

ΕΘΝΙΚΟ ΜΕΤΣΟΒΙΟ ΠΟΛΥΤΕΧΝΕΙΟ
ΣΧΟΛΗ ΕΦΑΡΜΟΣΜΕΝΩΝ ΜΑΘΗΜΑΤΙΚΩΝ ΚΑΙ ΦΥΣΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΤΟΜΕΑΣ ΜΑΘΗΜΑΤΙΚΩΝ
2011

ΔΙΠΛΩΜΑΤΙΚΗ ΕΡΓΑΣΙΑ

ΠΡΟΒΛΗΜΑΤΑ ΥΠΟΛΟΓΙΣΜΟΥ
ΙΔΙΟΤΙΜΩΝ ΚΑΙ ΙΔΙΟΔΙΑΝΥΣΜΑΤΩΝ
ΣΕ ΣΥΜΜΕΤΡΙΚΟΥΣ ΠΙΝΑΚΕΣ
ΚΑΙ
ΕΦΑΡΜΟΓΕΣ ΣΕ ΤΡΙΔΙΑΓΩΝΙΟΥΣ
ΠΙΝΑΚΕΣ

Τσαμπαρδούκα Θεοδώρα

Επιβλέπων καθηγητής
Χρυσάφινος Κώστας

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ

ΚΕΦΑΛΑΙΟ 1	<i>Εισαγωγή</i>	9
ΚΕΦΑΛΑΙΟ 2:	<i>Το χαρακτηριστικό πολυώνυμο</i>	13
ΚΕΦΑΛΑΙΟ 3:	<i>Μέθοδος Jacobi</i>	17
ΚΕΦΑΛΑΙΟ 4:	<i>Τα θεωρήματα του Gerschgorin</i>	26
ΚΕΦΑΛΑΙΟ 5:	<i>Η μέθοδος Householder</i>	34
ΚΕΦΑΛΑΙΟ 6:	<i>Ιδιοτιμές τριδιαγώνιων πινάκων</i>	41
ΚΕΦΑΛΑΙΟ 7:	<i>Ο Αλγόριθμος QR</i>	
	<i>7.1 Η παραγοντοποίηση QR</i>	49
	<i>7.2 Ο ορισμός του αλγορίθμου QR</i>	51
ΚΕΦΑΛΑΙΟ 8:	<i>Αντίστροφη επανάληψη για τα ιδιοδιανύσματα</i>	
	<i>8.1 Η μέθοδος των δυνάμεων</i>	55
	<i>8.2 Αντίστροφη επανάληψη</i>	56
ΚΕΦΑΛΑΙΟ 9:	<i>Το πηλίκο Rayleigh</i>	62
ΚΕΦΑΛΑΙΟ 10:	<i>Ανάλυση της διαταραχής</i>	65
ΚΕΦΑΛΑΙΟ 11:	<i>Κώδικες Matlab και παραδείγματα</i>	69

ΠΑΡΑΡΤΗΜΑ

1.	<i>Βασικοί ορισμοί</i>	85
----	------------------------	----

ΒΙΒΛΙΟΓΡΑΦΙΑ

ΠΡΟΛΟΓΟΣ

Το πρόβλημα του υπολογισμού των ιδιοτιμών και των ιδιοδιανυσμάτων πινάκων προκύπτει σε πολλές περιοχές της εφαρμοσμένης επιστήμης. Ο υπολογισμός αυτών των χαρακτηριστικών μεγεθών ενός πίνακα είναι απαραίτητος σε πολλά προβλήματα που αφορούν τα Μαθηματικά, τη Φυσική, τη Μηχανική αλλά και την Τεχνολογία. Αρχικά, ιδιοτιμές και ιδιοδιανύσματα χρησιμοποιούμε στην επίλυση συνήθων και μερικών διαφορικών εξισώσεων. Επίσης, στην περιστροφή άκαμπτου σώματος στο χώρο απαιτείται υπολογισμός ιδιοτιμών που στην ουσία μας δίνουν τις ροπές αδράνειας των αξόνων. Επίσης, όταν μελετάμε ταλαντώσεις δυναμικά συνδεδεμένων σωμάτων προκύπτουν γραμμικά συστήματα για την επίλυση των οποίων είναι απαραίτητη η εύρεση ιδιοτιμών. Τέλος, οι ιδιοτιμές και τα ιδιοδιανύσματα παίζουν αποφασιστικό ρόλο σε προβλήματα συντονισμού που προκύπτουν σε ηλεκτρικά ή μηχανικά συστήματα όπως πχ οι φυσικές συχνότητες δονήσεων δοκών σε κατασκευές που στην πραγματικότητα είναι ιδιοτιμές ή οι συχνότητες λήψεως μιας ραδιοφωνικής εκπομπής που είναι επίσης ιδιοτιμές.

Σε αυτήν την εργασία θα ασχοληθούμε με μεθόδους που υπολογίζουν ιδιοτιμές και ιδιοδιανύσματα συμμετρικών πινάκων καθώς επίσης και μεθόδους που υπολογίζουν τα χαρακτηριστικά μεγέθη τριδιαγώνιων πινάκων. Πιο συγκεκριμένα θα δούμε μεθόδους που υπολογίζουν ιδιοτιμές και ιδιοδιανύσματα πλήρων συμμετρικών πινάκων. Θα δούμε, επίσης, τον μετασχηματισμό πλήρους συμμετρικού πίνακα σε τριδιαγώνια μορφή και τη χρήση μεθόδων που υπολογίζουν τις ιδιοτιμές αυτού του τριδιαγώνιου πίνακα γιατί πολλές φορές είναι πιο αποδοτικό να μετασχηματίσουμε τον πλήρη συμμετρικό πίνακα σε τριδιαγώνια μορφή και στην συνέχεια να βρούμε τις ιδιοτιμές του παρά να βρω απευθείας τις ιδιοτιμές του πλήρη συμμετρικού πίνακα. Τέλος, θα αναπτύξουμε μεθόδους οι οποίες υπολογίζουν μέγιστες ή ελάχιστες ιδιοτιμές όπως πχ η μέθοδος των δυνάμεων. Αναλυτικά:

Στο κεφάλαιο **1** κάνουμε μια μικρή εισαγωγή με κάποιους ορισμούς, κάποια παραδείγματα και ένα βασικό θεώρημα τα οποία μας βοηθούν να καταλάβουμε τον υπολογισμό ιδιοτιμών και ιδιοδιανυσμάτων και τη χρησιμότητά τους.

Στο κεφάλαιο **2** παρουσιάζονται δύο παραδείγματα που υπολογίζουν τα χαρακτηριστικά μεγέθη σε μικρού μεγέθους πίνακες χρησιμοποιώντας το χαρακτηριστικό πολυώνυμο και ένα παράδειγμα που δείχνει την ανάγκη ανάπτυξης προσεγγιστικών επαναληπτικών μεθόδων σε μεγάλου μεγέθους πίνακες που να μας υπολογίζουν ιδιοτιμές και ιδιοδιανύσματα αριθμητικά.

Στο κεφάλαιο **3** παρουσιάζεται η κλασική μέθοδος Jacobi που υπολογίζει ιδιοτιμές και ιδιοδιανύσματα σε συμμετρικούς πίνακες. Παρουσιάζεται αναλυτικά ο αλγόριθμος της μεθόδου αυτής και επίσης συζητείται εκτενώς η σύγκλιση της.

Στο κεφάλαιο **4** αναλύονται τα Θεωρήματα Gerschgorin τα οποία μας επιτρέπουν να προσδιορίζουμε την περιοχή στην οποία θα βρίσκονται οι ιδιοτιμές ενός πίνακα.

Στην ουσία μας βοηθούν ώστε να έχουμε μια πρώτη αρχική ιδέα για το που περίπου θα βρίσκονται οι ιδιοτιμές ενός πίνακα.

Στο κεφάλαιο **5** αναπτύσσεται η μέθοδος *Householder* η οποία μετασχηματίζει ένα πλήρη συμμετρικό πίνακα σε τριδιαγώνια μορφή. Στο κεφάλαιο αυτό περιέχονται κάποια πράγματα για τους πίνακες *Householder*, εξηγείται ο *Householder* μετασχηματισμός και εξηγείται αναλυτικά ο αλγόριθμος της μεθόδου.

Στο κεφάλαιο **6** αναπτύσσεται η μέθοδος που στηρίζεται στην ακολουθιακή ιδιότητα του *Sturm* η οποία υπολογίζει ιδιοτιμές τριδιαγώνιων συμμετρικών πινάκων οι οποίες είναι μεγαλύτερες από ένα δοθέν πραγματικό αριθμό θ . Η μέθοδος αυτή είναι πολύ ενδιαφέρουσα καθώς δεν μας υπολογίζει μόνο ιδιοτιμές αλλά και να μας δείχνει ποια από τις ιδιοτιμές είναι. Αυτό είναι πολύ σημαντικό γιατί σε πολλά πρακτικά προβλήματα δεν χρειάζεται να βρούμε όλες τις ιδιοτιμές ενός πίνακα αλλά κάποια ή κάποιες από αυτές πχ την μεγαλύτερη ή τη μικρότερη.

Στο κεφάλαιο **7** παραθέτεται η μέθοδος *QR* που στηρίζεται στην παραγοντοποίηση *QR* και υπολογίζει ιδιοτιμές σε συμμετρικούς τριδιαγώνιους πίνακες. Η μέθοδος αυτή εξηγείται αναλυτικά και επίσης παρουσιάζεται ο αλγόριθμος που την υλοποιεί.

Στο κεφάλαιο **8** παρουσιάζονται δύο μέθοδοι. Η πρώτη είναι η μέθοδος των δυνάμεων που υπολογίζει την μεγαλύτερη κατά μέτρο ιδιοτιμή και το αντίστοιχο ιδιοδιάνυσμα ενός πίνακα ενώ η δεύτερη είναι η αντίστροφη επανάληψη η οποία μας δίνει ιδιοτιμή και το αντίστοιχο ιδιοδιάνυσμα πίνακα όταν είναι διαθέσιμες μία καλή προσέγγιση της ιδιοτιμής αυτής και μια καλή προσέγγιση του αντίστοιχου ιδιοδιανύσματος.

Στο κεφάλαιο **9** αναπτύσσεται μια τεχνική, η οποία βασίζεται στο πηλίκο *Reyleigh*, που μας δίνει μια ακριβή προσέγγιση ιδιοτιμής ενός πίνακα όταν είναι διαθέσιμη μια ακριβή προσέγγιση του συσχετιζόμενου ιδιοδιανύσματος.

Στο κεφάλαιο **10** παρουσιάζουμε περιληπτικά κάποια Θεωρήματα τα οποία μας επιτρέπουν να έχουμε μια εκτίμηση για το πόσο επηρεάζονται οι ιδιοτιμές και τα ιδιοδιανύσματα ενός πίνακα από διαταραχές στα στοιχεία του.

Τέλος, στο κεφάλαιο **11** παραθέτουμε του κώδικες των παραπάνω μεθόδων σε *Matlab* καθώς και παραδείγματα για την κατανόησή τους.

Στο τέλος της εργασίας παραθέτουμε ένα παράρτημα με κάποιους ορισμούς και βασικές έννοιες για την κατανόηση της εργασίας καθώς και τη βιβλιογραφία που χρησιμοποιήθηκε.

Αρωγός μου στη μελέτη και τη συγγραφή της εργασίας αυτής ήταν ο κ. Χρυσάφινος Κώστας τον οποίο ευχαριστώ θερμά για την εξαιρετική συνεργασία και την πολύτιμη βοήθειά του.

1. ΕΙΣΑΓΩΓΗ

Το πρόβλημα των ιδιοτιμών και των ιδιοδιανυσμάτων σε συμμετρικούς πίνακες προκύπτει σε πολλές περιοχές της εφαρμοσμένης επιστήμης (όπως στην μηχανική, στις συνήθεις διαφορικές εξισώσεις κτλ). Η ορολογία της λέξης ιδιοτιμή (eigenvalue) προέρχεται από την γερμανική λέξη Eigenwert η οποία σημαίνει κατάλληλη ή χαρακτηριστική τιμή. Η ιδέα των ιδιοτιμών εμφανίζεται για πρώτη φορά σε ένα άρθρο που ασχολήθηκε με τα συστήματα γραμμικών διαφορικών εξισώσεων από τον Γάλλο μαθηματικό Jean le Rond d'Alembert στη διαδικασία της μελέτης πάνω στην κίνηση χορδής σε διάφορα σημεία του χώρου.

Μετά την μικρή αυτή εισαγωγή θα δώσουμε τον ορισμό της ιδιοτιμής και του ιδιοδιανύσματος ενός πίνακα.

ΟΡΙΣΜΟΣ 1.1: Έστω ότι έχουμε έναν πίνακα A $n \times n$ ($A \in \mathbb{R}^{n \times n}$). Ένας μιγαδικός αριθμός λ για τον οποίο η γραμμική εξίσωση

$$Ax = \lambda x \quad (1.1)$$

έχει μία λύση $x \in \mathbb{C}^n \setminus \{0\}$, ονομάζεται **ιδιοτιμή** του A και η λύση $x \in \mathbb{C}^n$ καλείται **ιδιοδιάνυσμα** του A (σε αντιστοιχία με το λ).•

ΠΑΡΑΤΗΡΗΣΗ: Πιο απλά θα μπορούσαμε να πούμε ότι ένας $n \times n$ πίνακας A έχει ακριβώς n , όχι κατ' ανάγκη διακεκριμένες, ιδιοτιμές $\lambda_i \in \mathbb{C}$, $i = 1, 2, \dots, n$ οι οποίες είναι ρίζες του χαρακτηριστικού πολυωνύμου:

$$p(\lambda) = \det(A - \lambda I) \quad (1.2)$$

Θεωρητικά οι ιδιοτιμές του πίνακα A υπολογίζονται από τις ρίζες του $p(\lambda)$. Τότε τα αντίστοιχα γραμμικά συστήματα μπορούν να επιλυθούν για τον υπολογισμό των ιδιοδιανυσμάτων που αντιστοιχούν σε κάθε ιδιοτιμή.

Παρακάτω παραθέτουμε δύο οικεία παραδείγματα για το πρόβλημα των ιδιοτιμών και των ιδιοδιανυσμάτων:

ΠΑΡΑΔΕΙΓΜΑ 1.2: Λαμβάνοντας υπόψη την περιστροφή ενός άκαμπτου σώματος $\Omega \subset \mathbb{R}^3$ (κίνηση σε 3 διαστάσεις) παίρνουμε τον αδρανειακό πίνακα που είναι ένας συμμετρικός πίνακας 3×3 όπως ο παρακάτω:

$$J = \begin{bmatrix} I_{xx} & -I_{xy} & -I_{xz} \\ -I_{yx} & I_{yy} & -I_{yz} \\ -I_{zx} & -I_{zy} & I_{zz} \end{bmatrix}$$

του οποίου τα διαγώνια στοιχεία μας δίνουν τη ροπή αδράνειας των αξόνων δηλαδή

$$I_{xx} = \int_{\Omega} (y^2 + z^2) d\Omega, \quad I_{yy} = \int_{\Omega} (z^2 + x^2) d\Omega, \quad I_{zz} = \int_{\Omega} (x^2 + y^2) d\Omega$$

ενώ τα υπόλοιπα στοιχεία του πίνακα (δηλαδή τα μη διαγώνια) ορίζονται από τα αντίστοιχα γινόμενα:

$$I_{xy} = I_{yx} = \int_{\Omega} xy d\Omega, \quad I_{yz} = I_{zy} = \int_{\Omega} yz d\Omega, \quad I_{zx} = I_{xz} = \int_{\Omega} xz d\Omega$$

Στο παραπάνω παράδειγμα, τα ιδιοδιανύσματα του αδρανειακού πίνακα είναι η κατεύθυνση των πρωτευόντων αξόνων της αδράνειας του σώματος, για τα οποία η ελεύθερη σταθερή περιστροφή είναι πιθανή, και οι ιδιοτιμές είναι οι ροπές αδράνειας στους άξονες. •

ΠΑΡΑΔΕΙΓΜΑ 1.3: Ένα άλλο παράδειγμα που χρησιμοποιεί ιδιοτιμές και ιδιοδιανύσματα προκύπτει από τη επίλυση συστημάτων από γραμμικές συνήθεις διαφορικές εξισώσεις της μορφής:

$$\frac{dx}{dt} = Ax$$

όπου x είναι ένα διάνυσμα από n στοιχεία καθένα από τα οποία είναι συναρτήσεις της ανεξάρτητης μεταβλητής t και ο A είναι ένας πίνακας $n \times n$ τα στοιχεία του οποίου είναι σταθερές. Στην περίπτωση που ο πίνακας A είναι διαγώνιος (δηλαδή πίνακας του οποίου τα μόνα μη μηδενικά στοιχεία βρίσκονται πάνω στην κύρια διαγώνιο του) με διαγώνια στοιχεία $a_{ii} = \lambda_i, i = 1, 2, \dots, n$, τότε η λύση του συστήματος είναι άμεση, ενώ κάθε μία από τις εξισώσεις μπορεί να λυθεί χωριστά δίνοντας

$$x_i(t) = x_i(0) \exp(\lambda_i t), \quad i = 1, 2, \dots, n$$

Στην περίπτωση, τώρα, που ο πίνακας A δεν είναι διαγώνιος, υποθέτουμε ότι μπορούμε να βρούμε έναν μη μοναδικό πίνακα M τέτοιον ώστε

$$M^{-1}AM = D$$

όπου D ένας διαγώνιος πίνακας. Έστω ότι $y = M^{-1}x$ τότε εύκολα παρατηρούμε ότι

$$\frac{dy}{dt} = M^{-1}AMy = Dy$$

Τότε, η επίλυση του συστήματος των διαφορικών εξισώσεων είναι άμεση και βρίσκουμε ότι:

$$x_i = (My)_i = \sum M_{ij} y_j(0) \exp(\lambda_i t), \quad j=1,2,\dots,n$$

Όπου $\lambda_j = d_{jj}$ είναι τα διαγώνια στοιχεία του πίνακα D . Οι αριθμοί λ_j , $j=1,2,\dots,n$ είναι οι ιδιοτιμές του πίνακα $A \in \mathbb{R}^{n \times n}$ και οι στήλες του πίνακα M αποτελούν τα ιδιοδιανύσματα του πίνακα A . Έτσι, η επίλυση του συστήματος των διαφορικών εξισώσεων προϋποθέτουν τον υπολογισμό των ιδιοτιμών και των ιδιοδιανυσμάτων του πίνακα A .

Στην περίπτωση των συστημάτων διαφορικών εξισώσεων όπου ο πίνακας A δεν είναι συμμετρικός το πρόβλημα είναι δυσκολότερο. Δηλαδή, αν για παράδειγμα οι ιδιοτιμές του A δεν είναι διαφορετικές τότε υπάρχει η πιθανότητα να μην υπάρχει ένα πλήρη σύνολο από γραμμικά ανεξάρτητα ιδιοδιανύσματα και έτσι ο πίνακας M δεν θα υπάρχει. •

ΣΗΜΕΙΩΣΗ: Εδώ πρόκειται να ασχοληθούμε με την εξέλιξη των αλγορίθμων για την επίλυση των αλγεβρικών προβλημάτων των ιδιοτιμών υποθέτοντας σε όλη τη διάρκεια ότι ο πίνακας $A \in \mathbb{R}^{n \times n}$ θα είναι συμμετρικός. Όπως σημειώθηκε παραπάνω, ανάλογα προβλήματα με μη συμμετρικούς πίνακες είναι περίπλοκα και δεν θα ασχοληθούμε με αυτά.

ΣΗΜΕΙΩΣΗ: Από εδώ και πέρα όταν θα αναφερόμαστε στο σύνολο των συμμετρικών πινάκων A $n \times n$ με $A = (a_{ij})$ θα συμβολίζουμε :

$$A \in \mathbb{R}_{sym}^{n \times n} \Leftrightarrow A \in \mathbb{R}^{n \times n} \text{ και } a_{ij} = a_{ji} \text{ (δηλαδή } A^T = A), \quad i, j = 1, 2, \dots, n \quad (1.3)$$

Παρακάτω παραθέτεται ένα βασικό θεώρημα που περιλαμβάνει κάποιες θεμελιώδεις ιδιότητες.

ΘΕΩΡΗΜΑ 1.4: Υποθέτουμε ότι έχουμε ένα πίνακα $A \in \mathbb{R}_{sym}^{n \times n}$. Τότε ισχύουν οι επόμενες προτάσεις.

