακριβής ιδιοτιμή ώστε το R να είναι μικρό. Στη πραγματικότητα, το μέρος της απόδειξης όπου αποδεικνύεται ότι οι ιδιοτιμές του M είναι ίδιες με αυτές του A+E μπορεί να μετατραπεί σε αλγόριθμο για τη λήψη κατά προσέγγιση ιδιοδιανυσμάτων από τον πίνακα X, μια διαδικασία που ονομάζεται «βελτίωση Rayleigh-Ritz».

Προσέγγιση με πίνακα χαμηλής τάξης.

Μια δεύτερη εφαρμογή του θεωρήματος του Mirsky είναι ο προσδιορισμός προσεγγίσεων χαμηλής τάξης σε έναν σταθερό πίνακα. Όπως παραπάνω, έστω Φ είναι μια συμμετρική συνάρτηση και $\|.\|_\Phi$ η αντίστοιχη αμετάβλητη νόρμα. Έστω ότι ο ερμιτιανός πίνακας X∈ C^{mxn} γράφεται ως X=UΣV^H, όπου $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_m \geq 0$. Το ζητούμενο είναι να βρεθεί ένας πίνακας Y τάξης όχι μεγαλύτερης από k που να είναι όσο το δυνατόν πιο κοντά στο πίνακα X στη Φ-νόρμα.

Πρώτον, έστω Υ οποιοσδήποτε πίνακας τάξης όχι μεγαλύτερης από k. Τότε οι μοναδικές τιμές του Υ είναι $\tau_1 \geq ... \geq \tau_k \geq ... \geq 0 = ... = 0$, δηλαδή οι τελευταίες m-k μοναδικές τιμές είναι όλες μηδέν. Από το θεώρημα του Mirsky προκύπτει ότι

$\|\Upsilon - \mathrm{X}\|_\Phi \geq \Phi(\tau_1 - \sigma_1, \ldots, \tau_k - \sigma_k, \sigma_{k+1}, \ldots, \sigma_m) \geq \Phi(0, \ldots 0, \sigma_{k+1}, \ldots, \sigma_m)$. Με άλλα λόγια, οποιαδήποτε προσέγγιση τάξης όχι μεγαλύτερη από k πρέπει να αφαιρείται τουλάχιστον κατά $\Phi(0, \ldots 0, \sigma_{k+1}, \ldots, \sigma_m)$ από τον X. Τώρα έστω Σ_k = diag $(\sigma_1, ..., \sigma_k, 0, \ldots, 0)$ και $\mathrm{X}_k = \mathrm{U}\, \Sigma_k V^H$. Τότε επαληθεύεται εύκολα ότι ο X_k έχει τάξη όχι μεγαλύτερη από το k και $\|\Upsilon - \mathrm{X}\|_\Phi = \Phi(0, \ldots 0, \sigma_{k+1}, \ldots, \sigma_m)$. Έτσι αποδείχθηκε το ακόλουθο προσεγγιστικό θεώρημα.

ΘΕΩΡΗΜΑ Mirsky – Schmidt. Έστω ερμιτιανός πίνακας X με, $\mathrm{X} = \mathrm{U}\Sigma V^H$ όπου $\sigma_1 \geq \sigma_2 \geq ... \geq \sigma_m \geq 0$. Έστω Φ συμμετρική συνάρτηση και $\|.\|_\Phi$ αντίστοιχη μοναδιαία αμετάβλητη νόρμα. Εάν το Υ είναι ένας πίνακας τάξης μικρότερης ή ίσης του k, τότε $\|\Upsilon - \mathrm{X}\|_\Phi = \Phi(0, \ldots 0, \sigma_{k+1}, \ldots, \sigma_m)$. Επιπλέον επιτυγχάνεται ισότητα για τον πίνακα που ορίζεται από τις σχέσεις Σ_k = diag $(\sigma_1, ..., \sigma_k, 0, \ldots, 0)$ και $\mathrm{X}_k = \mathrm{U}\, \Sigma_k V^H$.

Έχουν ήδη αναφερθεί παραπάνω οι προσεγγίσεις Rayleigh-Ritz στις ιδιοτιμές ενός πίνακα.Α. Τόσο ο Rayleigh όσο και ο Ritz ασχολήθηκαν με την προσέγγιση των ιδιοτιμών ενός άπειρου τελεστή αντικαθιστώντας τον με ένα πρόβλημα ιδιοτιμής πίνακα. Ο Rayleigh βρήκε τα συστήματα δόνησης των φυσικών συχνοτήτων περιορίζοντας τους βαθμούς ελευθερίας τους σε συγκεκριμένο αριθμό τρόπων λειτουργίας, οι οποίοι θα επιλεγούν για να τονίσουν τη θεμελιώδη συχνότητα. Κανένας όμως από τους δύο δεν έδωσε επίσημη αιτιολόγηση για τη μέθοδό του. Στη βιβλιογραφία είναι σύνηθες να αποκαλείται η προσέγγιση ιδιοδιανυσμάτων που λαμβάνεται από τους πίνακες M και X ως "διανύσματα Ritz", αν και ο ίδιος ο Ritz είχε πει ότι δεν ήταν σε θέση να καθορίσει σε απόλυτο βαθμό τη σύγκλιση τους.

Μια άλλη σημαντική παρατήρηση είναι ότι όταν ένας Ερμιτιανός πίνακας διαταράσσεται τυχαία, μια πολλαπλή ιδιοτιμή θα τείνει να διασπαστεί σε άλλες πιο απλές ιδιοτιμές.

ΑΣΚΗΣΗ.

Έστω A και B ερμιτιανοί πίνακες με τον B να είναι τάξης k. Τότε η μεγαλύτερη ιδιοτιμή του πίνακα (A-B) δεν είναι μικρότερη από την (k+1)-η μεγαλύτερη ιδιοτιμή του A.