- i. Υπάρχουν n γραμμικά ανεξάρτητα ιδιοδιανύσματα $x^{(i)} \in \mathbb{R}^n$ και αντίστοιχες ιδιοτιμές $\lambda_i \in \mathbb{R}$ τέτοια ώστε να ισχύει $Ax^{(i)} = \lambda_i x^{(i)}$ για κάθε $i = 1, 2, \dots, n$.
- ii. Η συνάρτηση

$$\lambda \rightarrow \det(A - \lambda I)$$

είναι πολυωνυμική βαθμού n με πρώτο όρο $(-1)^n \lambda^n$, καλείται **χαρακτηριστικό πολυώνυμο** του πίνακα A . Οι ιδιοτιμές του πίνακα A είναι οι ρίζες του χαρακτηριστικού πολυωνύμου.

iii. Αν οι ιδιοτιμές λ_i και λ_j του πίνακα A είναι διαφορετικές τότε τα αντίστοιχα ιδιοδιανύσματα $x^{(i)}$ και $y^{(j)}$ είναι ορθογώνια στον \mathbb{R}^n δηλαδή

$$\text{Αν } \lambda_i \neq \lambda_j \text{ τότε } x^{(i)T} x^{(j)} = 0, \quad i, j = 1, 2, \dots, n.$$

iv. Αν λ_i είναι μια ρίζα πολλαπλότητας m τότε υπάρχει ένας γραμμικός υπόχωρος στον \mathbb{R}^n διάστασης m , αποτελούμενος από m ανά δύο ορθογώνια μεταξύ τους ιδιοδιανύσματα που συνδέονται με την ιδιοτιμή λ_i .

v. Υποθέτουμε ότι κάθε ιδιοδιάνυσμα $x^{(i)}$ του πίνακα A είναι κανονικοποιημένο δηλαδή $x^{(i)T} x^{(i)} = 1$ για $i = 1, 2, \dots, n$ και έστω τετραγωνικός πίνακας X , στις στήλες του οποίου εισάγουμε τα κανονικοποιημένα (ορθογώνια) ιδιοδιανύσματα. Τότε ο πίνακας $\Lambda = X^T A X$ είναι διαγώνιος και τα διαγώνια στοιχεία του Λ αποτελούν τις ιδιοτιμές του A .

vi. Έστω $Q \in \mathbb{R}^{n \times n}$ είναι ένας ορθογώνιος πίνακας και ορίζουμε συμμετρικό πίνακα $B \in \mathbb{R}_{sym}^{n \times n}$ ώστε $B = Q^T A Q$ με $\det(B - \lambda I) = \det(A - \lambda I)$ για κάθε $\lambda \in \mathbb{R}$. Τότε οι ιδιοτιμές του είναι ίδιες με τις ιδιοτιμές του A και τα ιδιοδιανύσματα του B είναι τα διανύσματα $Q^T x^{(i)}, i = 1, 2, \dots, n$.

vii. Κάθε διάνυσμα $u \in \mathbb{R}^n$ μπορεί να εκφραστεί σαν γραμμικός συνδυασμός των ορθοκανονικοποιημένων ιδιοδιανυσμάτων $x^{(i)}, i = 1, 2, \dots, n$ του A δηλαδή

$$u = \sum_{i=1}^n a_i x^{(i)}, \quad a_i = x^{(i)T} u.$$

viii. Το ίχνος (trace) του πίνακα A , όπου $\text{Trace}(A) = \sum_{i=1}^n a_{ii}$, είναι ίσο με το άθροισμα των ιδιοτιμών του πίνακα A .

(Για την απόδειξη του παραπάνω Θεωρήματος βλέπε πίσω βιβλιογραφία το [2])

2. ΤΟ ΧΑΡΑΚΤΗΡΙΣΤΙΚΟ ΠΟΛΥΩΝΥΜΟ

Αρχικά θα παραθέσουμε δύο απλά παραδείγματα υπολογισμού των χαρακτηριστικών μεγεθών ενός πίνακα μικρού μεγέθους, χρησιμοποιώντας το χαρακτηριστικό πολυώνυμο. Στο πρώτο παράδειγμα (πίνακας 2×2) έχουμε ιδιοτιμή πολλαπλότητας 2 ενώ στο δεύτερο παράδειγμα (πίνακας 3×3) έχουμε μία ιδιοτιμή πολλαπλότητας 2 και μία πολλαπλότητας 1.

ΠΑΡΑΔΕΙΓΜΑ 2.1: Να βρεθούν οι ιδιοτιμές και τα ιδιοδιανύσματα του παρακάτω πίνακα 2×2

$$A = \begin{bmatrix} 1 & -1 \\ 1 & 3 \end{bmatrix}$$

Οι ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα A θα ικανοποιούν την χαρακτηριστική εξίσωση $(A - \lambda I)x = 0$ δηλαδή

$$\begin{bmatrix} 1-\lambda & -1 \\ 1 & 3-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \quad (2.1)$$

Για να βρούμε τις ιδιοτιμές του πίνακα A θα βρούμε τις ρίζες της εξίσωσης:

$$\det(A - \lambda I) = \begin{vmatrix} 1-\lambda & -1 \\ 1 & 3-\lambda \end{vmatrix} = \lambda^2 - 4\lambda + 4 = 0$$

Έχουμε δύο ιδιοτιμές (όχι διακεκριμένες) $\lambda_1 = 2$ και $\lambda_2 = 2$ δηλαδή $\lambda = 2$ με πολλαπλότητα 2.

Για $\lambda = 2$ το σύστημα (2.1) δίνει: $\begin{bmatrix} -1 & -1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

Από το παραπάνω σύστημα παίρνουμε: $-x_1 - x_2 = 0$ και $x_1 + x_2 = 0$. Δηλαδή, $x_1 + x_2 = 0$. Άρα αν $x_1 = c$ τότε $x_2 = -c$. Δηλαδή το αντίστοιχο ιδιοδιάνυσμα $x^{(1)}$ είναι:

$$x^{(1)} = c \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \quad c \neq 0 \quad (2.2)$$

Μπορούμε, βέβαια, να παραλείψουμε την αυθαίρετη σταθερά c όταν βρίσκουμε τα ιδιοδιανύσματα έτσι αντί της (2.2) μπορούμε να γράψουμε:

$$x^{(1)} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

Παρατήρηση: Όταν έχουμε ιδιοτιμή λ πολλαπλότητας m τότε μπορούμε να βρούμε q γραμμικώς ανεξάρτητα ιδιοδιανύσματα όπου $1 \leq q \leq m$. Δηλαδή παραπάνω είχαμε την ιδιοτιμή $\lambda=2$ με πολλαπλότητα 2 και έχουμε ένα γραμμικώς ανεξάρτητο ιδιοδιάνυσμα που αντιστοιχεί σε αυτή την διπλή ιδιοτιμή. (Το παράδειγμα αυτό είναι από τη βιβλιογραφία [3]-κεφάλαιο 7, σελίδα 409)•

ΠΑΡΑΔΕΙΓΜΑ 2.2: Να βρεθούν τα χαρακτηριστικά μεγέθη του παρακάτω πίνακα 3×3

$$A = \begin{bmatrix} 4 & 6 & 6 \\ 1 & 3 & 2 \\ -1 & -5 & -2 \end{bmatrix}$$

Οι ιδιοτιμές και τα ιδιοδιανύσματα πρέπει να ικανοποιούν το παρακάτω σύστημα:

$$\begin{bmatrix} 4-\lambda & 6 & 6 \\ 1 & 3-\lambda & 3 \\ 1 & -5 & -2-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \quad (2.3)$$

Οι ιδιοτιμές είναι οι ρίζες της εξίσωσης:

$$\det(A - \lambda I) = (\lambda - 1)(\lambda - 2)^2$$

Δηλαδή, οι ιδιοτιμές είναι $\lambda_1 = 1$ και $\lambda_2 = 2$ πολλαπλότητας 2. Παρακάτω θα βρούμε τα ιδιοδιανύσματα για κάθε μία από της ιδιοτιμές.

$$\text{Για } \lambda_1 = 1 \text{ από το σύστημα (2.3) παίρνουμε: } \begin{cases} 3x_1 + 6x_2 + 6x_3 = 0 \\ x_1 + 2x_2 + 2x_3 = 0 \\ x_1 + 5x_2 + 3x_3 = 0 \end{cases}$$

Η λύση του παραπάνω συστήματος μας δίνει το ιδιοδιάνυσμα που αντιστοιχεί στην πρώτη ιδιοτιμή το οποίο είναι $x^{(1)} = [-4/3 \quad -1/3 \quad 1]^T$.

$$\text{Για } \lambda_2 = 2 \text{ (πολλαπλότητα 2) από την (2.3) παίρνουμε: } \begin{cases} 2x_1 + 6x_2 + 6x_3 = 0 \\ x_1 + x_2 + 2x_3 = 0 \\ x_1 + 5x_2 + 4x_3 = 0 \end{cases}$$

Η λύση του παραπάνω συστήματος είναι το αντίστοιχο ιδιοδιάνυσμα της ιδιοτιμής λ_2 που είναι το $x^{(2)} = [-3/2 \quad -1/2 \quad 1]^T$. (Το παράδειγμα αυτό είναι από το βιβλιογραφία [4]-κεφάλαιο 6, σελίδα 137). •

Στα παραπάνω παραδείγματα φαίνεται ότι δοθέντος ενός πίνακα $A \in \mathbb{R}^{n \times n}$ με $n \leq 4$ είναι αρκετά εύκολο να καταγράψουμε το χαρακτηριστικό πολυώνυμο $\det(A - \lambda I)$ αναπτύσσοντας την ορίζουσα, στη συνέχεια βρίσκουμε τις ρίζες αυτού του πολυωνύμου βαθμού n και ορίζουμε τις ιδιοτιμές του πίνακα A . Στην περίπτωση όμως που έχουμε $n \geq 5$ δεν υπάρχει σαφής μέθοδος για τις ρίζες του πολυωνύμου και έτσι πρέπει να καταφύγουμε σε αριθμητικές τεχνικές. Μία επιπρόσθετη δυσκολία είναι ότι οι ρίζες μπορεί να είναι πολύ ευαίσθητες στις μικρές αλλαγές στους συντελεστές του πολυωνύμου και αυτό γιατί η επίδραση των στρογγυλοποιημένων σφαλμάτων στην κατασκευή του χαρακτηριστικού πολυωνύμου είναι συνήθως καταστροφική.

ΠΑΡΑΔΕΙΓΜΑ 2.3: Θεωρούμε ένα διαγώνιο πίνακα μεγέθους 16 του οποίου τα διαγώνια στοιχεία είναι $j+1/3, j=1,2,\dots,16$. Οι ιδιοτιμές είναι τα διαγώνια στοιχεία του πίνακα. Κατασκευάζοντας το χαρακτηριστικό πολυώνυμο δουλεύοντας με 10 σημαντικά ψηφία δίνεται το παρακάτω αποτέλεσμα:

$$\lambda^{16} - 141.3333333\lambda^{15} + 9193.333333\lambda^{14} - \dots$$

Χρησιμοποιώντας ένα σταθερό αριθμητικό αλγόριθμο (μέθοδος Newton) για να υπολογίσουμε τις ρίζες του πολυωνύμου και δουλεύοντας με 10 σημαντικά ψηφία δίνουν τη μικρότερη ρίζα 1.33333331 η οποία είναι η πιο κοντινή λύση στα 10 σημαντικά ψηφία. Οι 3 μεγαλύτερες λύσεις έχουν υπολογιστεί κατά προσέγγιση και είναι οι $15.5 \pm 1.3i$ και 16.7 οι οποίες είναι διαφορετικές από τις πραγματικές τιμές 14.3, 15.3, 16.3 αντίστοιχα, αν και ο πίνακας σε αυτό παράδειγμα είναι σχετικά μικρού μεγέθους και οι ιδιοτιμές έχουν καλή απόσταση. Από αυτό το παράδειγμα συμπεραίνουμε ότι αυτή η αριθμητική μέθοδος η οποία κατασκευάζει το χαρακτηριστικό πολυώνυμο και βρίσκει τις ρίζες του είναι τελείως ανεπαρκής για γενική χρήση, παρά μόνο για πίνακες με πολύ μικρό μέγεθος. •

ΣΗΜΕΙΩΣΗ: Το γεγονός ότι στην γενική περίπτωση οι ρίζες του χαρακτηριστικού πολυωνύμου δεν μπορούν να δοθούν με κάποια συγκεκριμένη φόρμα δείχνει ότι η μέθοδος πρέπει να συνεχίζει με διαδοχικές προσεγγίσεις. Ακόμη και αν δεν μπορεί να παράγει τις αναγκαίες ιδιοτιμές σε πεπερασμένο αριθμό βημάτων τότε θα δούμε ότι υπάρχουν γρήγορες, συγκλίνουσες, επαναληπτικές μέθοδοι για τον υπολογισμό των ιδιοτιμών και των ιδιοδιανυσμάτων αριθμητικά.

ΣΗΜΕΙΩΣΗ: Είναι πολύ δύσκολο να υπολογίσουμε το χαρακτηριστικό πολυώνυμο και στη συνέχεια τις ρίζες του για πολύ μεγάλο n . Άρα είναι ανάγκη να χρησιμοποιήσουμε προσεγγιστικές μεθόδους για τον υπολογισμό των ιδιοτιμών και των αντίστοιχων ιδιοδιανυσμάτων ενός πίνακα.

Επίσης, στη συνέχεια οι πίνακες που θα χρησιμοποιούμε θα είναι συμμετρικοί.

3. ΜΕΘΟΔΟΣ JACOBI

Αυτή η μέθοδος χρησιμοποιεί διαδοχικούς ορθογώνιους μετασχηματισμούς για να παράγει μια ακολουθία από πίνακες η οποία να προσεγγίζει ένα διαγώνιο πίνακα στο όριο. Κάθε βήμα στη διαδικασία προϋποθέτει ένα πίνακα που απεικονίζει μια επίπεδη περιστροφή. Παρακάτω ξεκινάμε με ένα απλό παράδειγμα.

ΠΑΡΑΔΕΙΓΜΑ 3.1: (Η επίπεδη περιστροφή πίνακα στον \mathbb{R}^2) Ας υποθέσουμε ότι $\varphi \in [-\pi, \pi]$ και θεωρούμε ένα πίνακα $R(\varphi) \in \mathbb{R}^{2 \times 2}$ που ορίζεται ως εξής:

$$R(\varphi) = \begin{bmatrix} \cos \varphi & \sin \varphi \\ -\sin \varphi & \cos \varphi \end{bmatrix}$$

Για ένα διάνυσμα $x \in \mathbb{R}^2$, $R(\varphi)x$ είναι η επίπεδη περιστροφή του x από την αρχή έως τη γωνία φ (όταν $\varphi > 0$ φορά σύμφωνα με τους δείκτες του ρολογιού και όταν $\varphi < 0$ αντίθετη)

Υπενθυμίζουμε ότι $\cos(-\varphi) = \cos \varphi$, $\sin(-\varphi) = -\sin(\varphi)$ και $\sin^2 \varphi + \cos^2 \varphi = 1$.

Έτσι έχουμε:

$$(R(\varphi))^T = R(-\varphi) \quad \text{και} \quad R(\varphi)R(-\varphi) = I$$

Συνεπώς, ο $R(\varphi)$ είναι ορθογώνιος πίνακας αφού $R(\varphi)R(\varphi)^T = R(\varphi)^T R(\varphi) = I$ όπου I είναι ένας 2×2 μοναδιαίος πίνακας. •

Ο επόμενος ορισμός επεκτείνει την έννοια της επίπεδης περιστροφής πίνακα από τον \mathbb{R}^2 στον \mathbb{R}^n .

ΟΡΙΣΜΟΣ 3.2: (Η επίπεδη περιστροφή πίνακα στον \mathbb{R}^n) Υποθέτουμε ότι $n \geq 2$, $1 \leq p < q \leq n$ και $\varphi \in [-\pi, \pi]$. Θεωρούμε ένα πίνακα $R^{(pq)} \in \mathbb{R}^{n \times n}$ του οποίου τα στοιχεία είναι ίδια με αυτά του μοναδιαίου πίνακα $I \in \mathbb{R}^{n \times n}$ εκτός από τα παρακάτω στοιχεία:

$$r_{pp} = c, \quad r_{pq} = s, \quad r_{qp} = -s, \quad r_{qq} = c \quad \text{όπου} \quad c = \cos \varphi \quad \text{και} \quad s = \sin \varphi$$

Όπως στο προηγούμενο παράδειγμα, είναι άμεσο να δούμε ότι

$$(R^{(pq)}(\varphi))^T = R^{(pq)}(-\varphi), \quad R^{(pq)}(\varphi)R^{(pq)}(-\varphi) = I$$

Και επομένως, $R^{(pq)}(\varphi)(R^{(pq)}(\varphi))^T = (R^{(pq)}(\varphi))^T R^{(pq)}(\varphi) = I$

Συνεπώς, $R^{(pq)} \in \mathbb{R}^{n \times n}$ είναι ένας ορθογώνιος πίνακας για p, q που $1 \leq p < q \leq n$ και για γωνίες φ όπου $\varphi \in [-\pi, \pi]$.

Το βασικό αποτέλεσμα που αποτελεί το θεμέλιο της μεθόδου Jacobi παρουσιάζεται στο παρακάτω θεώρημα.

ΘΕΩΡΗΜΑ 3.3: Υποθέτουμε ότι έχουμε πίνακα $A \in \mathbb{R}_{sym}^{n \times n}$. Για κάθε ζεύγος ακεραίων (p, q) με $1 \leq p < q \leq n$, υπάρχει $\varphi \in [-\pi/4, \pi/4]$ τέτοια ώστε το (p, q) -στοιχείο του συμμετρικού πίνακα $(R^{(pq)}(\varphi))^T A R^{(pq)}(\varphi)$ να είναι ίσο με το μηδέν.

Απόδειξη: Αρχικά για χάρη απλοποίησης θα γράφουμε R αντί για $R^{(pq)}(\varphi)$ σε όλη την απόδειξη και επίσης $c = \cos \varphi$ και $s = \sin \varphi$.

Θεωρούμε το γινόμενο $A' = AR$. Προφανώς, η μόνη διαφορά ανάμεσα στον A' και στον A είναι στις στήλες p και q , καθώς οι στήλες του A' είναι γραμμικός συνδυασμός των δύο ίδιων στηλών του πίνακα A (λόγω συμμετρίας):

$$\begin{cases} a_{ip}' = a_{ip}c - a_{iq}s \\ a_{iq}' = a_{ip}s + a_{iq}c \end{cases}, \quad i = 1, 2, \dots, n \quad (3.1)$$

Πολλαπλασιασμός του A' με το R^T (από αριστερά) δίνει παρόμοιο αποτέλεσμα, αλλά επηρεάζει τις γραμμές p και q . Γράφοντας $B = R^T A'$ δίνει:

$$\begin{cases} b_{pj} = a_{pj}'c - a_{qj}'s \\ b_{qj} = a_{pj}'s + a_{qj}'c \end{cases}, \quad j = 1, 2, \dots, n \quad (3.2)$$

Συνδυάζοντας αυτές τις συναρτήσεις δείχνουμε ότι $B = R^T A R$, όπου

$$\begin{cases} b_{pp} = a_{pp}c^2 - 2a_{pq}sc + a_{qq}s^2 \\ b_{qq} = a_{pp}s^2 + 2a_{pq}sc + a_{qq}c^2 \\ b_{pq} = (a_{pp} - a_{qq})sc + a_{pq}(c^2 - s^2) = b_{qp} \end{cases} \quad (3.3)$$

Τα υπόλοιπα στοιχεία του πίνακα $B = R^T A R$ στις στήλες p και q δίνονται από τις παρακάτω εκφράσεις:

$$\begin{cases} b_{ip} = a_{ip}c - a_{iq}s \\ b_{iq} = a_{ip}s - a_{iq}c \end{cases}, \quad i = 1, 2, \dots, n, \quad i \neq p, q$$

Ο πίνακας $B = R^T A R$ είναι προφανώς συμμετρικός, και έτσι τα μη διαγώνια στοιχεία του B δίνονται από τις ίδιες εκφράσεις.

Τέλος, σημειώνεται ότι όλα τα στοιχεία του B που δεν βρίσκονται είτε στη γραμμή p ή q ή στη στήλη p ή q , είναι ίδια με τα αντίστοιχα στοιχεία του πίνακα A δηλαδή:

$$b_{ij} = a_{ij} \text{ αν } i \neq p, q \text{ και } j \neq p, q.$$

Από τη σχέση (3.3) παρατηρούμε ότι για να διασφαλίσουμε ότι το b_{pq} , δηλαδή το (p, q) -στοιχείο του πίνακα $B = R^T A R$, είναι ίσο με το μηδέν, αρκεί να διαλέξουμε γωνία φ τέτοια ώστε:

$$\tan 2\varphi = \frac{2a_{pq}}{a_{qq} - a_{pp}}.$$

$$\text{Έτσι, επιλέγουμε } \varphi = \frac{1}{2} \tan^{-1} \frac{2a_{pq}}{a_{qq} - a_{pp}} \in [-\pi/4, \pi/4] \quad (3.4)$$

Για να το δούμε, χρησιμοποιούμε τις τριγωνομετρικές ταυτότητες $c^2 - s^2 = \cos(2\varphi)$ και $sc = \frac{1}{2} \sin(2\varphi)$ στο b_{pq} στη σχέση (3.3) με $b_{pq} = 0$. Αυτό ολοκληρώνει την απόδειξη. •

ΣΗΜΕΙΩΣΗ : Παρατηρούμε ότι για να αποφύγουμε τους περίπλοκους τριγωνομετρικούς υπολογισμούς στον τύπο (3.4) για το φ μπορούμε να γράψουμε $t = s/c$ και βλέπουμε ότι το t πρέπει να ικανοποιεί την

$$(a_{pp} - a_{qq})t + a_{pq}(1 - t^2) = 0 \quad (3.5)$$

Αν $a_{pq} = 0$ μπορούμε να διασφαλίσουμε ότι η σχέση (3.5) κρατάει την επιλογή $t = 0$ (που αντιστοιχεί με την επιλογή $\varphi = 0$). Αν $a_{pq} \neq 0$ και $a_{pp} = a_{qq}$ θέτουμε $t = 1$ (αντίστοιχα $\varphi = \pi/4$). Τέλος, αν $a_{pq} \neq 0$ και $a_{pp} \neq a_{qq}$ τότε λύνουμε την δευτεροβάθμια εξίσωση (3.5) που θα έχει δύο διακεκριμένες πραγματικές ρίζες, οπότε ορίζουμε ως t την μικρότερη από αυτές σε απόλυτη τιμή. Έχοντας επιλέξει το t , στη συνέχεια χρησιμοποιούμε τη σχέση $\sec^2 \varphi = 1 + \tan^2 \varphi$ για να υπολογίσουμε το c από $c = 1 / (1 + t^2)^{1/2}$ και μετά το s από τη σχέση $s = ct$.

ΟΡΙΣΜΟΣ 3.4: (Η κλασική μέθοδος Jacobi) Έστω πίνακας $A \in \mathbb{R}_{sym}^{n \times n}$ και ορίζουμε $A^{(0)} = A$. Δοθέντος $k \geq 0$ και $A^{(k)} \in \mathbb{R}_{sym}^{n \times n}$, το βασικό βήμα της μεθόδου Jacobi υπολογίζει πίνακα $A^{(k+1)} \in \mathbb{R}_{sym}^{n \times n}$ επιλέγοντας πρώτα το μεγαλύτερο σε απόλυτη τιμή μη διαγώνιο στοιχείο $(A^{(k)})_{pq} \equiv a_{pq}^{(k)}$ του πίνακα $A^{(k)}$ και μετά θέτουμε

$A^{(k+1)} = (R^{(pq)}(\varphi_k))^T A^{(k)} R^{(pq)}(\varphi_k)$ με φ_k τέτοια ώστε $(A^{(k+1)})_{pq} = 0$. Αυτή η διαδικασία στη συνέχεια επαναλαμβάνεται μέχρι όλα τα μη διαγώνια στοιχεία να είναι μικρότερα από ένα δοθέν θετικό όριο ανοχής ε .

Παρακάτω, παραθέτεται ο αλγόριθμος της μεθόδου Jacobi:

ΑΛΓΟΡΙΘΜΟΣ JACOBI:

Για $k=0$ θέτουμε $A^{(0)} = A$

Για $k=1,2,\dots$ υπολογίζουμε $A^{(k+1)} = (R^{(pq)}(\varphi))^T A^{(k)} R^{(pq)}(\varphi)$. (3.6)

όπου τα pq επιλέγονται από το μέγιστο κατά μέτρο μη διαγώνιο στοιχείο του $A^{(k)}$ και γωνία φ τέτοια ώστε το $(A^{(k+1)})_{pq} = 0$.

Με αυτόν τον τρόπο ανακτούμε τη ζητούμενη ακολουθία πινάκων $(A^{(k)})$.

Για να δείξουμε ότι όταν $k \rightarrow \infty$ η ακολουθία πινάκων $(A^{(k)})$, η οποία παράγεται από τα διαδοχικά βήματα της κλασικής μεθόδου Jacobi, συγκλίνει σε ένα διαγώνιο πίνακα (του οποίου τα διαγώνια στοιχεία θα αποτελούν τις ιδιοτιμές του αρχικού πίνακα A), χρειαζόμαστε το παρακάτω αποτέλεσμα.

ΛΗΜΜΑ 3.5: Το άθροισμα των τετραγώνων των στοιχείων του συμμετρικού πίνακα είναι αμετάβλητο κάτω από τους ορθογώνιους μετασχηματισμούς. Έτσι εάν $A \in \mathbb{R}_{sym}^{n \times n}$ και $B = R^T A R$ όπου $R \in \mathbb{R}^{n \times n}$ ορθογώνιος πίνακας τότε:

$$\sum_{i=1}^n \sum_{j=1}^n b_{ij}^2 = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 \quad (3.7)$$

Η ποσότητα

$$\|A\|_F = \left(\sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 \right)^{1/2} \quad (3.8)$$

καλείται **Frobenius νόρμα** του πίνακα $A \in \mathbb{R}^{n \times n}$. Η Frobenius νόρμα του πίνακα A είναι μία 2-νόρμα του A , με A να αντιμετωπίζεται σαν ένα στοιχείο του γραμμικού χώρου διάστασης n^2 πάνω στο σώμα των πραγματικών αριθμών. Επισημαίνεται ότι η Frobenius νόρμα στον $\mathbb{R}^{n \times n}$ δεν παράγεται ως φυσική νόρμα πίνακα που αντιστοιχεί σε διανυσματική νόρμα.

Ένα μπορούμε να διατυπώσουμε, λέγοντας ότι η Frobenius νόρμα ενός συμμετρικού πίνακα A παραμένει αμετάβλητη παρά τους ορθογώνιους μετασχηματισμούς, ότι $\|B\|_F = \|A\|_F$ όπου $B = R^T A R$.

Απόδειξη:(λήμματος) Το άθροισμα των τετραγώνων των στοιχείων του πίνακα A είναι ίσο με το trace (ίχνος) του πίνακα A^2 (επειδή ο πίνακας A είναι συμμετρικός), έτσι

$$\text{Trace}(A^2) = \sum_{i=1}^n (A^2)_{ii} = \sum_{i=1}^n \sum_{j=1}^n a_{ij} a_{ji} = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^2 \quad (3.9)$$

Ομοίως, με $B = R^T A R$ συμμετρικό θα έχουμε:

$$\text{Trace}(B^2) = \sum_{i=1}^n \sum_{j=1}^n b_{ij}^2$$

Άρα τώρα απομένει να δείξουμε ότι $\text{Trace}(A^2) = \text{Trace}(B^2)$. Όμως, παρατηρούμε ότι

$$B^2 = (R^T A R) (R^T A R) = R^T A^2 R \text{ όπου ο } R \text{ ορθογώνιος πίνακας.}$$

Συνεπώς, ο πίνακας B^2 είναι ένας ορθογώνιος μετασχηματισμός του A^2 το οποίο σημαίνει (από το Θεώρημα 1.4 (vi)) ότι οι πίνακες B^2 και A^2 έχουν τις ίδιες ιδιοτιμές και συνεπώς τα ίδια ίχνη (trace) αφού γνωρίζουμε ότι το ίχνος είναι το άθροισμα των ιδιοτιμών. •

Αναφέραμε παραπάνω ότι μας ενδιαφέρει στις επαναληπτικές μεθόδους το θέμα της σύγκλισης. Παρακάτω θα αναλύσουμε το θέμα της σύγκλισης στην μέθοδο Jacobi.

ΘΕΩΡΗΜΑ 3.6: Υποθέτουμε ότι έχουμε ένα πίνακα $A \in \mathbb{R}_{sym}^{n \times n}$, $n \geq 2$. Στην κλασσική μέθοδο Jacobi τα μη διαγώνια στοιχεία της ακολουθίας των πινάκων $(A^{(k)})$, που παράγεται από το $A^{(0)} = A$ σύμφωνα με τον ορισμό (3.4), συγκλίνουν στο μηδέν υπό την έννοια ότι

$$\lim_{k \rightarrow \infty} \sum_{i,j=1, i \neq j}^n [(A^{(k)})_{ij}]^2 = 0 \quad (3.10)$$

Επίσης,

$$\lim_{k \rightarrow \infty} \sum_{i=1}^n [(A^{(k)})_{ii}]^2 = \text{Trace}(A^2) \quad (3.11)$$

ΣΗΜΕΙΩΣΗ: Επιπλέον, γνωρίζουμε ότι κάθε διαγώνιο στοιχείο της ακολουθίας $(A^{(k)})$ δηλαδή κάθε $(a_{ii}^{(k)})$ συγκλίνει σε μία ιδιοτιμή (αλλά δεν ξέρω σε ποια). (Η απόδειξη της τελευταίας πρότασης αυτού του θεωρήματος δεν υπάρχει στην παρακάτω απόδειξη)

Απόδειξη: Έστω a_{pq} είναι το μη διαγώνιο στοιχείο του πίνακα A με την μεγαλύτερη κατά απόλυτη τιμή και έστω $B = (R^{(pq)}(\varphi))^T A R^{(pq)}(\varphi)$, όπου φ ορίζεται από τον τύπο (3.4). Επίσης, θέτουμε ότι $c = \cos \varphi$ και $s = \sin \varphi$. Έτσι έχουμε:

$$\begin{pmatrix} b_{pp} & b_{pq} \\ b_{qp} & b_{qq} \end{pmatrix} = \begin{pmatrix} c & s \\ -s & c \end{pmatrix}^T \begin{pmatrix} a_{pp} & a_{pq} \\ a_{qp} & a_{qq} \end{pmatrix} \begin{pmatrix} c & s \\ -s & c \end{pmatrix}$$

και από το Λήμμα 3.5 (σχέση 3.7) θα έχω ότι: $b_{pp}^2 + 2b_{pq}^2 + b_{qq}^2 = a_{pp}^2 + 2a_{pq}^2 + a_{qq}^2$

Θέτοντας ότι:

$$S(A) = \sum_{i,j=1}^n a_{ij}^2, \quad D(A) = \sum_{i=1}^n a_{ii}^2, \quad L(A) = \sum_{i,j=1, i \neq j}^n a_{ij}^2 \quad (3.12)$$

Λαμβάνουμε, $S(A) = D(A) + L(A)$. Όμως από το λήμμα (3.5) ξέρουμε ότι ισχύει $S(B) = S(A)$ και επομένως $D(B) + L(B) = D(A) + L(A)$.

Τα διαγώνια στοιχεία του πίνακα B είναι τα ίδια με αυτά του πίνακα A εκτός από τα στοιχεία στις γραμμές p και q ($1 \leq p \leq q \leq n$).

Επίσης, επειδή το $b_{pq} = 0$ (λόγω κατασκευής του αλγορίθμου) θα έχουμε ότι :

$$b_{pp}^2 + b_{qq}^2 = a_{pp}^2 + 2a_{pq}^2 + a_{qq}^2$$

Επομένως, από την τελευταία ισότητα συμπεραίνουμε ότι

$$D(B) = D(A) + 2a_{pq}^2$$

και συνεπώς,

$$L(B) = L(A) - 2a_{pq}^2$$

Το a_{pq} είναι το μεγαλύτερο μη διαγώνιο στοιχείο του πίνακα A . Συνεπώς,

$L(A) \leq N a_{pq}^2$, όπου $N = n(n-1)$ είναι ο αριθμός των μη διαγώνιων στοιχείων και

άρα

$$L(B) \leq (1 - 2/N) L(A) \quad (3.13)$$

Γράφοντας $A^{(0)} = A$, $A^{(1)} = B$ και παράγοντας υπακολουθίες της ακολουθίας $(A^{(k)})$ με τον ίδιο τρόπο, όπως εμφανίζεται στον αλγόριθμο του ορισμού (3.2), από τη σχέση (3.13) συμπεραίνουμε ότι

$$0 \leq L(A^{(k)}) \leq (1 - 2/N)^k L(A), \quad k = 1, 2, 3, \dots \quad (3.14)$$

Όπου $N \geq 2$. Έτσι, συμπεραίνουμε ότι $\lim_{k \rightarrow \infty} L(A^{(k)}) = 0$.

Τώρα από τις σχέσεις (3.10), (3.11) και (3.9) έχουμε ότι:

$$\text{Trace}(A^2) = S(A) = S(A^{(k)}) = D(A^{(k)}) + L(A^{(k)}), \quad \forall k \geq 0$$

Και οριακά όταν $k \rightarrow \infty$ έχουμε: $\text{Trace}(A^2) = \lim_{k \rightarrow \infty} D(A^{(k)})$. •

Σύμφωνα με το Θεώρημα 1.4 (viii) το ίχνος του πίνακα A^2 είναι ίσο με το άθροισμα των ιδιοτιμών του A^2 και οι ιδιοτιμές του A^2 είναι τα τετράγωνα των ιδιοτιμών του πίνακα A . Έτσι, έχουμε δείξει ότι το άθροισμα των τετραγώνων των διαγώνιων στοιχείων στην ακολουθία πινάκων $(A^{(k)})$, που παράγεται από την κλασική μέθοδο Jacobi, συγκλίνει στο άθροισμα των τετραγώνων των ιδιοτιμών του πίνακα A . Περισσότερη δουλειά απαιτείται όμως για να δείξουμε ότι για κάθε $i = 1, 2, \dots, n$ ακολουθία των διαγώνιων στοιχείων $(a_{ii}^{(k)})$ συγκλίνει σε μία ιδιοτιμή του πίνακα A όταν $k \rightarrow \infty$.

Παρακάτω, περιγράφεται μία παραλλαγή της μεθόδου Jacobi.

ΟΡΙΣΜΟΣ 3.7: (Η σειριακή μέθοδος Jacobi) Αυτή η παραλλαγή της μεθόδου Jacobi προχωρά με μία συστηματική σειρά, χρησιμοποιώντας μετασχηματισμούς $R^{(pq)}(\rho)$ που ανάγουν στο μηδέν τα στοιχεία $(1, 2), (1, 3), \dots, (1, n), (2, 3), (2, 4), \dots, (2, n), \dots, (n-1, n)$ με αυτή τη σειρά. Το πλήρες βήμα στη συνέχεια επαναλαμβάνεται.

Δεν είναι δύσκολο να αποδείξουμε ότι η μέθοδος αυτή επίσης συγκλίνει. Και οι δύο μορφές της μεθόδου Jacobi συγκλίνουν αρκετά γρήγορα. Ο ρυθμός της σύγκλισης είναι στην πραγματικότητα πολύ πιο γρήγορος από ότι αναφέρθηκε στον τύπο (3.14) και πράγματι μπορεί να δειχθεί ότι η σύγκλιση είναι τελικά τετραγωνική.

Παρακάτω παραθέτουμε ένα παράδειγμα.

ΠΑΡΑΔΕΙΓΜΑ 3.8: Έστω ένας συμμετρικός πίνακας 5×5

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

Οι τιμές των $D(A^{(k)})$ και $L(A^{(k)})$ (που υπολογίζονται από τη σχέση 3.12) μετά από κάθε επανάληψη της σειριακής μεθόδου Jacobi, με $A^{(0)} = A$, φαίνονται στον πίνακα (3.1). Τα μη-διαγώνια στοιχεία στην τρίτη επανάληψη, $A^{(3)}$, είναι μηδέν. Τα διαγώνια στοιχεία του $A^{(3)}$, τα οποία μας δίνουν τις ιδιοτιμές, είναι τα παρακάτω:

$$8.094, 1.690, -0.671, 7.170, -3.282.$$

(Σημειώνουμε ότι οι ιδιοτιμές δεν εμφανίζονται με κάποια συγκεκριμένη σειρά)

Πίνακας (3.1)

k	$D(A^{(k)})$	$L(A^{(k)})$
0	43.000	88.00000000
1	126.309	4.69087885
2	130.981	0.01948855
3	131.000	0.00000000

Σε αυτό το κεφάλαιο εξετάσαμε τη χρήση της μεθόδου Jacobi για τον υπολογισμό των ιδιοτιμών συμμετρικού πίνακα A . Συνεχίζουμε με τον προσδιορισμό των ιδιοδιανυσμάτων του πίνακα A . Είναι αρκετά εύκολο να υπολογίσουμε τα ιδιοδιανύσματα του A . Σημειώνουμε ότι αν ο M είναι ένας ορθογώνιος πίνακας τέτοιος ώστε $M^T A M = D$, όπου D διαγώνιος, τότε τα διαγώνια στοιχεία του πίνακα D είναι οι ιδιοτιμές του A και οι στήλες του πίνακα M τα αντίστοιχα ιδιοδιανύσματα του A .

Στην πορεία για τις Jacobi επαναλήψεις (είτε την κλασσική είτε την σειριακή μέθοδο), έχουμε κατασκευάσει την επίπεδη περιστροφή $R^{(p,q_j)}(\varphi_j)$, $j=1,2,\dots,k$. Έτσι, η προσέγγιση $M^{(k)}$ ενός ορθογώνιου πίνακα M μπορεί να αποκτάται

λαμβάνοντας υπόψη το γινόμενο N πινάκων στρέψης: Αρχικά , θέτω $M^{(0)} = I$ και μετά εφαρμόζουμε τον μετασχηματισμό στήλης $R^{(p_j q_j)}(\varphi_j)$ σε κάθε βήμα

$$j = 1, 2, \dots, k.$$

Αυτό αντιστοιχεί στον υπολογισμό $M^{(j-1)}$ και καταλήγουμε στον ορθογώνιο πίνακα

$$M^{(k)} = R^{(p_1 q_1)}(\varphi_1) \dots \dots \dots R^{(p_k q_k)}(\varphi_k)$$

ο οποίος απεικονίζει την απαιτούμενη προσέγγιση σε ένα ορθογώνιο πίνακα M . Οι στήλες του $M^{(k)}$ θα προσεγγίζουν τα ιδιοδιανύσματα του πίνακα A αντίστοιχα με την προσέγγιση ιδιοτιμών οι οποίες εμφανίζονται στην διαγώνιο του $A^{(k)}$.

Η μέθοδος του Jacobi συνήθως συγκλίνει σε λογικό αριθμό επαναλήψεων και είναι ικανοποιητική για μικρούς και μέτριους σε μέγεθος πίνακες. Όμως , υπάρχουν αρκετά προβλήματα , ειδικά στην περιοχή της αριθμητικής επίλυσης των μερικών διαφορικών εξισώσεων , οι οποίες κινούνται σε πολύ μεγάλους πίνακες που είναι αραιοί , με τα περισσότερα από τα στοιχεία τους να είναι ίσα με το μηδέν. Μία πρόσθετη εκτίμηση είναι ότι σε πολλές πρακτικές καταστάσεις δεν χρειάζεται να υπολογίζουμε όλες τις ιδιοτιμές. Συχνότερα απαιτούνται μερικές από τις μεγαλύτερες κατά μέτρο ιδιοτιμές και τα αντίστοιχα ιδιοδιανύσματα ή αντίθετα μερικές από τις μικρότερες κατά μέτρο ιδιοτιμές. Η μέθοδος του Jacobi δεν είναι κατάλληλη για τέτοιου είδους προβλήματα, επειδή πάντα μας δίνει όλες τις ιδιοτιμές και δεν θα διατηρεί την αραιή κατασκευή του πίνακα κατά τη διάρκεια της εξέλιξης των επαναλήψεων. Για παράδειγμα είναι εύκολο να δούμε ότι αν η μέθοδος Jacobi εφαρμοστεί σε ένα τριδιαγώνιο συμμετρικό πίνακα στη συνέχεια μετά το τέλος μίας σάρωσης δύο από τα στοιχεία του πίνακα θα είναι γενικά διάφορα του μηδενός και , αν και παραμένει συμμετρικός , ο μετασχηματισμένος πίνακας δεν είναι πλέον τριδιαγώνιος. Παρακάτω θα εξετάσουμε αριθμητικούς αλγόριθμους που υπολογίζουν επιλεγμένες ιδιοτιμές ενός πίνακα.

4. ΤΑ ΘΕΩΡΗΜΑΤΑ ΤΟΥ GERSCHGORIN

Το Θεώρημα Gerschgorin παρέχει ένα πολύ απλό τρόπο για τον προσδιορισμό μιας περιοχής η οποία θα περιέχει τις ιδιοτιμές ενός πίνακα. Είναι πολύ γενικό και δεν χρειάζεται να υποθέσουμε ότι ο πίνακας είναι συμμετρικός. Πράγματι, επιτρέπουμε στα στοιχεία τετραγωνικού πίνακα τάξης n να είναι μιγαδικά και μπορούμε να δηλώνουμε $A \in \mathbb{C}^{n \times n}$.

ΟΡΙΣΜΟΣ 4.1: Υποθέτουμε ότι $n \geq 2$ και $A \in \mathbb{C}^{n \times n}$. Ο δίσκος Gerschgorin D_i , $i=1,2,\dots,n$ του πίνακα A ορίζεται ως ένα κλειστό κυκλικό χωρίο

$$D_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq R_i\} \quad (4.1)$$

στο μιγαδικό επίπεδο, όπου

$$R_i = \sum_{j=1, j \neq i}^n |a_{ij}| \quad (4.2)$$

είναι η ακτίνα του D_i (Το R_i είναι το άθροισμα των στοιχείων της i γραμμής εξαιρουμένου του στοιχείου a_{ii}).