Λύση. Έστω $\lambda_1 \geq \ldots \geq \lambda_n$ οι ιδιοτιμές του ερμιτιανού πίνακα A και $\mu_1 \geq \ldots \geq \mu_k$ οι ιδιοτιμές του ερμιτιανού πίνακα B. Οι ιδιοτιμές του πίνακα (A-B) θα είναι $\nu_1, \ldots, \nu_n$ όπου ν_i = λ_i - μ_i για i= 1,2,…,k και ν_i = λ_i για i > k. Έστω (λ_{k+1}) η (k+1)-η μεγαλύτερη ιδιοτιμή του A. Τότε θα ισχύει ότι $(\nu_{k+1}$ = λ_{k+1} - $\mu_{k+1})$ και αφού ο B είναι θετικά ορισμένος θα πρέπει $(\mu_k$ > $\mu_{k+1})$. Από τις παραπάνω ανισότητες προκύπτει ότι $(\nu_{k+1}$> λ_{k+1} - $\mu_k)$. Η ν_{k+1} είναι η (k+1)-η μεγαλύτερη ιδιοτιμή του (A-B) και λ_1 η μεγαλύτερη ιδιοτιμή του A. Οπότε η μεγαλύτερη ιδιοτιμή του (A-B) θα είναι τουλάχιστον η $(\lambda_1$ - $\mu_1)$. Συνεπάγεται ότι $(\lambda_1$ - $\mu_1 >$ $\lambda_{k+1})$. Έτσι αποδείχθηκε το ζητούμενο.

ΑΣΚΗΣΗ.

Έστω ο πίνακας A = $\begin{pmatrix} B & c \\ c^H & \delta \end{pmatrix}$. Να δειχθεί ότι υπάρχει μια ιδιοτιμή (λ) του πίνακα A που να ικανοποιεί τη σχέση $|\lambda - \delta| \leq \|c\|_2$. Ο πίνακας B είναι ερμιτιανός.

Λύση. Δημιουργείται ο πίνακας A – λI = $\begin{pmatrix} B-\lambda & c \\ c^H & \delta-\lambda \end{pmatrix}$και από αυτόν υπολογίζεται το χαρακτηριστικό πολυώνυμο p(λ). Τότε p(λ) = $(B-\lambda)(\,\delta-\lambda)$ –

$c \ c^H$. Προκειμένου να βρεθούν οι ιδιοτιμές (λ) τίθεται p(λ) = 0. Από το θεώρημα Gerschgorin θα ισχύει ότι $|λ - δ| \leq c \ c^H$, δηλαδή $|λ - δ| \leq \|c\|_2$.

ΚΕΦΑΛΑΙΟ ΠΕΜΠΤΟ

<u>Ανάδειξη επιπλέον αποτελεσμάτων.</u>

Στη συνέχεια θα αναδειχθούν ορισμένα χρήσιμα αποτελέσματα.

Αρχικά θα αντιμετωπιστεί το πρόβλημα των μη ερμιτιανών διαταραχών των ερμιτιανών πινάκων και στη συνέχεια θα αντιμετωπιστεί η διαταραχή ιδιοτιμών πινάκων που είναι παρόμοιοι με ερμιτιανούς ή κανονικούς πίνακες.

Μη ερμιτιανές διαταραχές.

Θα μελετηθεί η μη- ερμιτιανή διαταραχή των ερμιτιανών πινάκων που αποδίδεται στον Kahan. Θα υποτεθεί ότι ο A είναι ένας ερμιτιανός πίνακας με ιδιοτιμές $\lambda_1 \geq \cdots \geq \lambda_n$ $ενώ$ θα υποτεθεί επιπλέον ότι ο $\tilde{A}$ είναι μια μη- ερμιτιανή διαταραχή του A, δηλαδή ο E=$\tilde{A}$ - A δεν είναι ερμιτιανός. Οι ιδιοτιμές του $\tilde{A}$, που μπορεί να είναι μιγαδικές, θα γράφονται (μ_κ+iν_κ) όπου $\mu_1 \geq \cdots \geq \mu_n$. Τέλος θα γραφτεί $E_R = \frac{E+E^H}{2}$ και $E_I = \frac{E-E^H}{2i} = \frac{\tilde{A}-\tilde{A}^H}{2i}$ για τα «πραγματικά» και «φανταστικά» μέρη του E.

Μπορεί να επαληθευτεί άμεσα ότι $\|\mathrm{E}\|_F^2 = \|E_R\|_F^2 + \|E_I\|_F^2$.

ΘΕΩΡΗΜΑ (Wilkinson).

Έστω D_k = { μ + iν: $|$μ + iν - $\lambda_k| \leq \|E\|_2$ και $|\nu| \leq \|E_I\|_2$}. Τότε θα ισχύει λ($\widetilde{A}$) $\subset U_{k=1}^{n}$ D_k.

Απόδειξη. Έχει δειχθεί ότι για κάθε (μ + iν) ∈ λ($\widetilde{A}$) θα υπάρχει μια ιδιοτιμή (λ_k) τέτοια ώστε $|$μ + iν - $\lambda_k| \leq \|E\|_2$. Απομένει τώρα να αποδειχθεί ότι $|\nu| \leq \|E_I\|_2$.

Έστω x ένα κανονικοποιημένο ιδιοδιάνυσμα του $\tilde{A}$ που αντιστοιχεί σε μ + iν, δηλ. $\tilde{A}$ = (μ + iν)x , με $\|\mathrm{x}\|_2$=1. Τότε θα ισχύουν τα εξής : (x^H $\tilde{A}$ x = μ + iν) και (x^H $\tilde{A}^H$ x = μ – iν).