ΘΕΩΡΗΜΑ 4.2: (Θεώρημα Gerschgorin) Έστω $n \geq 2$ και $A \in \mathbb{C}^{n \times n}$. Όλες οι ιδιοτιμές του πίνακα A βρίσκονται στην περιοχή $D = \bigcup_{i=1}^n D_i$, όπου D_i , $i=1,2,\dots,n$ είναι οι δίσκοι Gerschgorin του A όπως ορίστηκαν στις σχέσεις (4.1) και (4.2).

Απόδειξη: Υποθέτουμε ότι $\lambda \in \mathbb{C}$ και $x \in \mathbb{C}_*^n$ είναι μία ιδιοτιμή και το αντίστοιχο ιδιοδιάνυσμα του A έτσι ώστε :

$$\sum_{j=1}^n a_{ij} x_j = \lambda x_i, i=1,2,\dots,n \quad (4.3)$$

Υποθέτουμε ότι x_k με $k \in \{1,2,\dots,n\}$, είναι η συντεταγμένη του x με την μεγαλύτερη κατά απόλυτο τιμή ή ένα από αυτά αν περισσότερα από ένα στοιχεία έχουν την ίδια απόλυτη τιμή. Σημειώνουμε ότι $x_k \neq 0$ αφού $x \neq 0$. Επίσης,

$$|x_j| \leq |x_k|, j=1,2,\dots,n \quad (4.4)$$

Αυτό σημαίνει ότι:

$$|\lambda - a_{kk}| |x_k| = |\lambda x_k - a_{kk} x_k| = \left| \sum_{j=1}^n a_{kj} x_j - a_{kk} x_k \right| = \left| \sum_{j=1, j \neq k}^n a_{kj} x_j \right| \leq |x_k| R_k \quad (4.5)$$

το οποίο, αν διαιρέσουμε με $|x_k|$, δείχνει ότι η λ βρίσκεται στον δίσκος Gerschgorin D_k με ακτίνα R_k και κέντρο το a_{kk} . Άρα, $\lambda \in D = \bigcup_{i=1}^n D_i$.

ΘΕΩΡΗΜΑ 4.3: (2^ο Θεώρημα Gerschgorin) Έστω $n \geq 2$. Υποθέτουμε ότι $1 \leq p \leq n-1$ και ότι οι δίσκοι Gerschgorin του πίνακα $A \in \mathbb{C}^{n \times n}$ μπορούν να χωριστούν σε δύο ξένα υποσύνολα $D^{(p)}$ και $D^{(q)}$, τα οποία περιέχουν p και $q = n-p$ τέτοιους δίσκους αντίστοιχα. Στη συνέχεια, η ένωση των δίσκων του συνόλου $D^{(p)}$ περιέχει p ιδιοτιμές και η ένωση των δίσκων του συνόλου $D^{(q)}$ περιέχει $n-p$ ιδιοτιμές. Ιδιαίτερα, αν ένας δίσκος είναι ξένος σε σχέση με τους άλλους, περιέχει ακριβώς μία ιδιοτιμή και αν όλοι οι δίσκοι είναι ξένοι μεταξύ τους, τότε ο κάθε δίσκος περιέχει ακριβώς μία ιδιοτιμή.

Απόδειξη: Πρόκειται να χρησιμοποιήσουμε το λεγόμενο ομοτοπικό εγχείρημα. Για $0 \leq \varepsilon \leq 1$, θεωρούμε τον πίνακα $B(\varepsilon) = (b_{ij}(\varepsilon)) \in \mathbb{C}^{n \times n}$ όπου:

$$b_{ij}(\varepsilon) = a_{ii} \text{ αν } i = j \text{ και } b_{ij}(\varepsilon) = \varepsilon a_{ij} \text{ αν } i \neq j \quad (4.6)$$

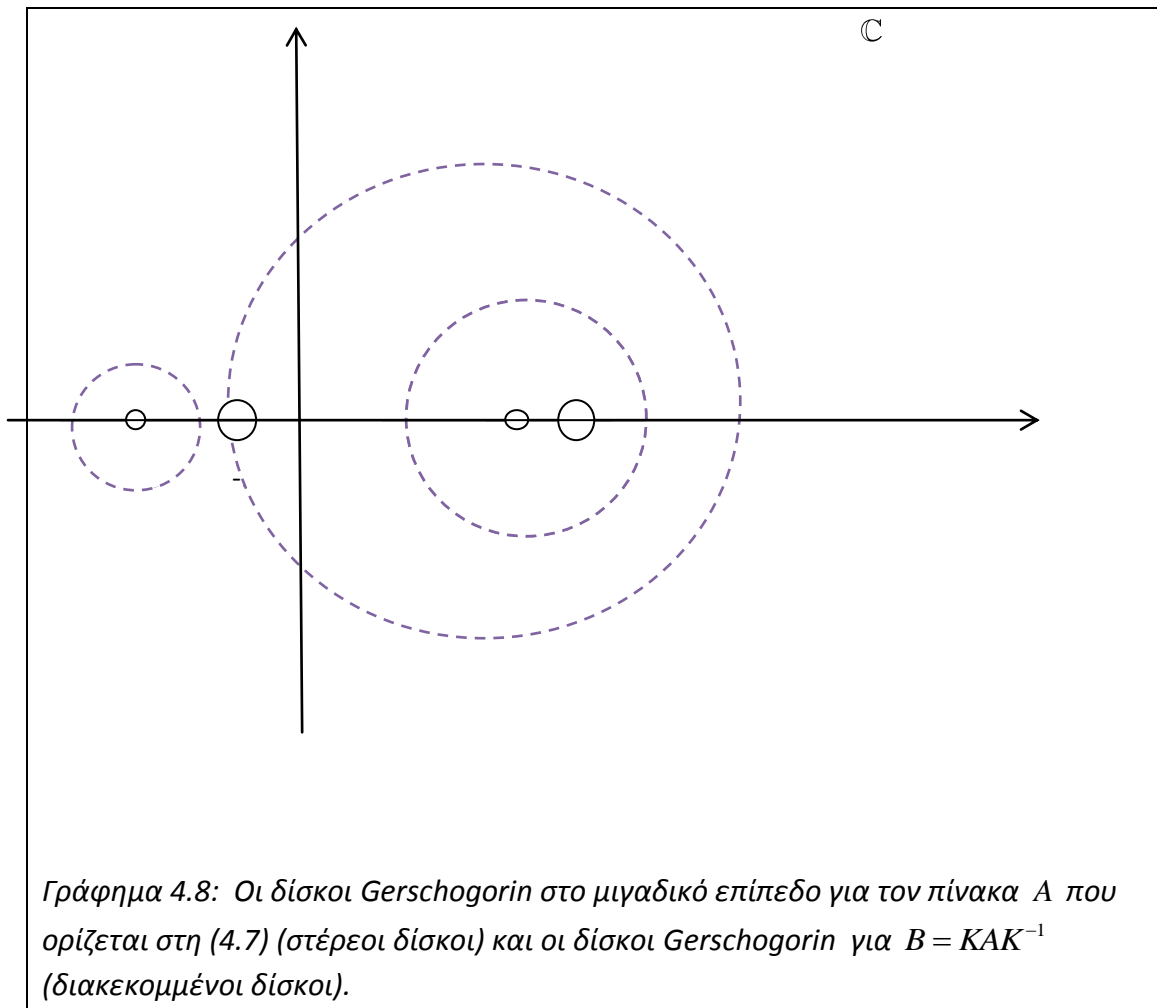
Στη συνέχεια θέτουμε $B(1) = A$ και $B(0)$ είναι ο διαγώνιος πίνακας του οποίου τα διαγώνια στοιχεία του ταυτίζονται με αυτά του πίνακα A . Κάθε μία από τις ιδιοτιμές του $B(0)$ είναι συνεπώς το κέντρο καθενός από τους δίσκους Gerschgorin του πίνακα A . Έτσι ακριβώς p από τις ιδιοτιμές του $B(0)$ βρίσκεται στην ένωση των δίσκων του συνόλου $D^{(p)}$. Τώρα, οι ιδιοτιμές του $B(\varepsilon)$ είναι οι ρίζες του χαρακτηριστικού πολυωνύμου, το οποίο είναι ένα πολυώνυμο του οποίου οι συντελεστές είναι συνεχείς συναρτήσεις του ε . Συνεπώς, οι ρίζες του πολυωνύμου είναι επίσης συνεχείς συναρτήσεις του ε . Έτσι, όσο το ε πηγαίνει από το 0 στο 1 οι ιδιοτιμές του $B(\varepsilon)$ κινούνται κατά μήκος συνεχών μονοπατιών στο μιγαδικό επίπεδο και την ίδια στιγμή οι ακτίνες του δίσκων Gerschgorin αυξάνονται από το μηδέν στις ακτίνες των δίσκων Gerschgorin του πίνακα A . Αφού p από τις ιδιοτιμές βρίσκονται στην ένωση των δίσκων του συνόλου $D^{(p)}$ όταν $\varepsilon = 0$, αυτοί οι δίσκοι είναι ξένοι προς όλους τους δίσκους του συνόλου, αυτές οι p ιδιοτιμές πρέπει ακόμα να βρίσκονται στην ένωση των δίσκων του συνόλου $D^{(p)}$ όταν $\varepsilon = 1$ και έτσι το θεώρημα έχει αποδειχθεί.

Τέλος , μπορούμε να αναφέρουμε , ότι η ίδια απόδειξη ισχύει και όταν οι δίσκοι χωρίζονται σε περισσότερα από δύο ξένα υποσύνολα. •

ΠΑΡΑΔΕΙΓΜΑ 4.4: Θεωρούμε τον πίνακα

$$A = \begin{pmatrix} 4.00 & 0.20 & -0.10 & 0.10 \\ 0.20 & -1.00 & -0.10 & 0.05 \\ -0.10 & -0.10 & 3.00 & 0.10 \\ 0.10 & 0.05 & 0.10 & -3.00 \end{pmatrix} \quad (4.7)$$

Το γράφημα 4.8 δείχνει , σαν στέρεους δίσκους , τους δίσκους Gerschgorin του παραπάνω πίνακα. Για παράδειγμα , ένα από αυτούς τους δίσκους έχει κέντρο το 4.00 και ακτίνα το 0.40 . Οι δίσκοι είναι ξεκάθαρα ξένοι μεταξύ τους , έτσι κάθε δίσκος περιέχει μία ιδιοτιμή του πίνακα. Η έννοια των διακεκομμένων κύκλων που εμφανίζονται στο γράφημα (4.8) εξηγείται στο επόμενο παράδειγμα.



ΠΑΡΑΔΕΙΓΜΑ 4.5: Ας θεωρήσουμε ότι έχουμε τον πίνακα A που ορίζεται στην (4.7) και στη συνέχεια τον μετασχηματίζουμε σε $B = KAK^{-1}$ όπου $K \in \mathbb{R}^{4 \times 4}$ είναι ένας πίνακας που είναι όπως ο μοναδιαίος εκτός από το στοιχείο $k_{22} = \kappa > 0$.

Αυτός ο μετασχηματισμός έχει σαν αποτέλεσμα τον πολλαπλασιασμό των στοιχείων στη γραμμή 2 με κ και τον πολλαπλασιασμό των στοιχείων στη στήλη 2 με $1/\kappa$. Το διαγώνιο στοιχείο a_{22} , έτσι, παραμένει το ίδιο. Μικρότερη τιμή του κ σημαίνει ότι ο δεύτερος δίσκος του B είναι μικρότερος από το δεύτερο δίσκο του A , αλλά σχεδιάζεται μεγαλύτερος. Οι διακεκομμένοι δίσκοι του γραφήματος 4.8 είναι για τον πίνακα B με $\kappa = 1/23$. Για αυτή την τιμή οι άλλοι δίσκοι είναι ξένοι με το δίσκο κέντρου -1.00 . Ο δίσκος με κέντρο 4.00 ίσα που ακουμπάει το δίσκο με κέντρο -1.00 . Ο δίσκος με κέντρο το -1.00 και ακτίνα 0.014 και είναι πολύ μικρός για να είναι ορατός στο γράφημα. Η ιδιοτιμή σε αυτό το δίσκο είναι η -1.009 . Η ίδια διαδικασία μπορεί να χρησιμοποιηθεί για να ελαττώσει το μέγεθος για κάθε έναν από τους δίσκους στη σειρά. •

Η παραπάνω ιδέα επισημοποιείται στο επόμενο θεώρημα.

ΘΕΩΡΗΜΑ 4.6: Έστω $n \geq 2$ και υποθέτουμε ότι στον πίνακα $A \in \mathbb{C}^{n \times n}$ όλα τα μη-διαγώνια στοιχεία του είναι μικρότερα, κατά απόλυτη τιμή, από το ε , δηλαδή $|a_{ij}| < \varepsilon$ για όλα τα $i, j \in \{1, 2, \dots, n\}$ με $i \neq j$. Υποθέτουμε, επίσης, ότι για συγκεκριμένο ακέραιο $r \in \{1, 2, \dots, n\}$ το διαγώνιο στοιχείο a_{rr} είναι κατά δ απομακρυσμένο από τα υπόλοιπα διαγώνια στοιχεία, δηλαδή, $|a_{rr} - a_{ii}| > \delta$ για όλα τα i με $i \neq r$. Στη συνέχεια υπό τον όρο ότι

$$\varepsilon < \frac{\delta}{2(n-1)} \tag{4.9}$$

υπάρχει μια ιδιοτιμή λ του πίνακα A τέτοια ώστε:

$$|\lambda - a_{rr}| < \frac{2(n-1)}{\delta} \varepsilon^2 \tag{4.10}$$

Απόδειξη: Θα εφαρμόσουμε το **μετασχηματισμό ομοιότητας**

$$A \in \mathbb{C}^{n \times n} \mapsto A' = KAK^{-1} \in \mathbb{C}^{n \times n},$$

όπου $K \in \mathbb{R}^{n \times n}$ είναι ίδιος με τον μοναδιαίο πίνακα εκτός από το διαγώνιο στοιχείο της γραμμής r το οποίο διαλέγεται να είναι το $k_{rr} = \kappa > 0$. Αυτό έχει σαν αποτέλεσμα τον πολλαπλασιασμό των μη-διαγώνιων στοιχείων της γραμμής r με κ και το στοιχείο της στήλης r της γραμμής i , με το $1/\kappa$. Ο δίσκος Gerschgorin από

τη γραμμή r έχει κέντρο το a_{rr} και ακτίνα που δεν ξεπερνά την ποσότητα $k(n-1)\varepsilon$ ενώ ο δίσκος αντίστοιχα της γραμμής $i \neq r$ έχει κέντρο το a_{ii} και ακτίνα που δεν ξεπερνά την ποσότητα $(n-2)\varepsilon + \varepsilon / \kappa$.

Τώρα θέλουμε να μικρύνουμε το μέγεθος του δίσκου r επιλέγοντας μία μικρή τιμή του κ , ενώ τον κρατάμε ξένο από το υπόλοιπο πλήθος των δίσκων. Αυτό είναι εύκολο να γίνει επιλέγοντας $\kappa = 2\varepsilon / \delta$. Η ακτίνα του δίσκου r δεν ξεπερνά την ποσότητα $2(n-1)\varepsilon^2 / \delta$ και η ακτίνα του δίσκου $i \neq r$ δεν ξεπερνά την ποσότητα $(n-2)\varepsilon + \frac{1}{2}\delta$.

Το άθροισμα αυτών των ακτίνων, συνεπώς, ικανοποιεί το:

$$R_r + R_i \leq 2(n-1)\varepsilon^2 / \delta + (n-2)\varepsilon + \frac{1}{2}\delta \tag{4.11}$$

$$< \varepsilon + (n-2)\varepsilon + \frac{1}{2}\delta < \delta$$

όπου έχουμε χρησιμοποιήσει τη συνθήκη (4.9) δύο φορές. Όσο τα κέντρα a_{rr} και a_{ii} αυτών των δίσκων απέχουν περισσότερο από το δ το ένα με το άλλο, η σχέση (4.11) δείχνει ότι οι δύο δίσκοι είναι μεταξύ τους ξένοι. Έτσι, αποδείχθηκε το ζητούμενο. •

Το Θεώρημα 4.6 είναι επαρκές για να δείξει ότι για ένα πίνακα, για τον οποίο επιτυγχάνεται η υπόθεσή του θεωρήματος, μπορούμε να βρούμε ένα δίσκο Gerschgorin του οποίου η ακτίνα είναι της τάξης του ε^2 , υπό τον όρο ότι το ε είναι αρκετά μικρό. Επίσης, δείχνεται ότι το διάστημα ανάμεσα στα διαγώνια στοιχεία είναι σημαντικό.

Ιδιαίτερως, το Θεώρημα 4.6 εφαρμόζεται στον πίνακα $A^{(k)}$, ο οποίος είναι αποτέλεσμα της μεθόδου Jacobi μετά από k επαναλήψεις. Αν σε αυτό το στάδιο όλα τα μη διαγώνια στοιχεία έχουν μέγεθος μικρότερο του ε τότε υπάρχει μία ιδιοτιμή σε κάθε ένα από τα διαστήματα $[a_{ii}^{(k)} - (n-1)\varepsilon, a_{ii}^{(k)} + (n-1)\varepsilon]$, υπό την προϋπόθεση ότι αυτά τα διαστήματα είναι ξένα μεταξύ τους. Αυτό όμως έπεται από το Θεώρημα 4.6. Εάν το ε είναι αρκετά μικρό, συγκρινόμενο με τις αποστάσεις ανάμεσα στα διαγώνια στοιχεία του πίνακα $A^{(k)}$ το Θεώρημα 4.6 μπορεί να χρησιμοποιηθεί για να δώσει κλειστότερα φράγματα για τις ιδιοτιμές.

Παραθέτουμε, επίσης, μερικά σχόλια για την σύγκλιση της κλασικής Jacobi επανάληψης. Σύμφωνα με την ανισότητα Cauchy-Schwarz,

$$\left(\sum_{i,j=1,i \neq j}^n |a_{ij}^{(k)}| \right)^2 \leq \sum_{i,j=1,i \neq j}^n 1^2 \sum_{i,j=1,i \neq j}^n |a_{ij}^{(k)}|^2 = n(n-1) \sum_{i,j=1,i \neq j}^n |a_{ij}^{(k)}|^2$$

Επομένως,

$$\left(\max_{i=1}^n \sum_{i,j=1,i \neq j}^n |a_{ij}^{(k)}| \right)^2 \leq n(n-1) \sum_{i,j=1,i \neq j}^n |a_{ij}^{(k)}|^2$$

έτσι από την σχέση (3.10) συνεπάγεται ότι ,

$$\lim_{k \rightarrow \infty} \max_{i=1}^n \sum_{j=1,j \neq i}^n |a_{ij}^{(k)}| = 0 .$$

Με απλά λόγια , οι ακτίνες των δίσκων Gerschgorin για τους πίνακες της ακολουθίας $(A^{(k)})$ συγκλίνουν στο μηδέν όταν $k \rightarrow \infty$. Όσο ο $A^{(k)}$ και ο A έχουν ταυτιζόμενες ιδιοτιμές για όλα τα k , έπεται από τα Θεωρήματα 4.2 και 4.3 ότι το σύνολο των ορίων των διαγώνιων στοιχείων δηλαδή το $\left\{ \lim_{k \rightarrow \infty} a_{11}^{(k)}, \dots, \lim_{k \rightarrow \infty} a_{nn}^{(k)} \right\}$, το οποίο δίνεται από την Jacobi επανάληψη είναι ισοδύναμο με το σύνολο των ιδιοτιμών του πίνακα A . Αυτό διατηρείται ανεξαρτήτως του διαστήματος ανάμεσα στα διαγώνια στοιχεία.

Στο τελευταίο μέρος αυτού του κεφαλαίου θα αναφέρουμε σε θεωρητικό επίπεδο κάποιες ακόμα ιδιότητες καθώς και το τρίτο Θεώρημα του Gerschgorin.

Παρατήρηση: Το σύνολο των ιδιοτιμών ενός πίνακα λέγεται **φάσμα** και συμβολίζεται με $\sigma(A)$.

Από τα προηγούμενα έχουμε ότι

$$\sigma(A) \subseteq D = \bigcup_{i=1}^n D_i \quad , \quad \text{όπου} \quad D_i = \{z \in \mathbb{C} : |z - a_{ii}| \leq R_i\} \quad (4.12)$$

Η παραπάνω σχέση σε συνδυασμό με το Θεώρημα 4.2 διασφαλίζει ότι κάθε ιδιοτιμή του πίνακα A βρίσκεται στην ένωση των κύκλων Gerschgorin D_i . Επίσης , επειδή ο A και A^T μοιράζονται το ίδιο φάσμα η παραπάνω σχέση μπορεί να γραφεί στη μορφή

$$\sigma(A) \subseteq C = \bigcup_{j=1}^n C_j \quad , \quad \text{όπου} \quad C_j = \left\{ z \in \mathbb{C} : |z - a_{jj}| \leq \sum_{i=1,i \neq j}^n |a_{ij}| \right\} \quad (4.13)$$

Οι κύκλοι D_i στο μιγαδικό επίπεδο καλούνται κύκλοι γραμμής και οι κύκλοι C_j κύκλοι στήλης. Η άμεση συνέπεια από τις σχέσεις (4.12) και (4.13) είναι το ακόλουθο Θεώρημα.