Τότε θα συνεπάγεται ότι: $x^H \ E_I$ x = $\frac{x^H(\tilde{A}-\tilde{A}^H)x}{2i}$ = ν, οπότε συμπεραίνεται ότι $|v| \leq \|E_I\|_2$.

Εάν μία από τις περιοχές D_k είναι απομονωμένη από τις υπόλοιπες, θα περιέχει μόνο μία ιδιοτιμή, την $(\mu_\kappa$+i$\nu_\kappa)$ η οποία θα είναι υποχρεωτικά πραγματική. Έτσι, το θεώρημα πρωτοτυπεί μόνο για ομάδες ιδιοτιμών των οποίων οι περιοχές αλληλεπικαλύπτονται. Συγκεκριμένα, εάν οι περιοχές $D_k, \ldots, D_{k+m-1}$ επικαλύπτονται, τότε θα περιέχουν ακριβώς m ιδιοτιμές του $\tilde{A}$, τις μ_κ+ iν_κ, …, μ_{k+m-1}+ iν_{k+m-1}. Οι περιοχές αυτές είναι δίσκοι που έχουν τμηθεί στο επάνω και στο κάτω τμήμα τους από τις οριζόντιες γραμμές I(z) = $\pm \|E_I\|_2$. Όσο η διαταραχή γίνεται όλο και περισσότερο ερμιτιανή, οι συγκεκριμένες γραμμές πλησιάζουν η μία την άλλη, περιορίζοντας με αυτό τον τρόπο τα μεγέθη των φανταστικών τμημάτων των ιδιοτιμών του $\tilde{A}$.

Παρακάτω παρουσιάζεται μια άλλη εκδοχή του θεωρήματος που παραπέμπει στο αντίστοιχο θεώρημα των Hoffman-Wielandt.

ΘΕΩΡΗΜΑ.

Δεδομένου ότι $\sqrt{\sum_{k=1}^{n} \nu_k^2} \leq \|E_I\|_F$ και

$\sqrt{\sum_{k=1}^{n} (\mu_k - \lambda_k)^2)} \leq \|E_R\|_F + \sqrt{\|E_I\|_F^2 - \sum_{k=1}^{n} \nu_k^2}$, τότε

θα ισχύει ότι $\sqrt{\sum_{k=1}^{n} |(\mu_k + i\nu_\kappa) - \lambda_k|^2} \leq \sqrt{2}\ \|E\|_F$.

Απόδειξη. Χρησιμοποιώντας τη μορφή Schur του πίνακα $\tilde{A}$, είναι δυνατόν να υποτεθεί ότι $\tilde{A}$ = M + iN + R, όπου M = diag$(\mu_1, \ldots, \mu_n), N =$ diag$(\nu_1, \ldots, \nu_n)$ $\kappa\alpha\iota$ R ένας άνω τριγωνικός πίνακας. Τότε A + E_R = M + $\frac{R+R^H}{2}$ και

$E_I =$ N + $\frac{R-R^H}{2i}$. Αφού τα N και $\frac{R-R^H}{2i}$ είναι μη συνδεδεμένα σύνολα που αποτελούνται από μη μηδενικά στοιχεία, θα ισχύει ότι

$\|E_I\|_F^2 = \|N\|_F^2 + \left\|\frac{R-R^H}{2i}\right\|_F^2 = \|N\|_F^2 + \frac{1}{\sqrt{2}} \|R\|_F^2 \geq \|N\|_F^2 = \sum_{k=1}^{n} \nu_k^2$. H τελευταία σχέση επαληθεύεται από τα δεδομένα.

Εφόσον είναι γνωστό ότι οι πίνακες Α και Μ είναι ερμιτιανοί τότε : $\sqrt{\sum_{k=1}^{n} (\mu_k - \lambda_k)^2)} \leq \|M - A\|_F = \left\|E_R - \frac{R+R^H}{2}\right\|_F$
$\leq \|E_R\|_F + \left\|\frac{R+R^H}{2}\right\|_F = \|E_R\|_F + \frac{1}{\sqrt{2}} \|R\|_F^2 = \|E_R\|_F + (\|E_I\|_F^2 - \|N\|_F^2)^{1/2} = \|E_R\|_F + \sqrt{\|E_I\|_F^2 - \sum_{k=1}^{n} \nu_k^2}$. Η σχέση αυτή ισχύει εξ υποθέσεως.

Προκειμένου να επαληθευτεί η προς απόδειξη σχέση ακολουθείται η παρακάτω σειρά πράξεων:

$\sum_{k=1}^{n} |(\mu_k + i\nu_\kappa) - \lambda|^2 = \sum_{k=1}^{n} (\mu_k - \lambda_k)^2 + \sum_{k=1}^{n} \nu_k^2 \leq (\|E_R\|_F + \sqrt{\|E_I\|_F^2 - \sum_{k=1}^{n} \nu_k^2}\ )^2 + \sum_{k=1}^{n} \nu_k^2 = \|E_R\|_F^2 +$
$+ 2\, \|E_R\|_F \sqrt{\|E_I\|_F^2 - \sum_{k=1}^{n} \nu_k^2}\ ) + \|E_I\|_F^2 \leq (\|E_R\|_F + \|E_I\|_F)^2 \leq 2\, (\|E_R\|_F + \|E_I\|_F)^2 = 2\, \|E\|_F^2$.

ΟΡΙΑ ΟΜΟΙΟΤΗΤΑΣ.

Από το θεώρημα Bauer –Fike υπό την υπόθεση ότι ο πίνακας Α είναι διαγωνοποιήσιμος (δηλαδή ισχύει X^{-1}AX=Λ, όπου ο Λ είναι διαγώνιος πίνακας) το $sv_A(\widetilde{A}) \leq \|X^{-1}\text{ E X}\|$, που προέρχεται από τον διαγωνοποιημένο πίνακα Α, εξαρτάται από τη συνθήκη του μετασχηματισμού.

Στη συνέχεια θα υποτεθεί ότι οι πίνακες Α και $\widetilde{A}$ μπορούν α αναχθούν είτε σε μετασχηματισμούς ομοιότητας ερμιτιανών πινάκων είτε σε μετασχηματισμούς κανονικών πινάκων και να ληφθούν αντίστοιχα όρια διαταραχής.

Αρχικά θα αναλυθεί η περίπτωση ερμιτιανού πίνακα. Το βασικό αποτέλεσμα στηρίζεται στο παρακάτω λήμμα: έστω Η και Κ είναι δύο nxn ερμιτιανοί πίνακες και Σ = diag($\sigma_1, \ldots, \sigma_n$) με $\sigma_1 \geq \ldots \geq \sigma_n \geq 0$. Τότε $\|\mathrm{H}\Sigma - \Sigma\mathrm{K}\|_2 \geq \sigma_n \|\mathrm{H} - \mathrm{K}\|_2$.

Το συγκεκριμένο λήμμα αποδεικνύεται ως εξής: έστω (λ) είναι η ιδιοτιμή του πίνακα (H – K) με τη μεγαλύτερη απόλυτη τιμή. Αυτό σημαίνει ότι $\|\mathrm{H} - \mathrm{K}\|_2 = |\lambda|$. Έστω (x) το αντίστοιχο κανονικοποιημένο ιδιοδιάνυσμα. Τότε x^H $(\mathrm{H}\,\Sigma - \Sigma\,\mathrm{K}\,)$ x = x^H $(\mathrm{H} - \mathrm{K})$ Σx + x^H (ΚΣ – ΣΚ) x =

= λx^H Σx + iτ, όπου (τ) πραγματικός αριθμός. Στη συγκεκριμένη απόδειξη θεωρείται ότι ο πίνακας ΚΣ – ΣΚ είναι ερμιτιανός. Ως εκ τούτου

$\|\mathrm{H}\,\Sigma - \Sigma\,\mathrm{K}\|_2 = max_{\|\mathrm{v}\|_2}^{\|\mathrm{u}\|_2} \; |u^H\ (\mathrm{H}\,\Sigma - \Sigma\,\mathrm{K})\ \mathrm{u}| \;\geq\; max_{\|\mathrm{u}\|_2} \; |u^H\ (\mathrm{H}\,\Sigma - \Sigma\,\mathrm{K})$ $\mathrm{u}| \geq |x^H\ (\mathrm{H}\,\Sigma - \Sigma\,\mathrm{K})\ \mathrm{x}| \geq |\lambda|\ |x^H\ \Sigma\ \mathrm{x}| = \|H - K\|_2\ |x^H\ \Sigma\ \mathrm{x}| \geq$

$\geq \sigma_n \|\mathrm{H} - \mathrm{K}\|_2.$

ΘΕΩΡΗΜΑ.

Έστω A, $\tilde{A} \in C^{nxn}$ και έστω ότι υπάρχουν δύο μη- κανονικοί πίνακες P, Q τέτοιοι ώστε οι πίνακες P^{-1} A P και $\mathrm{Q}^{-1}\tilde{A}$ Q να είναι ερμιτιανοί. Στη συνέχεια, έστω οι ιδιοτιμές των A και $\tilde{A}$ (που είναι υποχρεωτικά πραγματικοί αριθμοί) είναι $\lambda_1 \geq\ldots\geq \lambda_n$ και $\tilde{\lambda}_1 \geq\ldots\geq \tilde{\lambda}_n$. Τότε θα ισχύει ότι

$|\tilde{\lambda}_i - \lambda_i| \leq \kappa_2$ (P) κ_2 (Q) $\|\tilde{A} - \mathrm{A}\|_2$, i = 1, 2, …, n.

Απόδειξη. Έστω $A_* = \mathrm{P}^{-1}$ A P και $\tilde{A}_* = \mathrm{Q}^{-1}$ A Q. Τότε

$\|\tilde{A} - \mathrm{A}\|_2 = \|\mathrm{Q}^{-1}\,\tilde{A}\ \mathrm{Q} - \mathrm{P}^{-1}\,\mathrm{A}\,\mathrm{P}\|_2 = \|\mathrm{Q}\ (\tilde{A}_*\mathrm{Q}^{-1}\,\mathrm{P} - \mathrm{Q}^{-1}\ \mathrm{P}A_*)\mathrm{P}^{-1}\|_2 \geq$ $\|\mathrm{Q}^{-1}\|_2^{-1}\ \|\mathrm{P}\|_2^{-1}\ \|\tilde{A}_*\,(\mathrm{Q}^{-1}\,\mathrm{P}) - (\mathrm{Q}^{-1}\,\mathrm{P})A_*\|_2.$

Θεωρώντας UΣV^H ως μετατροπή της μοναδικής τιμής του Q^{-1} P και με το σ_n να υποδηλώνει τη μικρότερη διαγώνιο του Σ, θα ισχύει με βάση το προαναφερθέν λήμμα ότι

$\|\tilde{A}_*\,(\mathrm{Q}^{-1}\,\mathrm{P}) - (\mathrm{Q}^{-1}\,\mathrm{P})A_*\|_2 = \|(U^H\,\tilde{A}_*\,\mathrm{U})\Sigma - \Sigma(V^H A_* \mathrm{V})\|_2 \geq$

$\geq \sigma_n\ \|U^H\,\tilde{A}_*\,\mathrm{U} - V^H A_* \mathrm{V}\|_2 \geq \sigma_n\ |\tilde{\lambda}_i - \lambda_i|$, i = 1, 2, …, n. Στη συνέχεια

$|\tilde{\lambda}_i - \lambda_i| \leq\leq \sigma_n^{\ -1}\ \|P\|_2\ \|Q^{-1}\|_2\ \|\tilde{A} - \mathrm{A}\|_2$, i=1, 2, …, n.