ΘΕΩΡΗΜΑ 4.7: Για ένα δοθέν πίνακα $A \in \mathbb{C}^{n \times n}$, ισχύει

$$\forall \lambda \in \sigma(A), \quad \lambda \in D \cap C \quad (4.14)$$

ΠΑΡΑΤΗΡΗΣΗ 4.8: Το Θεώρημα του Gershgorin (Θεώρημα 4.2) καθώς και το Θεώρημα (4.7) δεν αποκλείουν την πιθανότητα ότι υπάρχουν κύκλοι που δεν περιέχουν ιδιοτιμές.

ΟΡΙΣΜΟΣ 4.9: Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ καλείται αναγώγιμος (*reducible*) αν υπάρχει ένας πίνακας μεταθέσεων P τέτοιος ώστε

$$PAP^T = \begin{bmatrix} B_{11} & B_{12} \\ 0 & B_{22} \end{bmatrix},$$

όπου B_{11} και B_{22} είναι τετραγωνικοί πίνακες. Ο πίνακας A είναι μη αναγώγιμος ή ανάγωγος (*irreducible*) αν δεν είναι αναγώγιμος.

Για να δούμε αν ένας πίνακας είναι αναγώγιμος τότε ο πιο βολικός τρόπος είναι το **προσανατολισμένο γράφημα** (*oriented graph*) του πίνακα. Το προσανατολισμένο γράφημα ενός πραγματικού πίνακα A προκύπτει από την ένωση n στοιχείων P_1, \dots, P_n τραβώντας μία γραμμή από το P_i στο P_j αν το αντίστοιχο στοιχείο του πίνακα είναι διάφορο του μηδενός δηλαδή αν $a_{ij} \neq 0$. Ένα προσανατολισμένο γράφημα είναι **ισχυρά συνδεδεμένο** αν για κάθε διακεκριμένα P_i και P_j υπάρχει ένα κατευθυνόμενο μονοπάτι από το P_i και P_j .

ΘΕΩΡΗΜΑ 4.10: Ένας πίνακας $A \in \mathbb{C}^{n \times n}$ είναι μη αναγώγιμος αν το προσανατολισμένο γράφημα του είναι ισχυρά συνδεδεμένο.

ΘΕΩΡΗΜΑ 4.11: (3^ο Θεώρημα Gershgorin) Έστω ο πίνακας $A \in \mathbb{C}^{n \times n}$ είναι ένας μη αναγωγίμος πίνακας. Μία ιδιοτιμή $\lambda \in \sigma(A)$ (όπου $\sigma(A) \subseteq D = \bigcup_{i=1}^n D_i$) δεν μπορεί να βρίσκεται στο σύνορο του D αν δεν ανήκει στο σύνορο κάθε κύκλου D_i για $i = 1, \dots, n$.

5. Η ΜΕΘΟΔΟΣ HOUSEHOLDER

Η γενική μέθοδος για τον υπολογισμό των ιδιοτιμών ενός πραγματικού συμμετρικού πίνακα ξεκινάει εφαρμόζοντας ορθογώνιο μετασχηματισμό για να τον μετασχηματίσει σε ένα τριδιαγώνιο πίνακα. Αυτό μπορεί να γίνει σε πεπερασμένο αριθμό βημάτων χρησιμοποιώντας πίνακες Householder.

ΟΡΙΣΜΟΣ 5.1: Δοθέντος ενός διανύσματος $u \in \mathbb{R}_*^n$, ο αντίστοιχος **Householder** πίνακας $H = H(u)$ τάξης \mathbb{R}^n ορίζεται από τον παρακάτω τύπο

$$H = I - \frac{2}{u^T u} uu^T \quad (5.1)$$

όπου I είναι ο μοναδιαίος πίνακας τάξης n .

Ολοφάνερα, για κάθε διάνυσμα $x \in \mathbb{R}^n$, έχουμε

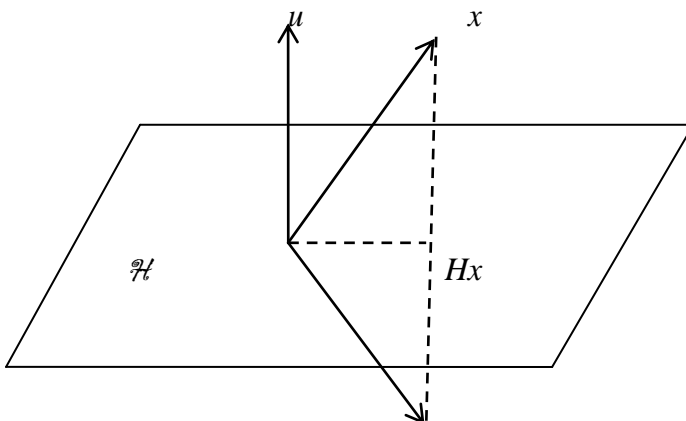
$$Hx = x - 2 \frac{u^T x}{u^T u} u \quad (5.2)$$

και έτσι τα διανύσματα Hx , x και u είναι συνεπίεδα. Ιδιαίτερως, αν $x \in \mathbb{R}^n$ και $u^T x = 0$ τότε $Hx = x$ και συνεπώς το $(n-1)$ -διάστασης υπερεπίπεδο \mathcal{H} , που απαρτίζεται από όλα τα διανύσματα x που είναι κάθετα στο u στον \mathbb{R}^n , μένει αμετάβλητο κάτω από την απεικόνιση $x \mapsto Hx$. Τέλος, για κάθε διάνυσμα $x \in \mathbb{R}^n$,

$$u^T Hx = -u^T x \quad (5.3)$$

Έτσι, αν η γωνία των διανυσμάτων x και u συμβολίζεται με φ , τότε η γωνία των διανυσμάτων u και Hx θα είναι ίση με $\pi + \varphi$. Συμπεραίνουμε από αυτές τις παρατηρήσεις ότι το διάνυσμα Hx είναι η ανάκλαση του x στο υπερεπίπεδο \mathcal{H} .

Γι' αυτό το λόγο, η απεικόνιση $x \mapsto Hx$ συχνά αναφέρεται ως **Householder μετασχηματισμός**, αντίστοιχα με το διάνυσμα $u \in \mathbb{R}_*^n$ (βλέπε γράφημα 5.4)



Γράφημα 5.4

Στο γράφημα αυτό παρατηρούμε την επίδραση του Householder μετασχηματισμού $H : x \mapsto Hx$, σε αντιστοιχία με το $u \in \mathbb{R}_*^n$, στο διάνυσμα $x \in \mathbb{R}^n$. Το Hx είναι η αντανάκλαση του x στο υπερεπίπεδο \mathcal{H} που είναι κάθετο στο u .

ΛΗΜΜΑ 5.2: Κάθε πίνακας Householder είναι συμμετρικός και ορθογώνιος.

Απόδειξη: Επειδή $I^T = I$, $(uu^T)^T = (u^T)^T u^T = uu^T$ και $u^T u$ είναι ένας (θετικός πραγματικός) αριθμός, συνεπάγεται ότι H συμμετρικός. Η ορθογωνιότητα του πίνακα H είναι συνέπεια από την ταυτοτική

$$H^T H = HH^T = H^2 = I - \frac{4}{u^T u} uu^T + \frac{4}{(u^T u)^2} (uu^T)(uu^T) = I$$

αφού $(uu^T)(uu^T) = u(u^T u)u^T = (u^T u)uu^T$ από την προσεταιριστικότητα του πολλαπλασιασμού πινάκων. •

ΛΗΜΜΑ 5.3: Έστω $1 \leq k < n$ και υποθέτουμε ότι H_k είναι ένας $k \times k$ πίνακας Householder. Στη συνέχεια, ο πίνακας $H \in \mathbb{R}^{n \times n}$, ο οποίος γράφεται σε σύνθετη μορφή ως:

$$H = \begin{pmatrix} I_{n-k} & 0 \\ 0^T & H_k \end{pmatrix}$$

όπου I_{n-k} είναι ο μοναδιαίος πίνακας τάξης $n-k$ και ο 0 είναι ένας $(n-k) \times k$ μηδενικός πίνακας, είναι επίσης πίνακας Householder.

(Η απόδειξη αυτού του λήμματος είναι απλή και άμεση).

ΛΗΜΜΑ 5.4: Δοθέντος ενός διανύσματος $x \in \mathbb{R}_*^n$, υπάρχει ένας πίνακας Householder $H \in \mathbb{R}_{sym}^{n \times n}$ τέτοιος ώστε όλα τα στοιχεία του διανύσματος Hx είναι μηδέν εκτός από το πρώτο, δηλαδή Hx είναι ένα μη μηδενικό πολλαπλάσιο του e_1 , με e_1 να είναι η πρώτη στήλη του μοναδιαίου πίνακα.

Σε γεωμετρικούς όρους αυτό το αποτέλεσμα μπορεί να αναδιατυπωθεί λέγοντας ότι για ένα διάνυσμα $x \in \mathbb{R}_*^n$ υπάρχει ένα $(n-1)$ -διάστασης υπερεπίπεδο \mathcal{H} που διέρχεται από την αρχή των αξόνων στο \mathbb{R}^n τέτοιο ώστε η ανάκλαση Hx του x στο υπερεπίπεδο \mathcal{H} να είναι ισοδύναμη με το μη μηδενικό πολλαπλάσιο του e_1 . Για να βρούμε το υπερεπίπεδο \mathcal{H} αρκεί να προσδιορίσουμε το διάνυσμα $u \in \mathbb{R}_*^n$ που είναι κάθετο στο \mathcal{H} . Αφού \mathcal{H} είναι ανεπηρέαστο από την αναγωγή του u (βλέπε ορισμό 5.1), το μέτρο του u δεν παίζει ρόλο στην κατασκευή. Όπως σημειώθηκε στην συζήτηση που ακολούθησε τον ορισμό 5.1, τα διανύσματα Hx , x και u είναι συνεπίπεδα. Άρα, πρόκειται να ψάξουμε το $u \in \mathbb{R}_*^n$, σαν κατάλληλο γραμμικό συνδυασμό του x και e_1 .

Απόδειξη: (λήμματος) Ψάχνουμε πίνακα $H = I - \left[2 / (u^T u) \right] uu^T$ με $u = x + ce_1$, όπου c ένας μη μηδενικός πραγματικός αριθμός που θα πρέπει να οριστεί. Συνεπώς,

$$\begin{aligned} u^T x &= x^T x + c\beta \\ u^T u &= x^T x + 2c\beta + c^2 \end{aligned}$$

όπου $\beta = e_1^T x$ είναι η πρώτη συντεταγμένη του x . Εκτελώντας υπολογισμούς λαμβάνουμε ότι

$$Hx = x - \frac{2}{u^T u} u (u^T x) = \frac{(c^2 - x^T x)x - 2c(x^T x + c\beta)e_1}{x^T x + 2c\beta + c^2}$$

Έτσι, Hx θα είναι πολλαπλάσιο του e_1 με την προϋπόθεση ότι διαλέγουμε το c έτσι ώστε $c^2 = x^T x$. Επίσης, για να αποφύγουμε την διαίρεση υπό το μηδέν, χρειάζεται να διασφαλίσουμε ότι $x^T x + 2c\beta + c^2 \neq 0$. Επίσης, το x θα είναι της μορφής $x = (\beta, x_2, \dots, x_{n-1})$ (αφού παραπάνω αναφέραμε ότι το β είναι η πρώτη συντεταγμένη του x και το έχουμε ορίσει και τα υπόλοιπα στοιχεία $x_2, \dots, x_{n-1} \in \mathbb{R}$) και έτσι έχουμε $x^T x = \beta^2 + x_2^2 + \dots + x_{n-1}^2$. Άρα, ισχύει ότι $c^2 = (x^T x)^2 = \beta^2 + x_2^2 + \dots + x_{n-1}^2 \geq \beta^2$. Άρα σημειώνουμε ότι $c^2 \geq \beta^2$ και συνεπώς

$$x^T x + 2c\beta + c^2 \geq \beta^2 + 2c\beta + c^2 = (\beta + c)^2 \neq 0$$

με την προϋπόθεση ότι $\beta + c \neq 0$, το οποίο μπορούμε να διασφαλίσουμε επιλέγοντας το κατάλληλο πρόσημο για το c το οποίο ορίζεται

$$c = (\text{sign}\beta) \sqrt{x^T x} \text{ όταν } \beta \neq 0 \text{ και } c = \sqrt{x^T x} \text{ όταν } \beta = 0.$$

Με αυτήν την επιλογή του c , έχουμε ότι $Hx = -ce_1$, όπως απαιτείται και έτσι η απόδειξη έχει τελειώσει. •

Στη συνέχεια θα δούμε πως οι πίνακες Householder μπορούν να χρησιμοποιηθούν για να μετασχηματίσουμε ένα δοθέν πίνακα σε τριδιαγώνια μορφή.

ΘΕΩΡΗΜΑ 5.5: Δίνεται πίνακας $A \in \mathbb{R}_{sym}^{n \times n}$ και $n \geq 3$, τότε υπάρχει ένας πίνακας $Q_n \in \mathbb{R}_{sym}^{n \times n}$, που είναι γινόμενο από $n-2$ πίνακες Householder $H_{(n,k)} \in \mathbb{R}_{sym}^{n \times n}$ $k=1,2,\dots,n-1$, που δίνεται από τον τύπο

$$Q_n = H_{(n,n-1)}H_{(n,n-2)}\dots H_{(n,2)}$$

τέτοιος ώστε ο $Q_n^T A Q_n = T_n$ να είναι τριδιαγώνιος. Ο πίνακας Q_n είναι ορθογώνιος.

Απόδειξη: Η απόδειξη του θεωρήματος θα προχωρήσει με επαγωγή. Πριν προχωρήσουμε σε αυτή, θα κάνουμε κάποιες παρατηρήσεις οι οποίες αναδεικνύουν το κλειδί για την απόδειξη.

Θεωρούμε τον πίνακα $A \in \mathbb{R}_{sym}^{n \times n}$, ο οποίος είναι διαμερισμένος από την πρώτη γραμμή και στήλη στην μορφή

$$A = \begin{pmatrix} a & b^T \\ b & C \end{pmatrix}$$

όπου $a \in \mathbb{R}$, $b \in \mathbb{R}^{n-1}$ και $C \in \mathbb{R}_{sym}^{(n-1) \times (n-1)}$ και ορίζουμε

$$E_1^T = \{u \in \mathbb{R}^n : u = (\lambda, 0, \dots, 0)^T, \lambda \in \mathbb{R}\}.$$

Αν b ανήκει στο \mathbb{R}_*^{n-1} , τότε, από το Λήμμα 5.4, υπάρχει ένας $(n-1) \times (n-1)$ Householder πίνακας H_{n-1} τέτοιος ώστε κάθε στοιχείο του $H_{n-1}b$, εκτός από το πρώτο, να είναι ίσο με το μηδέν. Αν $b=0$ τότε $H_{n-1}b=0$. Έτσι και αλλιώς $H_{n-1}b \in E_1^{n-1}$.

Ας επεκτείνουμε τον πίνακα Householder $H_{n-1} \in \mathbb{R}_{sym}^{(n-1) \times (n-1)}$, χρησιμοποιώντας το Λήμμα 5.3 με $k=n-1$, σε ένα Householder πίνακα $H_{(n,n-1)} \in \mathbb{R}_{sym}^{n \times n}$, ορίζοντας το $(1,1)$ -στοιχείο του $H_{(n,n-1)}$ ίσο με ένα και επιλέγοντας τα υπόλοιπα στοιχεία στην πρώτη γραμμή και στην πρώτη στήλη του $H_{(n,n-1)}$ ίσα με το μηδέν. Έτσι,

$$\begin{aligned} H_{(n,n-1)}^T A H_{(n,n-1)} &= \begin{pmatrix} 1 & 0^T \\ 0 & H_{n-1}^T \end{pmatrix} \begin{pmatrix} a & b^T \\ b & C \end{pmatrix} \begin{pmatrix} 1 & 0^T \\ 0 & H_{n-1} \end{pmatrix} \\ &= \begin{pmatrix} a & d^T \\ d & D \end{pmatrix} \end{aligned} \tag{5.5}$$

όπου $d = H_{n-1}^T b = H_{n-1} b \in E_1^{n-1}$ και $D = H_{n-1}^T C H_{n-1} \in \mathbb{R}_{sym}^{(n-1) \times (n-1)}$.

Όσο $d \in E_1^{n-1}$, η πρώτη γραμμή και η πρώτη στήλη του $H_{(n,n-1)}^T A H_{(n,n-1)}$ είναι εκτός της επιθυμητής μορφής. Έτσι απομένει να μετασχηματίσουμε τον υποπίνακα D σε τριδιαγώνια μορφή. Αυτό θα επιτευχθεί προχωρώντας επαγωγικά.

Αν $n=3$, τότε ο 3×3 πίνακας $H_{(n,n-1)}^T A H_{(n,n-1)}$ είναι αυτομάτως τριδιαγώνιος επειδή $d \in E_1^2$ και συμπληρώνουμε την απόδειξη παίρνοντας ότι $Q_3 = H_{(3,2)}$. Και σημειώνουμε, τέλος, ότι αν $d \in E_1^3$ τότε

$$Q_3^T f = H_{(3,2)}^T f = H_{(3,2)} f \in E_1^3$$

καθώς το (1,1)-στοιχείο του $H_{(3,2)}$ είναι ίσο με ένα και τα υπόλοιπα στοιχεία της πρώτης στήλης είναι όλα μηδέν.

Ας υποθέσουμε τώρα ότι $n \geq 4$ και $A \in \mathbb{R}_{sym}^{n \times n}$. Η επαγωγική μας υπόθεση είναι ότι το περιεχόμενο του θεωρήματος έχει ήδη θεμελιωθεί για κάθε πραγματικό συμμετρικό πίνακα τάξης $n-1$, δηλαδή $D \in \mathbb{R}_{sym}^{(n-1) \times (n-1)}$ μπορεί να μετασχηματιστεί σε τριδιαγώνια μορφή:

$$Q_{n-1}^T D Q_{n-1} = T_{n-1},$$

όπου $Q \in \mathbb{R}^{(n-1) \times (n-1)}$ είναι ένας ορθογώνιος πίνακας ο οποίος είναι το γινόμενο από $n-3$ πίνακες Householder, καθένας από τους οποίους είναι μεγέθους $(n-1) \times (n-1)$:

$$Q_{n-1} = H_{(n-1,n-2)} \cdots H_{(n-1,2)}$$

και $Q_{n-1}^T f \in E_1^{n-1}$ για κάθε διάνυσμα $f \in E_1^{n-1}$. Έτσι, αυτή η επαγωγική υπόθεση έχει επαληθευτεί για 3×3 πραγματικούς συμμετρικούς πίνακες.

Τώρα, θα επεκτείνουμε καθένα από τους $(n-1) \times (n-1)$ πίνακες $H_{(n-1,k)}$ για $k=2, \dots, n-2$, σε $n \times n$ πίνακες Householder $H_{(n,k)}$, $k=2, \dots, n-2$ αντίστοιχα, σύμφωνα με το Λήμμα 5.3, και ορίζουμε

$$Q_n = H_{(n,n-1)} H_{(n,n-2)} \cdots H_{(n,2)}.$$

Έτσι από την (5.5) έχουμε,

$$\begin{aligned} Q_n^T A Q_n &= \begin{pmatrix} 1 & 0^T \\ 0 & Q_{n-1}^T \end{pmatrix} \begin{pmatrix} a & d^T \\ d & D \end{pmatrix} \begin{pmatrix} 1 & 0^T \\ 0 & Q_{n-1} \end{pmatrix} \\ &= \begin{pmatrix} a & d^T Q_{n-1} \\ Q_{n-1}^T d & T_{n-1} \end{pmatrix} \end{aligned}$$

Όσο $d \in E_1^{n-1}$, παρατηρείται από την επαγωγική υπόθεση ότι $Q_{n-1}^T d$ επίσης ανήκει στο E_1^{n-1} και επομένως ο τελευταίος πίνακας είναι τριδιαγώνιος. Έτσι, αν Q_n είναι το γινόμενο από $n-2$ πίνακες Householder, καθένας είναι μεγέθους $n \times n$ και

ορθογώνιοι, τότε $Q_n \in \mathbb{R}^{n \times n}$ είναι ορθογώνιος. Επιπλέον, για κάθε διάνυσμα $f \in E_1^n$ έχουμε $Q_n^T f \in E_1^n$, επειδή το $(1,1)$ -στοιχείο του Q_n είναι το ένα και τα υπόλοιπα στοιχεία της πρώτης στήλης του Q_n είναι ίσα με το μηδέν. Αυτό περιέχει το επαγωγικό βήμα και έτσι ολοκληρώθηκε η απόδειξη. •

Ο αναδρομικός μετασχηματισμός ενός συμμετρικού πίνακα σε τριδιαγώνια μορφή αποτυπώνεται στην απόδειξη του Θεωρήματος 5.5 και καλείται **μέθοδος Householder**. Για να εφαρμοστεί αυτή η μέθοδος πρακτικά είναι σημαντικό να εκτελέσουμε τους μετασχηματισμούς αποτελεσματικά. Λογαριάζοντας τις αριθμητικές πράξεις που εμπεριέχονται είναι απλό και δείχνει ότι η τέλεια αναγωγή απαιτεί προσεγγιστικά $\frac{1}{3}n^3$ πολλαπλασιασμούς για μία μέτρια μεγάλη τιμή του n .

Ακριβώς παρακάτω παραθέεται ο αλγόριθμος που περιγράφει την μέθοδο Householder.

ΑΛΓΟΡΙΘΜΟΣ HOUSEHOLDER:

Δίνεται ένας πίνακας $A_0 \in \mathbb{R}_{sym}^{n \times n}$.

Βήμα 1^ο : Δοθέντος ενός διανύσματος $u \in \mathbb{R}_*^n$, υπολογίζουμε το γινόμενο από Householder πίνακες : $Q_n = H_{(n,n-1)}H_{(n,n-2)} \dots H_{(n,2)}$ και στη συνέχεια υπολογίζουμε το $Q_n^T A_0 Q_n = A_1$.

Βήμα 2^ο : Όμοια, $Q_{n-1} = H_{(n-1,n-2)}H_{(n-1,n-3)} \dots H_{(n-1,2)}$ και $Q_{n-1}^T A_1 Q_{n-1} = A_2$.

.....

Συνεχίζουμε την ίδια διαδικασία μέχρι να υπολογίσουμε τον πίνακα A_{n-2} και έτσι θέτουμε $A_{n-2} = T$, όπου T θα είναι ο ζητούμενος τριδιαγώνιος πίνακας.

ΠΑΡΑΔΕΙΓΜΑ 5.6: Για να εξηγήσουμε την μέθοδο Householder, γυρίζουμε πίσω στον πίνακα A που ορίζεται στο παράδειγμα 3.8. Το πρώτο στάδιο χρησιμοποιεί τον πίνακα Householder που ορίζεται από το διάνυσμα

$$u = (0.000, 4.162, 2.000, 1.000, 2.000)^T \tag{5.6}$$

Το αποτέλεσμα του μετασχηματισμού είναι ο πίνακας

$$\begin{pmatrix} 4.000 & -3.162 & 0.000 & 0.000 & 0.000 \\ -3.162 & 5.300 & 1.232 & -0.332 & 0.284 \\ 0.000 & 1.232 & 1.653 & 3.312 & 0.275 \\ 0.000 & -0.332 & 3.312 & 5.149 & 1.123 \\ 0.000 & 0.284 & 0.275 & 1.123 & -3.102 \end{pmatrix}$$

Το ηγετικό στοιχείο του πίνακα είναι αμετάβλητο (δηλαδή το (1,1)- στοιχείο) και παρατηρούμε ότι η πρώτη γραμμή και πρώτη στήλη έχουν τριδιαγώνια δομή.

Το δεύτερο στάδιο χρησιμοποιεί τον πίνακα Householder με το διάνυσμα

$$u = (0.000, 0.000, 2.540, -0.332, 0.284)^T \quad (5.7)$$

και δίνει τον νέο πίνακα

$$\begin{pmatrix} 4.000 & -3.162 & 0.000 & 0.000 & 0.000 \\ -3.162 & 5.300 & -1.308 & 0.000 & 0.000 \\ 0.000 & -1.308 & 0.057 & -2.166 & 0.792 \\ 0.000 & 0.000 & -2.166 & 6.610 & 0.420 \\ 0.000 & 0.000 & 0.792 & 0.420 & -2.967 \end{pmatrix}$$

Σε αυτή τη φάση το νέο ηγετικό στοιχείο το 2×2 είναι αμετάβλητο και παρατηρούμε ότι οι δύο πρώτες γραμμές και στήλες έχουν τριδιαγώνια δομή.

Το τελικό στάδιο χρησιμοποιεί τον πίνακα Householder με το διάνυσμα

$$u = (0.000, 0.000, 0.000, -4.471, 0.792)^T \quad (5.8)$$

και δίνει τον τριδιαγώνιο πίνακα

$$\begin{pmatrix} 4.000 & -3.162 & 0.000 & 0.000 & 0.000 \\ -3.162 & 5.300 & -1.308 & 0.000 & 0.000 \\ 0.000 & -1.308 & 0.057 & 2.306 & 0.000 \\ 0.000 & 0.000 & 2.306 & 5.208 & -3.411 \\ 0.000 & 0.000 & 0.000 & -3.411 & -1.565 \end{pmatrix}$$

Οι αριθμητικές τιμές εδώ παραθέτονται με τρία δεκαδικά ψηφία , χάριν απλότητας. •

Έχοντας δείξει πώς ένας συμμετρικός πίνακας μετασχηματίζεται σε τριδιαγώνια μορφή , μπορούμε τώρα να εξετάσουμε το πρόβλημα του προσδιορισμού των ιδιοτιμών ενός τριδιαγώνιου πίνακα.

Στο υπόλοιπο αυτού του κεφαλαίου πρόκειται να θεωρήσουμε ότι όλα τα μη-διαγώνια στοιχεία b_i είναι μη μηδενικά. Γιατί, αν υποθέσουμε ότι $b_k = 0$ για κάποιο k που ανήκει στο $\{2, 3, \dots, n\}$, τότε οι ιδιοτιμές του πίνακα T εμπεριέχουν τις ιδιοτιμές του πίνακα που αποτελείται από τις πρώτες $k-1$ γραμμές και στήλες μαζί με τις ιδιοτιμές του πίνακα που αποτελείται από τις τελευταίες $n-k+1$ γραμμές και στήλες. Αυτά τα δύο προβλήματα εξελίσσονται χωριστά και μπορούν να επεξεργαστούν ανεξάρτητα. Αν αρκετά από τα μη-διαγώνια στοιχεία είναι ίσα με το μηδέν, τότε ο πίνακας μπορεί να χωριστεί σε ένα αριθμό από μικρότερους πίνακες οι οποίοι μετά μπορούν να αντιμετωπιστούν ανεξάρτητα.

ΘΕΩΡΗΜΑ 6.1: (Θεώρημα Cauchy) Έστω $n \geq 3$. Οι ρίζες του p_r διαχωρίζουν αυτές του p_{r+1} για $r=1, 2, \dots, n-1$ δηλαδή ανάμεσα από δύο συνεχόμενες ρίζες του p_{r+1} υπάρχει ακριβώς μία ρίζα του πολυωνύμου p_r , $r=1, 2, \dots, n-1$.

Απόδειξη: Η απόδειξη θα γίνει με επαγωγή. Είναι τετριμμένο να δείξουμε ότι η ιδιότητα αυτή ισχύει για $r=1$: οι δύο ρίζες

$$\frac{1}{2} \left[a_1 + a_2 \pm \sqrt{(a_1 - a_2)^2 + 4b_2^2} \right],$$

του p_2 διαχωρίζονται από το a_1 , που είναι η μοναδική ρίζα του γραμμικού πολυωνύμου p_1 .

Υποθέτουμε ότι η πρόταση ισχύει για $r=i-1$, $2 \leq i \leq n-1$ έτσι ώστε οι ρίζες του p_{i-1} να διαχωρίζουν αυτές του p_i . Συμβολίζουμε με a και β δύο συνεχόμενες ρίζες του p_i , η επαγωγική υπόθεση συνεπάγεται ότι το p_{i-1} έχει ακριβώς μία ρίζα ανάμεσα από το a και β , το οποίο σημαίνει ότι τα $p_{i-1}(a)$ και $p_{i-1}(\beta)$ έχουν αντίθετα πρόσημα. Τώρα έχουμε

$$p_{i+1}(\lambda) = (a_{i+1} - \lambda) p_i(\lambda) - b_{i+1}^2 p_{i-1}(\lambda),$$

τέτοιο ώστε, επειδή a και β είναι ρίζες του p_i , έπεται ότι $p_{i+1}(a)$ και $p_{i+1}(\beta)$ έχουν επίσης αντίθετα πρόσημα. Έτσι, p_{i+1} έχει τουλάχιστον μία ρίζα ανάμεσα στα a και β . Διαλέγοντας a και β να είναι το ζεύγος από συνεχόμενες ρίζες του p_i με τη σειρά, θα έχουμε άρα εντοπίσει $i-1$ ρίζες του p_{i+1} .

Στη συνέχεια διαλέγουμε a ώστε να είναι η αλγεβρικά μικρότερη ρίζα του p_i . Είναι εύκολο να δούμε ότι καθένα από τα πολυώνυμα p_1, p_2, \dots, p_n τείνει στο ∞ όταν το $\lambda \rightarrow -\infty$. Από την επαγωγική υπόθεση, το p_{i-1} δεν έχει ρίζες μικρότερες από το a και έτσι $p_{i-1}(a)$ είναι θετικό. Έτσι, από την επαγωγική σχέση $p_{i+1}(a)$ είναι αρνητικό και συνεπώς p_{i+1} πρέπει να έχει μία ρίζα μικρότερη του a . Ένα παρόμοιο

επιχείρημα δείχνει ότι το p_{i+1} έχει μία ρίζα μεγαλύτερη από τη μεγαλύτερη ρίζα του p_i , έτσι ώστε έχουμε εντοπίσει όλες τις $i+1$ ρίζες του p_{i+1} . Υπάρχει ακριβώς μία ρίζα του p_i ανάμεσα από κάθε ζεύγος συνεχόμενων ριζών του p_{i+1} και η ιδιότητα έπεται. •

Έχουμε δείξει, ιδιαίτερα, ότι όλες οι ρίζες από κάθε p_r είναι διακεκριμένες. Επιπλέον, $p_i(\lambda)$ και $p_{i-1}(\lambda)$ δεν μπορούν να μηδενίζονται και τα δύο για το ίδιο λ γιατί αν αυτό συνέβαινε η αναδρομική σχέση θα έδειχνε ότι αυτή η τιμή του λ θα ήταν ρίζα του p_r , για όλες τις τιμές του $r \in \{0, 1, \dots, n\}$. Όμως, p_0 προφανώς ποτέ δεν μηδενίζεται (αφού παραπάνω όρισα ότι $p_0(\lambda) \equiv 1$).

ΘΕΩΡΗΜΑ 6.2: (Ακολουθιακή ιδιότητα Sturm) Έστω ότι $\theta \in \mathbb{R}$ και θεωρούμε την ακολουθία $p_i(\theta)$, $i=0, 1, \dots, n$ (όπως ορίστηκε στη σχέση (6.2)) όπου υποθέτουμε ότι κανένα από τα $p_i(\theta)$ δεν είναι μηδέν. Ο αριθμός των συμφωνιών στο πρόσημο ανάμεσα σε διαδοχικά στοιχεία της ακολουθίας είναι ίσο με τον αριθμό των ιδιοτιμών του πίνακα T (σχέση (6.1)), όπου οι ιδιοτιμές αυτές είναι αυστηρά μεγαλύτερες από το θ .

Απόδειξη: Δίνεται $\lambda \in \mathbb{R}$ και $1 \leq j \leq n$, και γράφουμε $s_j(\lambda)$ για τον αριθμό των συμφωνιών στο πρόσημο ανάμεσα σε διαδοχικά στοιχεία για την ακολουθία

$$p_0(\lambda), p_1(\lambda), \dots, p_j(\lambda),$$

και $g_j(\lambda)$ για τον αριθμό των ριζών του πολυωνύμου p_j , οι οποίες είναι αυστηρά μεγαλύτερες από το λ .

Είναι προφανές ότι $s_1(\theta) = g_1(\theta)$. Η απόδειξη τώρα προχωράει με επαγωγή. Ας υποθέσουμε ότι $2 \leq k \leq n$ και ακολουθούμε την επαγωγική υπόθεση $s_{k-1}(\theta) = g_{k-1}(\theta)$.

Πρόκειται να αποδείξουμε ότι $s_k(\theta) = g_k(\theta)$.

Σύμφωνα με την υπόθεση, είτε $s_k(\theta) = s_{k-1}(\theta) + 1$ αν $p_k(\theta)$ και $p_{k-1}(\theta)$ έχουν το ίδιο πρόσημο, είτε $s_k(\theta) = s_{k-1}(\theta)$ αν έχουν αντίθετο πρόσημο. Υποθέτουμε ότι θ βρίσκεται στο διάστημα ανάμεσα σε δύο ρίζες α και β του p_{k-1} . Σύμφωνα με το Θεώρημα 6.1, υπάρχει ακριβώς μία ρίζα του p_k ανάμεσα στα α και β , και συμβολίζουμε τη ρίζα αυτή με φ . Όπως είδαμε στην απόδειξη του προηγούμενου θεωρήματος το $p_k(\lambda)$ είναι θετικό όταν το λ είναι μεγάλο σε απόλυτη τιμή και αρνητικό και το πρόσημο του $p_k(\lambda)$ προσδιορίζεται από τον αριθμό των ριζών του

p_k , οι οποίες είναι μικρότερες του λ . Τώρα, αν $\theta < \varphi$ τότε p_k και p_{k-1} έχουν τον ίδιο αριθμό ριζών μικρότερων του θ , έτσι ώστε $p_k(\theta)$ και $p_{k-1}(\theta)$ έχουν το ίδιο πρόσημο και ισχύει $s_k(\theta) = s_{k-1}(\theta) + 1$. Επίσης, αν p_k και p_{k-1} έχουν τον ίδιο αριθμό ριζών μικρότερων του θ , τότε p_k πρέπει να έχει μία ακόμα ρίζα η οποία είναι μεγαλύτερη του θ . Αυτό σημαίνει ότι $g_k(\theta) = g_{k-1}(\theta) + 1$. Άρα έχουμε ότι $s_k(\theta) = g_k(\theta)$. Ένα παρόμοιο επιχείρημα δείχνει ότι $s_k(\theta) = g_k(\theta)$ στην εναλλακτική περίπτωση όπου $\theta > \varphi$. Είναι επίσης απλή υπόθεση να μετατρέψουμε το επιχείρημα ελαφρώς για τις περιπτώσεις όπου το θ είναι μικρότερο από τη μικρότερη ρίζα του p_{k-1} ή μεγαλύτερο από την μεγαλύτερη ρίζα του p_{k-1} και έτσι το επαγωγικό βήμα είναι πλήρες. •

Το θεώρημα και η απόδειξη δεν επιτρέπουν στα στοιχεία της ακολουθίας να είναι μηδέν, στην οποία περίπτωση το πρόσημο γίνεται ασαφές. Μία περισσότερο προσεκτική ανάλυση είναι βαρετή αλλά όχι δύσκολη, μας δείχνει ότι το θεώρημα ακόμα κρατάει αν υιοθετήσουμε τη σύμβαση ότι όταν το $p_j(\theta)$ είναι μηδέν του δίνεται το ίδιο πρόσημο με αυτού του $p_{j-1}(\theta)$. Όπως έχουμε δει, δύο συνεχόμενα στοιχεία της ακολουθίας δεν μπορεί να είναι αμφότερα ίσα με το μηδέν. Το επόμενο παράδειγμα θα επεξηγήσει την εφαρμογή της ακολουθιακής ιδιότητας του Sturm.

ΠΑΡΑΔΕΙΓΜΑ 6.3: Προσδιορίστε τη δεύτερη μεγαλύτερη ιδιοτιμή του πίνακα

$$A = \begin{pmatrix} 3 & 1 & 0 & 0 \\ 1 & -1 & 2 & 0 \\ 0 & 2 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{pmatrix} \quad (6.3)$$

Αν οι ιδιοτιμές είναι οι λ_j , όπου $\lambda_1 > \lambda_2 > \lambda_3 > \lambda_4$, εμείς επιθυμούμε να βρούμε τη λ_2 . Τώρα, είναι εύκολο να δούμε από το Θεώρημα 4.3 (το 2^ο Θεώρημα Gerschgorin) ότι όλες οι ιδιοτιμές βρίσκονται στο διάστημα $[-4, 4]$. Παίρνουμε το κεντρικό σημείο του διαστήματος και υπολογίζουμε την ακολουθία Sturm με $\theta = 0$, η οποία δίνει

$$p_0(0) = 1, \quad p_1(0) = 3, \quad p_2(0) = -4, \quad p_3(0) = -16, \quad p_4(0) = -12$$

Σε αυτήν την ακολουθία υπάρχουν τρεις συμφωνίες στα πρόσημα:

$$(1,3) , (-4,-16) , (-16,-12)$$

Έτσι $s_4(0)=3$ και ο πίνακας έχει τρεις ιδιοτιμές μεγαλύτερες από το μηδέν. Αυτό σημαίνει ότι λ_2 πρέπει να βρίσκεται στο δεξιό μισό του διαστήματος $[-4,4]$, το οποίο είναι το $[0,4]$. Στη συνέχεια κατασκευάζουμε την ακολουθία Sturm με $\theta=2$ που είναι το μεσαίο στοιχείο του διαστήματος $[0,4]$, η οποία δίνει

$$p_0(2)=1 , p_1(2)=1 , p_2(2)=-4 , p_3(2)=-0 , p_4(2)=4$$

Σημειώνουμε ,αρχικά , ότι εδώ το $p_3(2)$ είναι μηδέν και του δίνεται το αρνητικό πρόσημο που συμφωνεί με το $p_2(2)$. Ο αριθμός των συμφωνιών στο πρόσημο εδώ είναι δύο , άρα δύο από τις ιδιοτιμές είναι μεγαλύτερες του δύο και λ_2 πρέπει να βρίσκεται στο διάστημα $[2,4]$ που είναι το δεξιό μισό του $[0,4]$.

Στη συνέχεια για $\theta=3$ αποκτούμε την ακολουθία

$$1 , +0 , -1 , 2 , -3 ,$$

με μία μόνο συμφωνία στο πρόσημο , έτσι το λ_2 πρέπει να βρίσκεται στο αριστερό μισό $[2,3]$ του διαστήματος $[2,4]$ και επαναλαμβάνουμε τη διαδικασία παίρνοντας $\theta = \frac{5}{2}$, το μεσαίο σημείο του διαστήματος $[2,3]$. Έτσι η ακολουθία είναι

$$1 , \frac{1}{2} , \frac{-11}{4} , \frac{17}{8} , \frac{-7}{16} ,$$

με μία μόνο συμφωνία στο πρόσημο , δείχνοντας ότι $\lambda_2 < 2.5$.

Αυτή η διαδικασία της διχοτόμησης μπορεί να επαναληφθεί όσες φορές απαιτείται για να εντοπίσουμε την ιδιοτιμή με μια δοθέντα ακρίβεια. Μετά από 13 στάδια βρίσκουμε ότι $\lambda_2 = 2.450$, σωστό για ακρίβεια τριών δεκαδικών ψηφίων. •

Αυτή η μέθοδος είναι αρκετά όμοια με τη συνηθισμένη διαδικασία διχοτόμησης για να βρούμε τη λύση της $f(x)=0$, ξεκινώντας με ένα διάστημα $[a,\beta]$ τέτοιο ώστε τα $f(a)$ και $f(\beta)$ να έχουν αντίθετα πρόσημα. Ένα μεγάλο πλεονέκτημα της ακολουθιακής μεθόδου Sturm δεν είναι μόνο ότι προσδιορίζει την ιδιοτιμή , αλλά μας δείχνει και ποια από τις ιδιοτιμές είναι. Αν χρησιμοποιούσαμε την μέθοδο

Jacobi του κεφαλαίου 3 θα έπρεπε να υπολογίσουμε όλες τις ιδιοτιμές, να τις ταξινομήσουμε σε σωστή σειρά και μετά να διαλέξουμε σαν λ_2 εκείνη με τη δεύτερη μεγαλύτερη τιμή.

Η ακολουθιακή μέθοδος *Sturm* θα υπολογίζει επίσης πόσες από τις ιδιοτιμές του πίνακα υπάρχουν σε ένα δοθέν διάστημα (α, β) . Αυτό που χρειάζεται μόνο να κάνουμε είναι να κατασκευάσουμε τις *Sturm* ακολουθίες $(p_j(a))_{j=0,1,\dots,n}$ και $(p_j(\beta))_{j=0,1,\dots,n}$. Στη συνέχεια, ο απαιτούμενος αριθμός ιδιοτιμών είναι $s_n(a) - s_n(\beta)$.