Τώρα $\sigma_n{}^{-1}$ = $\left\|(Q^{-1}P)^{-1}\right\|_2 \leq \left\|P^{-1}\right\|_2 \left\|Q\right\|_2$. Με συνδυασμό των δύο τελευταίων ανισώσεων αποδεικνύεται η ζητούμενη σχέση.

Το επόμενο θεώρημα είναι παρόμοιο με το προηγούμενο (στη βιβλιογραφία αναφέρεται ως θεώρημα των Sun – Zhang) και αφορά πίνακες που μπορούν να μετασχηματιστούν σε αντίστοιχους κανονικούς .

ΘΕΩΡΗΜΑ.

Έστω A, $\tilde{A} \in C^{nxn}$ και έστω ότι υπάρχουν δύο μη- κανονικοί πίνακες P, Q τέτοιοι ώστε οι πίνακες P^{-1} A P και $\mathrm{Q}^{-1}\tilde{A}$ Q να είναι κανονικοί. Τότε θα ισχύει ότι md_2 (A, $\tilde{A}$) $\leq$ κ(P)κ(Q) $\left\|\tilde{A} - \mathrm{A}\right\|_F$.

Απόδειξη. Εάν καταστεί δυνατόν να καθοριστεί ένα αντίστοιχο λήμμα για κανονικούς πίνακες με το λήμμα που προαναφέρθηκ,ε τότε η προς απόδειξη σχέση εξάγεται κατ' αναλογία. Συγκεκριμένα πρέπει να αποδειχθεί ότι αν τα M και N είναι κανονικοί πίνακες και Σ = diag($\sigma_{1,\ldots,}\sigma_n$), με $\sigma_1 \geq\ldots\geq \sigma_n$, τότε $\left\|M\Sigma - \Sigma N\right\|_F \geq \sigma_n\left\|M - N\right\|_F$. Προκειμένου να επιτευχθεί αυτό, τίθεται ως

δ=$\left\|M\Sigma - \Sigma N\right\|_F^{\,2}$ -$\sigma_n{}^2 \left\|M - N\right\|_F^{\,2}$ και Ω = Σ - σ_nI.

Προφανώς, τα διαγώνια στοιχεία του Ω είναι μη αρνητικά. Εξάλλου, δ = $\left\|M(\Omega + \sigma_n I) - (\Omega + \sigma_n I)\,N\right\|_F^{\,2}$ $\sigma_n{}^2$ $\left\|M - -N\right\|_F^{\,2}$ = $\left\|M\Omega - \Omega N + \sigma_n(M - N\right\|_F^{\,2} - \sigma_n{}^2 \left\|M - N\right\|_F^{\,2}$ = $\left\|M\Omega - \Omega N\right\|_F^{\,2}$ + 2 σ_n R {trace [$(M\Omega - \Omega N)^H$ $(M - N)$]} = $\left\|M\Omega - \Omega N\right\|_F^{\,2}$ + σ_n trace{Ω[$(M - N)^H$ $(M - N)$ + $(M - N)$ [$(M - N)^H$]} $\geq$ 0, που είναι η ζητούμενη ανισότητα.

Σύμφωνα με τον Kahan, το θεώρημα του Wilkinson, παρότι αναφέρεται σε μη συμμετρικές διαταραχές, δεν προσδιορίζει όρια για τα φανταστικά μέρη. Εκτός από τα αποτελέσματα που προαναφέρθηκαν, ο Kahan απέδειξε ότι η αντίστοιχη απόσταση της μη ερμιτιανής διαταραχής είναι ανάλογη με το $logn\|E\|_2$.

ΕΦΑΡΜΟΓΗ.

Εφόσον λ(Z) είναι πραγματική τιμή τότε $\|Z - Z^H\|_2$ ≤(0.038+log_2n) $\|Z + Z^H\|_2$. Με δεδομένη τη παραπάνω πρόταση να δειχθεί ότι αν ο πίνακας A είναι ερμιτιανός τότε md(A, $\tilde{A}$) ≤ $\|E_R\|_2$ + ≤ (0.038+log_2 n) $\|E_I\|_2$.

Απόδειξη. Έστω A ένας ερμιτιανός πίνακας και $\tilde{A}$ = A + E όπου E είναι μια μικρή απόκλιση από το πίνακα A. Τα E_R και E_I αντιπροσωπεύουν το πραγματικό και το φανταστικό μέρος του E, δηλαδή E = E_R + iE_I . Επιπλέον $\|E_R\|_2 \leq \|E\|_2$ και $\|E_I\|_2 \leq \|E\|_2$. Τίθεται Z = E_R + iE_I οπότε $\|E\|_2$ = $\|Z\|_2$. Στη συνέχεια θα γίνουν οι υπολογισμοί των ποσοτήτων $\|Z + Z^H\|_2$ και $\|Z - Z^H\|_2$.

1. $\|Z + Z^H\|_2$ = $\|$ E_R + iE_I + (E_R + iE_I)^H $\|$ = $\|E_R$ + iE_I + + E_R^H + i $E_I^H\|_2$ = $\|(E_R + E_R^H) +$ i$(E_I + E_I^H)\|_2$. Θεωρώντας τους πίνακες E_R και E_I ως πραγματικούς έπεται ότι $\|E_R + E_R^H\|_2 \leq 2 \|E_R\|_2$ και $\|E_I + E_I^H\|_2 \leq 2 \|E_I\|_2$. Άρα $\|Z + Z^H\|_2 \leq 2 \|E_R\|_2$ + 2 $\|E_I\|_2$.