Είναι πολύ σημαντικό να υπολογίσουμε την ακολουθία $p_j(\theta)$ απευθείας από την επαγωγική σχέση. Για παράδειγμα, στο παράδειγμα 6.3, με $\theta = 2.445$ παίρνουμε:

$$\begin{aligned} p_0(2.445) &= 1, \\ p_1(2.445) &= 3 - 2.445 = 0.555, \\ p_2(2.445) &= (-1 - 2.445) \times 0.555 - 1 \times 1 = -2.9120, \\ p_3(2.445) &= (1 - 2.445) \times -2.9120 - 4 \times 0.555 = 1.9878, \\ p_4(2.445) &= (1 - 2.445) \times 1.9878 - 1 \times -2.9120 = 0.0396. \end{aligned}$$

Εναλλακτικά, για να κατασκευάσουμε άμεσες μορφές για τα πολυώνυμα $p_j(\lambda)$, $j = 0, 1, \dots, n$ και μετά να υπολογίσουμε $p_j(\theta)$ εισάγοντας $\lambda = \theta$ σε κάθε ένα από τα πολυώνυμα $p_j(\lambda)$, θα οδηγηθούμε στην κατασκευή της άμεσης μορφής του χαρακτηριστικού πολυωνύμου του πίνακα, που είναι το $p_n(\lambda)$ και θα έχουμε δει ότι αυτό επηρεάζεται ολέθρια από τα στρογγυλοποιημένα σφάλματα. Ο υπολογισμός από την απευθείας χρήση της επαγωγικής σχέσης δεν επηρεάζεται από τα στρογγυλοποιημένα σφάλματα.

ΠΑΡΑΔΕΙΓΜΑ 6.4: Σαν δεύτερο παράδειγμα, θεωρούμε πάλι τον πίνακα A του παραδείγματος 3.8, ο οποίος έχει μετασχηματιστεί σε τριδιαγώνια μορφή (παράδειγμα 5.6 – τελευταίος πίνακας)

Πίνακας 6.4: Διαδικασία διχοτόμησης για την μεγαλύτερη ιδιοτιμή. Η στήλη k δείχνει τον αριθμό των επαναλήψεων, θ_k είναι η k -οστή επανάληψη που προσεγγίζει την άγνωστη ιδιοτιμή λ_1 και $s_4(\theta_k)$ είναι ο αριθμός των συμφωνιών σε πρόσημο της *Sturm* ακολουθίας $p_0(\theta_k), \dots, p_4(\theta_k)$.

k	θ_k	$s_4(\theta_k)$
1	0.00000	3
2	5.46285	2
3	8.19428	0
4	6.82857	2
5	7.51142	1
6	7.85285	1
7	8.02357	1
8	8.10892	0
9	8.06625	1
10	8.08758	1
11	8.09825	0
12	8.09292	1
13	8.09559	0
14	8.09425	0
15	8.09359	1
16	8.09392	0
17	8.09375	0
18	8.09367	1

Ο πίνακας 6.4 δείχνει το αποτέλεσμα της διαδικασίας διχοτόμησης , χρησιμοποιώντας την Sturm ακολουθία. Η ∞ -νόρμα του τριδιαγώνιου πίνακα είναι 10.926 , έτσι η διαδικασία ξεκινά με το διάστημα $[-10.926,10.926]$.Η μεγαλύτερη ιδιοτιμή είναι η 8.094 , με τρία δεκαδικά ψηφία , η οποία συμφωνεί με το αποτέλεσμα της μεθόδου Jacobi στο κεφάλαιο 3. Αυτός ο πίνακας δείχνει πόσο λίγες αποθηκεύσεις είναι δυνατές ακόμα και όταν απαιτούνται όλες οι ιδιοτιμές. Παρατηρούμε από τον πίνακα ότι για $\theta = 7.51142$ δίνει μία συμφωνία στο πρόσημο ενώ για $\theta = 6.82857$ δίνει δύο συμφωνίες στο πρόσημο. Στρογγυλεύοντας σε τρία δεκαδικά ψηφία πάλι , η διαδικασία της διχοτόμησης για τη δεύτερη μεγαλύτερη ιδιοτιμή μπορεί να ξεκινήσει με το διάστημα $[6.829, 7.511]$. •

Η μέθοδος της διχοτόμησης μπορεί να παρουσιάζεται σχετικά αργή αλλά έχει το σημαντικό πλεονέκτημα της εγγυημένης επιτυχίας και επηρεάζεται πολύ λίγο από τα στρογγυλοποιημένα σφάλματα. Επιπλέον , η ποσότητα της δουλειάς που προϋποθέτεται δεν είναι μεγάλη. Αν έχουμε υπολογίσει τα τετράγωνα των μη διαγώνιων στοιχείων , b_r^2 , του πίνακα T εκ των προτέρων, κάθε υπολογισμός όλων των μελών της ακολουθίας απαιτεί $2n$ πολλαπλασιασμούς. Αν η διαδικασία της

διχοτόμησης συνεχιστεί για 40 στάδια, η ιδιοτιμή θα έχει υπολογιστεί με εννέα σημαντικά ψηφία και αν απαιτούμε να υπολογίσουμε m από τις ιδιοτιμές με αυτήν την ακρίβεια πρόκειται να χρειαστούμε $80mn$ πολλαπλασιασμούς. Αν m είναι μια καλή ποσότητα μικρότερη από το n , όπου n η τάξη του πίνακα, πρόκειται να είναι μία εξαιρετική ποσότητα μικρότερη από τη δουλειά που προϋποθέτεται για την διαδικασία αναγωγής σε τριδιαγώνια μορφή, η οποία όπως έχουμε δει χρειάζεται $\frac{1}{3}n^3$ πολλαπλασιασμούς. Στα περισσότερα πρακτικά προβλήματα η αρχική Householder αναγωγή σε τριδιαγώνια μορφή είναι η αιτία για την περισσότερη υπολογιστική δουλειά.

7. Ο ΑΛΓΟΡΙΘΜΟΣ QR

Σε αυτό το κεφάλαιο θα συζητήσουμε με συντομία τον αλγόριθμο QR, μία εναλλακτική μέθοδο για τον υπολογισμό των ιδιοτιμών ενός τριδιαγώνιου πίνακα. Κατ' αρχήν η μέθοδος αυτή μπορεί να εφαρμοστεί σε έναν πλήρη πίνακα, άλλα είναι πιο αποδοτικό να χρησιμοποιήσουμε την μέθοδο Householder για να μετασχηματίσουμε τον πίνακα σε τριδιαγώνια μορφή πρώτα. Η βάση της μεθόδου αυτής είναι η QR παραγοντοποίηση του πίνακα.

7.1 Η ΠΑΡΑΓΟΝΤΟΠΟΙΗΣΗ QR

Υποθέτουμε ότι $n \geq 3$ και $A \in \mathbb{R}^{n \times n}$ είναι ένας συμμετρικός τριδιαγώνιος πίνακας. Αρχικά, θα δείξουμε πως κατασκευάζουμε έναν ορθογώνιο πίνακα $Q \in \mathbb{R}^{n \times n}$ και έναν άνω τριγωνικό πίνακα $R \in \mathbb{R}^{n \times n}$ τέτοιους ώστε $A = QR$. Το πρόβλημα είναι όμοιο με την παραγοντοποίηση LU που χρησιμοποιείται στην επίλυση συστημάτων γραμμικών εξισώσεων, αλλά εδώ έχουμε ένα ορθογώνιο πίνακα Q αντί για ένα κάτω τριγωνικό πίνακα L.

Κατασκευάζουμε τον πίνακα Q σαν ένα γινόμενο από πίνακες επίπεδης περιστροφής $R^{(p,p+1)}(\varphi) \in \mathbb{R}^{n \times n}$ (βλέπε ορισμό 3.2), με μια κατάλληλη επιλεγμένη γωνία φ . Για να εξηγήσουμε τι σημαίνει εδώ 'σωστά επιλεγμένη', σημειώνουμε ότι στο γινόμενο

$$B = R^{(p,p+1)}(\varphi)A \quad (7.1)$$

το στοιχείο $b_{p+1,p}$ είναι εύκολο να βρεθεί ότι είναι

$$b_{p+1,p} = -sa_{pp} + ca_{p+1,p},$$

όπου $s = \sin \varphi$ και $c = \cos \varphi$. Μπορούμε να έχουμε $b_{p+1,p} = 0$ επιλέγοντας

$$s = \frac{a_{p+1,p}}{\rho}, \quad c = \frac{a_{pp}}{\rho}, \quad \rho = (a_{pp}^2 + a_{p+1,p}^2)^{1/2}. \quad (7.2)$$

Σημειώνουμε παρεμπιπτόντως ότι

$$b_{pp} = ca_{pp} + sa_{p+1,p},$$

$$\begin{aligned} b_{p,p+1} &= ca_{p,p+1} + sa_{p+1,p+1} \ , \\ b_{p+1,p+1} &= -sa_{p,p+1} + ca_{p+1,p+1} \ . \end{aligned}$$

Τα υπόλοιπα στοιχεία του πίνακα B είναι τα ίδια με αυτά του πίνακα A . Για να συνοψίζουμε τα πιο σημαντικά σημεία, μετά τον πολλαπλασιασμό του συμμετρικού τριδιαγώνιου πίνακα A από αριστερά με τον $R^{(pp+1)}(\varphi)$, όπου $c = \cos \varphi$ και $s = \sin \varphi$ στον $R^{(pp+1)}(\varphi)$ επιλέγονται σύμφωνα με τον ορισμό (7.2), αποκτούμε ένα τριδιαγώνιο πίνακα $B = (b_{ij}) \in \mathbb{R}^{n \times n}$ τέτοιο ώστε $b_{p+1,p} = 0$.

Μετά από αυτή τη σύντομη προετοιμασία, ξεκινάμε την περιγραφή της παραγοντοποίησης QR . Υποθέτουμε ότι πολλαπλασιάζουμε διαδοχικά τον πίνακα A από αριστερά με $n-1$ πίνακες επίπεδης περιστροφής,

$$Q_1 = R^{(12)}(\varphi_1) \ , \quad Q_2 = R^{(23)}(\varphi_2) \ , \quad \dots \ , \quad Q_{n-1} = R^{(n-1n)}(\varphi_{n-1}) \ ,$$

με $\varphi_1, \varphi_2, \dots, \varphi_{n-1}$ να διαλέγονται σύμφωνα με τη σχέση (7.2). Πιο συγκεκριμένα:

for $p = 1, 2, \dots, n-1$,
 φ_p διαλέγεται έτσι ώστε το $(p+1, p)$ -στοιχείο του πίνακα $Q_p \dots Q_1 A$ να είναι μηδέν.

Με δεδομένο ότι τα στοιχεία κάτω από τη διαγώνιο του πίνακα

$$Q_{p-1} \dots Q_1 A \ , \quad 2 \leq p \leq n-1 \ ,$$

τα οποία είναι ίσα με το μηδέν, παραμένουν μηδέν μετά τον πολλαπλασιασμό με τον επόμενο πίνακα επίπεδης περιστροφής Q_p της ακολουθίας, συμπεραίνουμε ότι μετά από διαδοχικούς πολλαπλασιασμούς του πίνακα A από αριστερά με τους Q_1, Q_2, \dots, Q_{n-1} , ο πίνακας

$$Q_{n-1} Q_{n-2} \dots Q_1 A = R \ , \tag{7.3}$$

είναι άνω τριγωνικός. Πράγματι, επειδή ο πίνακας A είναι τριδιαγώνιος, ο πίνακας R είναι τριδιαγώνιος και άνω τριγωνικός. Συνεπώς, ο πίνακας R είναι **διδιαγώνιος** υπό την έννοια ότι $R_{ij} = 0$ αν $i \neq j, j-1$. Με την ιδιότητα ότι οι πίνακες $Q_p = R^{(pp+1)}(\varphi_p)$, $p = 1, 2, \dots, n-1$ είναι ορθογώνιοι και συνεπώς

$Q_p^T Q_p = I$, πολλαπλασιάζοντας στην σχέση (7.3) από αριστερά με $Q_1^T Q_2^T \dots Q_{n-1}^T$, βρίσκουμε ότι

$$A = QR ,$$

όπου

$$Q = Q_1^T Q_2^T \dots Q_{n-1}^T$$

είναι ένας ορθογώνιος πίνακας (σαν γινόμενο από ορθογώνιους πίνακες). Η επόμενη ενότητα (7.2) του κεφαλαίου αυτού περιγράφει τον αλγόριθμο QR , ο οποίος βασίζεται την παραγοντοποίηση QR , για την αριθμητική λύση του προβλήματος ιδιοτιμών ενός πίνακα $A \in \mathbb{R}^{n \times n}$, ο οποίος είναι συμμετρικός και τριδιαγώνιος.

7.2 Ο ΟΡΙΣΜΟΣ ΤΟΥ ΑΛΓΟΡΙΘΜΟΥ ΤΟΥ QR

Υποθέτουμε ότι έχουμε ένα πίνακα $A \in \mathbb{R}^{n \times n}$ συμμετρικό και τριδιαγώνιο . Ο αλγόριθμος QR ορίζει μία ακολουθία από συμμετρικούς τριδιαγώνιους πίνακες $A^{(k)} \in \mathbb{R}^{n \times n}$, $k = 0, 1, 2, \dots$ ξεκινώντας με $A^{(0)} = A$.

Υποθέτουμε ότι $k \geq 0$. Το k -οστό βήμα του αλγορίθμου QR παίρνει τον συμμετρικό τριδιαγώνιο πίνακα $A^{(k)}$ και διαλέγει μία μετατόπιση $\mu_k \in \mathbb{R}$ (η επιλογή του μ_k θα συζητηθεί παρακάτω) και στη συνέχεια σχηματίζουμε την παραγοντοποίηση QR

$$A^{(k)} - \mu_k I = Q^{(k)} R^{(k)} .$$

Στη συνέχεια πολλαπλασιάζουμε τα $Q^{(k)}$ και $R^{(k)}$ με αντίστροφη σειρά (από την παραπάνω σχέση) και κατασκευάζουμε τον νέο πίνακα $A^{(k+1)}$, ο οποίος ορίζεται από την παρακάτω σχέση

$$A^{(k+1)} = R^{(k)} Q^{(k)} + \mu_k I .$$

Θυμίζοντας ότι ο πίνακας $Q^{(k)}$ είναι ορθογώνιος , είναι απλό ζήτημα να δούμε ότι $A^{(k+1)} = Q^{(k)T} A^{(k)} Q^{(k)}$, έτσι ώστε οι πίνακες $A^{(k+1)}$ και $A^{(k)}$ να έχουν τις ίδιες ιδιοτιμές. Επειδή $A^{(0)} = A$, όλοι οι πίνακες της ακολουθίας $(A^{(k)})$ έχουν τις ίδιες ιδιοτιμές με τον πίνακα A . Είναι , επίσης , εύκολο να δειχθεί ότι κάθε ένας από τους πίνακες $A^{(k)}$ είναι συμμετρικός και τριδιαγώνιος.

Η επιλογή της παραμέτρου μετατόπισης μ_k είναι πολύ σημαντική. Αν είναι ορθά επιλεγμένη η ακολουθία πινάκων $A^{(k)}$ τότε θα συγκλίνει γρήγορα σε ένα πίνακα, μέσα στον οποίο έναν από τα μη διαγώνια στοιχεία είναι μηδέν. Αν το στοιχείο αυτό βρίσκεται στην πρώτη ή στην τελευταία γραμμή, τότε θα έχουμε προσδιορίσει μία από τις ιδιοτιμές. Αν το στοιχείο αυτό είναι ένα από τα ενδιάμεσα στοιχεία, τότε μπορούμε να διασπάσουμε τον πίνακα σε δύο ξεχωριστούς πίνακες μικρότερης τάξης. Σε αυτήν την περίπτωση μπορούμε να επαναλάβουμε την επαναληπτική διαδικασία με μικρότερους πίνακες μέχρι να βρεθούν όλες οι ιδιοτιμές.

Η συνήθης απλή επιλογή για την παράμετρο μετατόπισης στο k -οστό βήμα είναι

$$\mu_k = a_{mm}^{(k)},$$

το τελευταίο διαγώνιο στοιχείο του πίνακα $A^{(k)}$. Γενικά, μετά από αρκετά βήματα της επανάληψης το στοιχείο στη θέση $(n, n-1)$ θα γίνει ασήμαντα μικρό. Μία από τις ιδιοτιμές του προκύπτον πίνακα είναι στη συνέχεια το τελευταίο διαγώνιο στοιχείο και μετά συνεχίζουμε την διαδικασία με τον πίνακα τάξης $n-1$ που προκύπτει αφαιρώντας την τελευταία γραμμή και στήλη. Πρέπει να σημειώσουμε επίσης ότι υπάρχουν ειδικές συνθήκες όπου η επιλογή της μετατόπισης είναι ανεπαρκής και άλλες καταστάσεις όπου κάποια άλλη επιλογή είναι πιο αποδοτική, αλλά δεν πρόκειται να συζητήσουμε λεπτομέρειες για αυτή την επιλογή περαιτέρω. Η απόδειξη για την σύγκλιση αυτής της μεθόδου είναι αρκετά μεγάλη και με πολλά τεχνικά στοιχεία και έτσι δεν την παραθέτουμε εδώ. (Για περαιτέρω μελέτη η απόδειξη αυτή υπάρχει πίσω-βιβλιογραφία [5]-κεφάλαιο 8).

Αυτή η μέθοδος δεν υπολογίζει τις ιδιοτιμές με κάποια συγκεκριμένη σειρά, και έτσι αν απαιτούμε μόνο ένα μικρό αριθμό από τις μεγαλύτερες ιδιοτιμές για παράδειγμα τότε η ακολουθιακή μέθοδος Sturm είναι προτιμότερη. Η πιο συνηθισμένη υπόδειξη είναι ότι ο αλγόριθμος QR χρησιμοποιείται σε πίνακες τάξης n αν απαιτείται περισσότερο από το $\frac{1}{4}n$ των ιδιοτιμών.

Αμέσως μετά βρίσκουμε την περιγραφή του αλγόριθμου QR

ΑΛΓΟΡΙΘΜΟΣ QR:

Δίνεται ένας πίνακας $A \in \mathbb{R}^{n \times n}$ συμμετρικός και τριδιαγώνιος.

Θέτουμε $A^{(0)} = A$, $\sigma_0 = (A^{(0)})_{nn}$ (η μετατόπιση μ_k) και για $k = 1, 2, \dots$

Βήμα 1^ο: Βρίσκουμε τον $A^{(k)} - \sigma_k I = Q^{(k)} R^{(k)}$

Βήμα 2^ο: Θέτουμε $A^{(k+1)} = R^{(k)} Q^{(k)} + \sigma_k I$

$$\text{και } \sigma_{k+1} = (A^{(k+1)})_{nn}$$

και συνεχίζουμε τη διαδικασία παίρνοντας την ακολουθία πινάκων $A^{(k)}$.

Παρακάτω δίνουμε ένα παράδειγμα.

ΠΑΡΑΔΕΙΓΜΑ 7.1: Εφαρμόζουμε τον αλγόριθμο QR στον τριδιαγώνιο πίνακα του παραδείγματος 5.6 (τελευταίος πίνακας). Ο πίνακας είναι:

$$\begin{pmatrix} 4.000 & -3.162 & 0.000 & 0.000 & 0.000 \\ -3.162 & 5.300 & -1.308 & 0.000 & 0.000 \\ 0.000 & -1.308 & 0.057 & 2.306 & 0.000 \\ 0.000 & 0.000 & 2.306 & 5.208 & -3.411 \\ 0.000 & 0.000 & 0.000 & -3.411 & -1.565 \end{pmatrix}$$

Μετά από ένα βήμα της επανάληψης ο πίνακας $A^{(1)} = R^{(0)}Q^{(0)} + \mu_0 I$, με $\mu_0 = a_{55}^{(0)} = a_{55}$ είναι ο παρακάτω:

$$A^{(1)} = \begin{pmatrix} 7.034 & -2.271 & 0 & 0 & 0 \\ -2.271 & 2.707 & -0.744 & 0 & 0 \\ 0 & -0.744 & 5.804 & 3.202 & 0 \\ 0 & 0 & 3.202 & -0.464 & 1.419 \\ 0 & 0 & 0 & 1.419 & -2.082 \end{pmatrix}$$

Στις αλληπάλληλες επαναλήψεις $k=1,2,3,4,5$ το στοιχείο $a_{54}^{(k)}$ έχει τις τιμές 1.419, -1.262, 0.965, -0.223, 0.002 και μετά την επόμενη επανάληψη το στοιχείο $a_{54}^{(k)}$ μηδενίζεται προς 10 δεκαδικά ψηφία. Το στοιχείο $a_{55}^{(6)}$ είναι -3.282, το οποίο είναι συνεπώς μία ιδιοτιμή.

Στη συνέχεια διαγράφουμε την τελευταία γραμμή και στήλη και συνεχίζουμε την διαδικασία στον 4×4 πίνακα που προκύπτει. Ύστερα από μία μόνο επανάληψη το στοιχείο στη θέση (4,3) μηδενίζεται προς 7 δεκαδικά ψηφία, δίνοντας την ιδιοτιμή -0.671. Μετά διαγράφουμε την τελευταία γραμμή και στήλη και συνεχίζουμε με τον 3×3 πίνακα που προκύπτει. Μετά από μία επανάληψη του 3×3 πίνακα το στοιχείο στη θέση (3,2) είναι το 0.0005 και ακόμη μία επανάληψη δίνει την ακριβή ιδιοτιμή 1.690. Έτσι τώρα μένουμε με έναν 2×2 πίνακα και ο υπολογισμός των δύο τελευταίων ιδιοτιμών είναι τετριμμένος.