2. $\|Z - Z^H\|_2$ ≤(0.038+log_2n) $\|Z + Z^H\|_2$ ≤
(0.038+log_2n) (2 $\|E_R\|_2$ + 2 $\|E_I\|_2$). Στη συνέχεια $\|$ A − $\tilde{A}$ $\|_2$ = $\|$ A - (A + E) $\|$ = $\|$ E $\|_2$. Οπότε md(A, $\tilde{A}$) = $\|$ E $\|_2$ ≤ $\|E_R\|_2$ + $\|E_I\|_2$ ≤(0.038+log_2n) (2 $\|E_R\|_2$ + 2 $\|E_I\|_2$) ≤(0.038+log_2n) $\|Z + Z^H\|_2$ ==> md(A, $\tilde{A}$)≤ $\|E_R\|_2$ + (0.038+log_2n) $\|E_I\|_2$.

ΣΥΖΗΤΗΣΗ

Η θεωρία των διαταραχών των ιδιοτιμών ενός πίνακα, εφαρμόζεται σε διάφορους επιστημονικούς τομείς. Παρακάτω θα γίνει μια προσπάθεια να χρησιμοποιηθούν τα μαθηματικά εργαλεία της προαναφερθείσας θεωρίας προκειμένου να γίνουν σαφείς οι τρόποι εφαρμογής της.

Αρχικά θα εξεταστεί η διαδικασία επεξεργασίας ενός σήματος και πως επηρεάζεται από τις διαταραχές των ιδιοτιμών. Έστω ότι ο πίνακας

X = $\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$ αντιπροσωπεύει ένα απλό ασπρόμαυρο σήμα. Αυτό το σήμα μπορεί να αποτελείται από μια σειρά δειγμάτων, όπου κάθε δείγμα μπορεί να είναι μια φωτεινότητα ενός pixel ή η τιμή ενός ήχου.

Με την είσοδο θορύβου στο σήμα δημιουργείται ένας νέος πίνακας , έστω

Y = $\begin{pmatrix} -0.05 & 0.12 & -0.18 & -0.22 & 0.03 \\ 0.03 & 1.05 & 1.08 & 0.94 & 0.12 \\ -0.09 & 0.92 & -0.03 & 1.06 & -0.04 \\ 0.02 & 0.93 & 1.15 & 1.10 & -0.09 \\ 0.11 & 0.05 & -0.15 & -0.07 & -0.01 \end{pmatrix}$. Υπολογίζοντας τις ιδιοτιμές των δύο πινάκων X, Y θα βρεθούν οι διαφορές στις τιμές αυτών των ιδιοτιμών λόγω της εμφάνισης του θορύβου. Οι ιδιοτιμές είναι σημαντικός παράγοντας στην ανάλυση σημάτων και εικόνων γιατί παρέχουν πληροφορίες αναφορικά με τη δομή και τα χαρακτηριστικά του σήματος. Η διαταραχή στις ιδιοτιμές, λόγω του θορύβου, θα οδηγήσει σε παραμορφώσεις του σήματος και στη συνέχεια θα εξαχθούν εσφαλμένες πληροφορίες που θα επηρεάσουν, με αρνητικό τρόπο, την επεξεργασία και την ανάλυση του συγκεκριμένου σήματος.

Σε δεύτερο επίπεδο θα εξεταστεί η επίδραση των διαταραχών των ιδιοτιμών ενός πίνακα σε προβλήματα μηχανικής. Έστω ένα γραμμικό δυναμικό σύστημα που περιγράφεται από τις εξισώσεις κατάστασης $\dot{x}$ = Ax., όπου (x) είναι το διάνυσμα κατάστασης και (A) είναι ο πίνακας κατάστασης. Οι ιδιοτιμές του πίνακα A καθορίζουν τη συμπεριφορά του συστήματος χωρίς διαταραχές. Αν υποτεθεί ότι το σύστημα υπόκειται σε διαταραχές στις ιδιοτιμές του λόγω εξωτερικών παραγόντων (π.χ θόρυβος, αστάθειες και αλλαγές στη δομή του συστήματος). Οι διαταραχές αυτές, πιθανότατα, θα μεταβάλλουν τις ιδιοτιμές του πίνακα A με αποτέλεσμα τα επηρεαστεί η κανονική λειτουργία του συστήματος.

Επεκτείνοντας το παράδειγμα, έστω ένα σύστημα μάζας – ελατηρίου – αποσβεστήρα. Το σύστημα αυτό περιγράφεται από τις εξισώσεις κίνησης
$m\ddot{x} + c\dot{x} + kx = 0$, όπου m = μάζα, c = ο συντελεστής απόσβεσης και k = η σταθερά του ελατηρίου. Το πρόβλημα εστιάζεται στην ανάλυση της συμπεριφοράς του συστήματος υπό συνθήκες διαφορετικών τιμών των μεγεθών m, c και k καθώς και υπό της επίδραση διαφόρων εξωτερικών διαταραχών. Το συγκεκριμένο σύστημα μπορεί να υποστεί διαταραχές στις ιδιοτιμές του π.χ λόγω της αλλαγής της σταθεράς απόσβεσης λόγω φθοράς του αποσβεστήρα. Η συγκεκριμένη μεταβολή θα διαφοροποιήσει τη δυναμική συμπεριφορά του συστήματος γιατί θα επέλθουν αλλαγές σε σημαντικές παραμέτρους όπως οι συχνότητες των αποκρίσεων, η ικανότητα απόσβεσης και η σταθερότητα.

Ένα άλλο παράδειγμα αναφοράς στο τομές της μηχανικής, είναι η ανάλυση ενός συστήματος ελέγχου που χρησιμοποιεί αισθητήρες και ελεγκτές προκειμένου το σύστημα να διατηρείται σε μια επιθυμητή κατάσταση. Οι διαταραχές των ιδιοτιμών που αφορούν τους πίνακες των μηχανικών μέσων του συστήματος όπως οι αισθητήρες και οι ελεγκτές, εξαιτίας φθοράς, θορύβου ή/και αλλαγών στην απόκριση των υποσυστημάτων θα επηρεάσουν τόσο την ακρίβεια όσο και τη σταθερότητα του συστήματος ελέγχου.

Περνώντας σε ένα άλλο πεδίο εφαρμογής, θα γίνει μνεία για διαταραχές ιδιοτιμών που αφορούν ανάλυση δεδομένων. Έστω ένας πίνακας δεδομένων που περιέχει μετρήσεις από διάφορα αισθητήρια όργανα. Οι στήλες αυτού του πίνακα, ας υποτεθεί ότι, περιέχουν τιμές από διαφορετικές μεταβλητές (θερμοκρασία, πίεση, ταχύτητα κλπ) ενώ οι γραμμές αντιστοιχούν σε διαφορετικά σημεία χρονικής ή χωρικής μέτρησης. Η ανάλυση τέτοιου πίνακα γίνεται προκειμένου να εντοπιστούν πρότυπα, τάσεις και συσχετίσεις μεταξύ των διαφορετικών αυτών μεταβλητών.

Ας υποτεθεί ότι ο πίνακας δεδομένων εκφράζει τις αναλύσεις της απόδοσης ενός μηχανικού συστήματος όπως είναι ένα αεροπλάνο. Κάθε γραμμή του πίνακα, υποθετικά, αντιπροσωπεύει μια διαφορετική πτήση και κάθε στήλη μια διαφορετική μετρική όπως π.χ το υψόμετρο, η κατανάλωση καυσίμων και η ταχύτητα. Οι ιδιοτιμές του πίνακα δεδομένων έχουν την ιδιότητα της ανάδειξης των προτύπων και των συσχετίσεων μεταξύ αυτών των μετρικών. Η ανάλυση των ιδιοτιμών μπορεί να απαντήσει σε ερωτήματα όπως ποιες μετρικές είναι σημαντικότερες για να περιγράψουν την διακύμανση των δεδομένων ή ποιες παράμετροι επηρεάζουν περισσότερο την απόδοση του αεροπλάνου. Οι διαταραχές των ιδιοτιμών του πίνακα δεδομένων είναι δυνατό να αντιπροσωπεύουν ανωμαλίες και αστάθειες των δεδομένων που επηρεάζουν τυην αξιοπιστία της ανάλυσης των συγκεκριμένων δεδομένων.

ΣΥΜΠΕΡΑΣΜΑΤΑ

Στην εργασία που εκπονήθηκε, παρουσιάστηκαν πολλά θεωρήματα που το καθένα συμβάλλει με το δικό του τρόπο στην ανάπτυξη της θεωρίας της ανάλυσης των διαταραχών των ιδιοτιμών. Υποκειμενικά πάντα, θα μπορούσε να ειπωθεί ότι μερικά από τα θεωρήματα αυτά είναι πιο σημαντικά από κάποια άλλα γιατί αποτελούν τη στέρεη βάση πάνω στην οποία εδραιώνεται η προαναφερθείσα θεωρία. Τέτοια θεωρήματα είναι του Gershgorin, του Mirsky, του Ostrowski, του Weyl και του Henrici.

Το θεώρημα του Gershgorin παρέχει μια γεωμετρική προσέγγιση που σχετίζεται με τον υπολογισμό των ιδιοτιμών του πίνακα. Συνεισφέρει στο περιορισμό του φάσματος όλων των δυνατών τιμών των ιδιοτιμών ορίζοντας περιοχές στο επίπεδο του φάσματος όπου είναι δυνατόν να βρεθούν αυτές οι τιμές. Ως θεώρημα είναι ιδιαίτερα χρήσιμό για τη γρήγορη εκτίμηση των ιδιοτιμών.

Το θεώρημα του Mirsky οδηγεί στον υπολογισμό μιας ανώτερης και μιας κατώτερης σειράς, σχετικά περιορισμένων, για τις ιδιοτιμές ενός πίνακα. Εκτιμά ακριβέστερα από το θεώρημα του Gershgorin τις ιδιοτιμές με αποτέλεσμα να υπάρχει καλύτερη αξιολόγηση των επιπτώσεων τους πάνω στα συστήματα.

Το θεώρημα του Ostrowski αποτελεί το εργαλείο παροχής ενός κριτηρίου σύγκλισης προκειμένου να προσδιοριστούν οι ιδιοτιμές ενός πίνακα. Το θεώρημα αυτό εστιάζει στον προσδιορισμό των ιδιοτιμών που σχετίζονται με σύγκλιση ή απόκλιση του συστήματος. Το αποτέλεσμα αυτό είναι ιδιαίτερα χρήσιμο προκειμένου να αξιολογηθεί η σταθερότητα και η συμπεριφορά του συστήματος κατά την εξέλιξη του χρόνου.

Το θεώρημα του Weyl εκτιμά τις ιδιοτιμές τις ιδιοτιμές του αθροίσματος δύο πινάκων με βάση τις ιδιοτιμές των πινάκων αυτών ξεχωριστά. Η εκτίμηση των ιδιοτιμών του αθροίσματος δύο πινάκων από τις ιδιοτιμές τους ξεχωριστά είναι αναγκαία σε περιπτώσεις ανάλυση πινάκων δεδομένων καθώς και σε προβλήματα της μηχανικής.

Το θεώρημα του Henrici αναφέρεται στη σχετική ακρίβεια των ιδιοτιμών ενός πίνακα σε σχέση με τους πιθανούς περιορισμούς που υπάρχουν στα στοιχεία του πίνακα. Αυτό έχει αντίκτυπο στην εκτίμηση της ευαισθησίας των ιδιοτιμών σε αλλαγές που υπόκεινται οι πίνακες όπως αλλαγές στις τιμές των στοιχείων τους καθώς και περιορισμό αυτών των τιμών.

Το θεώρημα Hoffman-Wielandt αναφέρεται στη συμπεριφορά των ιδιοτιμών όταν αυτός υφίσταται μικρές διαταραχές. Η συνεισφορά του συγκεκριμένου θεωρήματος στη θεωρία των διαταραχών των ιδιοτιμών αναλύεται σε τρείς συνιστώσες. Η πρώτη συνιστώσα σχετίζεται με τον καθορισμό της συμπεριφοράς των ιδιοτιμών και αποτελεί κρίσιμο εργαλείο για την κατανόηση της ευαισθησίας ενός συστήματος όταν αυτό υπόκειται σε μικρές αλλαγές. Η δεύτερη συνιστώσα αφορά τις πρακτικές εφαρμογές. Τομείς όπως ο έλεγχος ποιότητας, επεξεργασία σήματος και ανάλυση συστημάτων ελέγχου θεωρούνται πεδία αιχμής ειδικά σε θέματα καθορισμού της ευστάθειας όταν τα συστήματα τους υπόκεινται σε μικρές διαταραχές. Η παράμετρος της ευστάθειας αποτελεί από μόνη της τη τρίτη συνιστώσα συνεισφοράς του θεωρήματος στη θεωρία των διαταραχών των ιδιοτιμών αφού σχετίζεται με τις ιδιοτιμές του αντίστοιχου πίνακα.

Όλα τα παραπάνω θεωρήματα αποτελούν τα μαθηματικά εργαλεία της θεωρίας των διαταραχών των ιδιοτιμών ενός ή περισσότερων πινάκων που με τη σειρά της εφαρμόζεται σε μια σειρά επιστημονικών πεδίων, όπως έχει ήδη αναφερθεί.

Οι διαταραχές στις ιδιοτιμές επηρεάζουν σε σημαντικό βαθμό τη σταθερότητα και τη συμπεριφορά ενός συστήματος. Η εύρεση και η διαχείριση παρόμοιων διαταραχών οδηγεί στη διατήρηση της αξιοπιστίας και της απόδοσης του συστήματος.

Τα διάφορα θεωρήματα όπως αναλύθηκαν επιτρέπουν την αξιολόγηση του βαθμού επιρροής των διαταραχών στην απόδοση που είχε ή θα έχει κάποιο σύστημα.

Οι επιπτώσεις που έχουν οι διαταραχές των ιδιοτιμών εμφανίζονται σε κλάδους όπως είναι η μηχανική, η ανάλυση δεδομένων και η επεξεργασία σημάτων. Η πλήρης κατανόηση των εμφανιζόμενων επιπτώσεων είναι σημαντική προκειμένου να αναπτυχθούν αξιόπιστες και αποτελεσματικές μέθοδοι που θα διαχειρίζονται αυτές τις διαταραχές.

Η θεωρία των διαταραχών των ιδιοτιμών εφαρμόζεται πολλάκις τόσο στην έρευνα όσο και στην τεχνολογική ανάπτυξη. Οι βελτιώσεις που είναι δυνατόν να εφαρμοστούν σε διάφορους τομείς της τεχνολογίας και που θα βασίζονται στη θεωρία των διαταραχών των ιδιοτιμών, θα αναπτύξουν περαιτέρω τη συνεισφορά της επιστήμης στην ολοκλήρωση των επιδιώξεων της ανθρώπινης κοινωνίας αναφορικά με την ευημερίας αυτής. Απαραίτητη προϋπόθεση για μια τέτοια εξέλιξη είναι η συνεχής έρευνα που θα οδηγήσει σε καινοτομίες της τεχνολογίας.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. STEWART G.W, SUN JI-GUANG, MATRIX PERTURBATION THEORY, ACADEMIC PRESS INC., 1990.
2. ΨΑΡΡΑΚΟΣ ΠΑΝΑΓΙΩΤΗΣ, ΘΕΜΑΤΑ ΑΝΑΛΥΣΗΣ ΠΙΝΑΚΩΝ, ΕΚΔ. ΕΜΠ. 2020.
3. WILKINSON J.H, THE ALGEBRIC EIGENVALUE PROBLEM, OXFORD UNIVERSITY PRESS, NY. 1988.
4. ΚΑΔΙΑΝΑΚΗΣ Ν., ΚΑΡΑΝΑΣΙΟΣ Σ., ΓΡΑΜΜΙΚΗ ΑΛΓΕΒΡΑ-ΑΝΑΛΥΤΙΚΗ ΓΕΩΜΕΤΡΙΑ ΚΑΙ ΕΦΑΡΜΟΓΕΣ, ΕΚΔ. ΕΜΠ., ΑΘΗΝΑ 2003.
5. HORN R.A, JOHNSON C.R, MATRIX ANALYSIS, CAMBRIDGE UNIVERSITY PRESS, 1990.
6. ΧΡΥΣΑΚΗΣ Θ., ΓΡΑΜΜΙΚΗ ΑΛΓΕΒΡΑ –ΑΝΑΛΥΤΙΚΗ ΓΕΩΜΕΤΡΙΑ, ΑΘΗΝΑ 2013.