Ο αριθμός των επαναλήψεων που απαιτούνται για να απομονώσουμε κάθε ιδιοτιμή μειώνεται όταν ο αλγόριθμος μειώνει το μέγεθος του πίνακα. Αυτού του είδους η συμπεριφορά είναι αναμενόμενη.

Οι αριθμητικές τιμές συμφωνούν με αυτές που λάβαμε και από την μέθοδο Jacobi αλλά και από την μέθοδο της διχοτόμησης. •

8. ΑΝΤΙΣΤΡΟΦΗ ΕΠΑΝΑΛΗΨΗ ΓΙΑ ΤΑ ΙΔΙΟΔΙΑΝΥΣΜΑΤΑ

8.1 Η ΜΕΘΟΔΟΣ ΤΩΝ ΔΥΝΑΜΕΩΝ (POWER METHOD)

Πριν αναλύσουμε την μέθοδο της αντίστροφης επανάληψης θα κάνουμε μία εισαγωγή για μια παρεμφερή μέθοδο, την μέθοδο των δυνάμεων. Η μέθοδος αυτή είναι καλή στο να προσεγγίζει τις οριακές ιδιοτιμές του πίνακα, δηλαδή τις ιδιοτιμές με το μεγαλύτερο και το μικρότερο μέτρο, λ_n και λ_1 αντίστοιχα καθώς και τα σχετιζόμενα με αυτές ιδιοδιανύσματα. Λύνοντας ένα τέτοιο πρόβλημα παρατηρούμε ότι παρουσιάζεται μεγάλο ενδιαφέρον από πολλές πραγματικές εφαρμογές (όπως γεοσεισμολογία, μηχανικές ταλαντώσεις, ανάλυση ηλεκτρικών κυκλωμάτων, κβαντομηχανική κ.α.) όπου ο υπολογισμός της λ_n (και του αντίστοιχου ιδιοδιανύσματος x_n) προκύπτει από τον καθορισμό της σωστής συχνότητας (και της αντίστοιχης θεμελιώδης ταλάντωσης) ενός δοθέντος φυσικού συστήματος.

Έχοντας προσεγγίσεις από τις λ_n και λ_1 μπορεί επίσης να είναι χρήσιμο στην ανάλυση των αριθμητικών μεθόδων. Για παράδειγμα, αν A είναι συμμετρικός πίνακας και θετικά ορισμένος μπορεί να υπολογίσει τη βέλτιστη τιμή της επιτάχυνσης και να υπολογίσει το σφάλμα της καθώς και να εκτελεί την ανάλυση της ευστάθειας σε μεθόδους για τα συστήματα με συνήθεις διαφορικές εξισώσεις.

Ας δούμε τώρα την μέθοδο των δυνάμεων:

Έστω $A \in \mathbb{C}^{n \times n}$ είναι ένας διαγωνοποιήσιμος πίνακας και έστω $X \in \mathbb{C}^{n \times n}$ είναι ο πίνακας με τα σωστά ιδιοδιανύσματα x_i , $i=1,2,\dots,n$. Υποθέτουμε επίσης ότι οι ιδιοτιμές του πίνακα A διατάσσονται ως εξής

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \dots \geq |\lambda_n| \quad (8.1)$$

όπου λ_1 έχει αλγεβρική πολλαπλότητα ίση με 1. Κάτω από αυτές τις υποθέσεις η λ_1 καλείται κυρίαρχη ιδιοτιμή του πίνακα A .

Δοθέντος ενός αυθαίρετου αρχικού διανύσματος $u^{(0)} \in \mathbb{C}^n$ με $\|u^{(0)}\|_2 = 1$ και θεωρώντας για $k=1,2,\dots$ την ακόλουθη επανάληψη που βασίζεται στον υπολογισμό των δυνάμεων των πινάκων, έχουμε την μέθοδο των δυνάμεων που περιγράφεται από τους παρακάτω τύπους

$$\begin{aligned} z^{(k)} &= Au^{(k-1)} \\ u^{(k)} &= z^{(k)} / \|z^{(k)}\|_2 \\ v^{(k)} &= \left(u^{(k)}\right)^H Au^{(k)}. \end{aligned} \tag{8.2}$$

8.2 ΑΝΤΙΣΤΡΟΦΗ ΕΠΑΝΑΛΗΨΗ (INVERSE ITERATION)

Είδαμε στο κεφάλαιο 3 ότι με την μέθοδο Jacobi μπορούμε επίσης, αν απαιτείται, να παράγουμε και τα ιδιοδιανύσματα του πίνακα εκτός από τις ιδιοτιμές. Αλλά, η χρήση της μεθόδου Householder σε συνδυασμό με την ακολουθιακή ιδιότητα του Sturm ή τον αλγόριθμο QR δίνει μόνο τις ιδιοτιμές. Υποθέτουμε ότι ο $A \in \mathbb{R}^{n \times n}$ είναι ένας συμμετρικός πίνακας και εικάζουμε ότι έχουμε μία καλή προσέγγιση $\theta \in \mathbb{R}$ της απαιτούμενης ιδιοτιμής $\lambda \in \mathbb{R}$ του πίνακα A και κάποια προσέγγιση $u^{(0)} \in \mathbb{R}_*^n$, $\|u^{(0)}\|_2 = 1$ του συσχετιζόμενου ιδιοδιανύσματος $u \in \mathbb{R}_*^n$, $\|u\|_2 = 1$. Είναι αυτονόητο ότι $\theta \neq \lambda$ και έτσι θ δεν είναι ιδιοτιμή του πίνακα A , έτσι ώστε ο πίνακας $A - \theta I$ είναι ομαλός ή μη ιδιάζων (ένας πίνακας A είναι ομαλός αν υπάρχει πίνακας B τέτοιος ώστε $AB = BA = I$). Η μέθοδος της **αντίστροφης επανάληψης (inverse iteration)** (ή αλλιώς αντίστροφη μέθοδος των δυνάμεων) ορίζει μία ακολουθία από διανύσματα $u^{(k)}$, $k = 0, 1, \dots$, όπως ακολουθείται: Δοθέντος ενός $u^{(k)} \in \mathbb{R}_*^n$, βρίσκουμε το $w^{(k)} \in \mathbb{R}_*^n$ και στη συνέχεια το $u^{(k+1)} \in \mathbb{R}_*^n$ από τις σχέσεις

$$\begin{aligned} (A - \theta I)w^{(k)} &= u^{(k)}, \\ u^{(k+1)} &= c_k w^{(k)} \end{aligned} \tag{8.3}$$

όπου $c_k = 1 / \sqrt{w^{(k)T} w^{(k)}} = 1 / \|w^{(k)}\|_2$. Έτσι, συμπεραίνουμε ότι $\|u^{(k)}\|_2 = 1$, $k = 0, 1, 2, \dots$.

ΘΕΩΡΗΜΑ 8.1: Υποθέτουμε ότι $A \in \mathbb{R}_{sym}^{n \times n}$. Η ακολουθία των διανυσμάτων $(u^{(k)})$ του \mathbb{R}_*^n που ορίζεται από τη διαδικασία αντίστροφης επανάληψης (σχέση (8.3)) συγκλίνει σε ένα κανονικοποιημένο ιδιοδιάνυσμα $u \in \mathbb{R}_*^n$ αντίστοιχο της ιδιοτιμής $\lambda \in \mathbb{R}$ η οποία είναι κοντινότερη στο $\theta \in \mathbb{R}$ (με την προϋπόθεση ότι το θ δεν είναι

ιδιοτιμή), έτσι το λ είναι μία απλή ιδιοτιμή και το αρχικό διάνυσμα $u^{(0)} \in \mathbb{R}_*^n$ δεν είναι ορθογώνιο στο διάνυσμα u .

Απόδειξη: Σύμφωνα με το Θεώρημα 1.4 (vii), το διάνυσμα $u^{(0)}$ μπορεί να εκφραστεί σαν γραμμικός συνδυασμός των (ορθο)κανονικοποιημένων ιδιοδιανυσμάτων $x^{(j)}$, $j = 1, 2, \dots, n$ του πίνακα A στη μορφή

$$u^{(0)} = \sum_{j=1}^n a_j x^{(j)} \quad , \quad a_j = u^{(0)T} x^{(j)} \quad . \quad (8.4)$$

Έστω $\lambda_s \in \mathbb{R}$ υποδηλώνει την ιδιοτιμή του πίνακα A που είναι πιο κοντά στο $\theta \in \mathbb{R}$. Πρόκειται να αποδείξουμε ότι η ακολουθία $(u^{(k)})$ συγκλίνει, όσο $k \rightarrow \infty$, στο ιδιοδιάνυσμα $u = x^{(s)} \in \mathbb{R}_*^n$ που σχετίζεται με τη λ_s με την προϋπόθεση ότι $a_s = u^{(0)T} x^{(s)} \neq 0$.

Επεκτείνοντας με τη βοήθεια των (ορθο)κανονικοποιημένων ιδιοδιανυσμάτων $x^{(j)}$ έχουμε

$$w^{(0)} = \sum_{j=1}^n \beta_j x^{(j)}$$

τοποθετούμε αυτή την επέκταση στην πρώτη γραμμή της σχέσης (8.3) με $k = 0$ και έτσι έχουμε ότι

$$u^{(0)} = (A - \theta I) \sum_{j=1}^n \beta_j x^{(j)}$$

και συγκρίνοντας την με τη σχέση (8.4) για το $u^{(0)}$ προκύπτει ότι $(\lambda_j - \theta)\beta_j = a_j$. Επειδή $\lambda_j - \theta \neq 0$ για όλα τα $j \in \{1, 2, \dots, n\}$, βλέπουμε ότι

$$u^{(1)} = c_0 w^{(0)} = c_0 \sum_{j=1}^n \frac{a_j}{\lambda_j - \theta} x^{(j)}.$$

Επαναλαμβάνοντας την παραπάνω σχέση για $k = 1, 2, \dots, m-1$ λαμβάνουμε τη σχέση

$$u^{(m)} = c_{m-1} \dots c_0 \sum_{j=1}^n \frac{a_j}{(\lambda_j - \theta)^m} x^{(j)} \quad . \quad (8.5)$$

Έτσι, $u^{(m)T} u^{(m)} = 1$ και συνεπώς

$$c_{m-1} \dots c_0 = \left[\sum_{j=1}^n \frac{a_j^2}{(\lambda_j - \theta)^{2m}} \right]^{-1/2}. \quad (8.6)$$

Αντικαθιστώντας την σχέση (8.6) στη σχέση (8.5), λαμβάνουμε τη σχέση

$$u^{(m)} = \frac{\sum_{j=1}^n \frac{a_j}{(\lambda_j - \theta)^m} x^{(j)}}{\left[\sum_{j=1}^n \frac{a_j^2}{(\lambda_j - \theta)^{2m}} \right]^{1/2}} = \frac{x_s + \sum_{j \neq s} \left(\frac{a_j}{a_s} \right) \left(\frac{\lambda_s - \theta}{\lambda_j - \theta} \right)^m x^{(j)}}{\left[1 + \sum_{j \neq s} \left(\frac{a_j}{a_s} \right)^2 \left(\frac{\lambda_s - \theta}{\lambda_j - \theta} \right)^{2m} \right]^{1/2}}.$$

Επειδή, $\left| \frac{\lambda_s - \theta}{\lambda_j - \theta} \right| < 1, \quad \forall j \in \{1, 2, \dots, n\} / \{s\}$

βρίσκουμε ότι $\lim_{m \rightarrow \infty} u^{(m)} = x_s = u$ και έτσι η απόδειξη έχει ολοκληρωθεί. •

Στην αναπάντεχη περίπτωση που το θ που έχουμε είναι μία ακριβής ιδιοτιμή είναι πιθανό καλύτερα στην πράξη να προσθέσουμε μία πολύ μικρή διαταραχή στο θ πριν ξεκινήσουμε την επανάληψη. Αν η εκτίμηση του θ είναι εντός του σφάλματος στρογγύλευσης του λ_s και οι ιδιοτιμές είναι καλά διατεταγμένες τότε η σύγκλιση της ακολουθίας $(u^{(k)})$ θα είναι υπερβολικά γρήγορη. Συνήθως ένας πολύ μικρός αριθμός επαναλήψεων θα είναι επαρκής.

Η απόδειξη του Θεωρήματος 8.1 δεν ισχύει στην περίπτωση που $a_s = 0$, για παράδειγμα όταν το αρχικό διάνυσμα $u^{(0)}$ είναι ακριβώς ορθογώνιο στο απαιτούμενο ιδιοδιάνυσμα. Όμως, αυτό δεν σημαίνει ότι η επανάληψη (σχέση (8.3)) επίσης κατεδαφίζεται. Για την επίδραση του σφάλματος στρογγυλοποίησης σχεδόν πάντα θα παρουσιάζουμε ένα μικρό πολλαπλάσιο του διανύσματος $x^{(s)}$ στην επέκταση του $u^{(0)}$ σε αντιστοιχία με το $x^{(j)}$ με $j = 1, 2, \dots, n$ και το απαιτούμενο ιδιοδιάνυσμα θα αποκτάται μέσα σε ένα μικρό αριθμό επαναλήψεων. Αυτό είναι ένα χρήσιμο στοιχείο της μεθόδου, επειδή στη πράξη δεν είναι εφικτό να εξετάσουμε αν το $u^{(0)}$ είναι ή δεν είναι ορθογώνιο στο u , δοθέντος ότι το ιδιοδιάνυσμα u είναι άγνωστο.

Υπάρχει, επίσης, πρόβλημα αν υπάρχει πολλαπλή ιδιοτιμή ή αν δύο ιδιοτιμές είναι κοντά μεταξύ τους: στην πρώτη περίπτωση $|\lambda_s - \theta| / |\lambda_j - \theta| = 1$ για κάποιο $j \neq s$ και έτσι η απόδειξη του Θεωρήματος 8.1 κατεδαφίζεται, ενώ στην δεύτερη περίπτωση $|\lambda_s - \theta| / |\lambda_j - \theta| \approx 1$ για κάποιο $j \neq s$ και έτσι οδηγούμαστε σε πολύ αργή σύγκλιση.

Ο υπολογισμός του $w^{(k)}$ από τη σχέση (8.3) απαιτεί τη λύση από ένα σύστημα γραμμικών εξισώσεων του οποίου πίνακας είναι ο $A - \theta I$. Αυτός ο πίνακας θα είναι συνήθως σχεδόν μη ομαλός, η επιδίωξη μας στην επιλογή του θ ήταν να κάνουμε τον πίνακα $A - \theta I$ ακριβώς μη ομαλό. Γενικά η λύση ενός τέτοιου συστήματος είναι πολύ επικίνδυνο λόγω της επίδρασης των στρογγυλοποιημένων σφαλμάτων. Σε αυτήν την περίπτωση όμως η επίδραση του στρογγυλοποιημένου σφάλματος θα παρουσιάσει ένα πολλαπλάσιο του κυρίαρχου ιδιοδιανύσματος και αυτό είναι ακριβώς αυτό που απαιτείται. Η ανάλυση της επίδρασης των στρογγυλοποιημένων σφαλμάτων θα επιβεβαιώσει το γεγονός, αλλά εδώ δεν θα κάνουμε περισσότερη ανάλυση. (Για περισσότερη μελέτη πάνω σε αυτό το θέμα βλέπε βιβλιογραφία το [6]-(κεφάλαιο 4.1 και 9) και το [7]-(καφάλαιο 7.6.1)).

Υπάρχουν δύο διαδρομές που μπορεί να εφαρμοστεί η διαδικασία της αντίστροφης επανάληψης. Μία προφανή δυνατότητα είναι να χρησιμοποιήσουμε τον αρχικό πίνακα $A \in \mathbb{R}^{n \times n}$, όπως υποδηλώθηκε στη σχέση (8.3). Μία εναλλακτική είναι να αντικαταστήσουμε τον πίνακα A στην εξίσωση αυτή με τον τριδιαγώνιο πίνακα $T \in \mathbb{R}^{n \times n}$ που προκύπτει από τη Householder μέθοδο. Ο υπολογισμός, στη συνέχεια, είναι πιο γρήγορος αλλά παράγει τα ιδιοδιανύσματα του τριδιαγώνιου πίνακα T . Για να προκύψουν τα αντίστοιχα ιδιοδιανύσματα του πίνακα A πρέπει να εφαρμόσω σε αυτό το διάνυσμα την ακολουθία των Householder μετασχηματισμών οι οποίοι χρησιμοποιήθηκαν στην αρχική αναγωγή σε τριδιαγώνια μορφή. Είναι εύκολο να δούμε ότι αυτή είναι η πιο αποτελεσματική μέθοδος.

Η αντίστροφη επανάληψη με τον αρχικό πίνακα $A \in \mathbb{R}^{n \times n}$ απαιτεί την παραγοντοποίηση LU του πίνακα A η οποία απαιτεί προσεγγιστικά $\frac{1}{3}n^3$ πολλαπλασιασμούς. Η ίδια διαδικασία με τον τριδιαγώνιο πίνακα T , χρησιμοποιώντας τον αλγόριθμο του Thomas, προϋποθέτει μόνο μικρό πολλαπλάσιο του n πολλαπλασιασμούς.

Έχοντας βρει ένα ιδιοδιάνυσμα του τριδιαγώνιου πίνακα $T \in \mathbb{R}^{n \times n}$, έτσι ώστε

$$Tu = \lambda u,$$

χρησιμοποιούμε, επίσης, το γεγονός ότι $Q^T A Q = T$ για να λάβουμε ότι

$$A Q u = \lambda Q u,$$

έτσι ώστε το διάνυσμα $Q u$ να είναι ένα ιδιοδιάνυσμα του πίνακα A . Χρησιμοποιώντας το θεώρημα 5.5, αυτό σημαίνει ότι το απαιτούμενο ιδιοδιάνυσμα του πίνακα A είναι

$$H_{(n,n-1)} \dots H_{(n,2)} u,$$

όπου οι πίνακες $H_{(n,j)} \in \mathbb{R}^{n \times n}$, $j = 2, \dots, n-1$ είναι πίνακες Householder. Για να πολλαπλασιάσουμε ένα διάνυσμα u με ένα Householder πίνακα $H = H(u)$ γράφουμε

$$Hx = (I - \alpha uu^T)x = x - \alpha(u^T x)u.$$

Υποθέτοντας ότι το $\alpha = 2/(u^T u)$ είναι γνωστό, απαιτείται ο υπολογισμός του βαθμωτού γινόμενου $u^T x$ και στη συνέχεια αφαιρούμε ένα πολλαπλάσιο του διανύσματος u από το διάνυσμα x . Αυτό προφανώς προϋποθέτει $2n$ πολλαπλασιασμούς. Συνεπώς, ο υπολογισμός του Qu απαιτεί μόνο $2n(n-2)$ πολλαπλασιασμούς και η δουλειά που προϋποθέτεται, σε όλη τη διαδικασία, είναι ανάλογη του n^2 αντί για n^3 . Πράγματι το συνολικό είναι μικρότερο από $2n(n-2)$, επειδή μια πιο προσεκτική μέτρηση μπορεί να χρησιμοποιήσει το γεγονός ότι πολλά από τα στοιχεία του διανύσματος u γνωρίζουμε ότι είναι μηδέν.

ΠΑΡΑΔΕΙΓΜΑ 8.2 : Χρησιμοποιούμε τον τριδιαγώνιο πίνακα του παραδείγματος 7.1 (πρώτος πίνακας).

$$\begin{pmatrix} 4.000 & -3.162 & 0.000 & 0.000 & 0.000 \\ -3.162 & 5.300 & -1.308 & 0.000 & 0.000 \\ 0.000 & -1.308 & 0.057 & 2.306 & 0.000 \\ 0.000 & 0.000 & 2.306 & 5.208 & -3.411 \\ 0.000 & 0.000 & 0.000 & -3.411 & -1.565 \end{pmatrix}$$

Για αυτόν τον πίνακα η μέθοδος της διχοτόμησης (παράδειγμα 6.4) έχει δώσει μία σωστή ιδιοτιμή την 8.094 (με ακρίβεια τριών δεκαδικών ψηφίων). Ξεκινώντας την αντίστροφη επανάληψη (σχέση (8.1)) με ένα τυχαία διαλεγμένο διάνυσμα $u^{(0)} \in \mathbb{R}_*^5$ βρίσκουμε ότι

$$u^{(1)} = (-0.0249, -0.0574, -0.3164, 0.4256, 0.8455)^T.$$

Επιτυχημένες επαναλήψεις δεν αφήνουν αλλαγές σε αυτό το διάνυσμα, όπως θα περιμέναμε, επειδή η ιδιοτιμή που χρησιμοποιήθηκε ήταν εντός του στρογγυλοποιημένου σφάλματος.

Συνεπώς αυτό είναι το ιδιοδιάνυσμα του παραπάνω τριδιαγώνιου πίνακα. Για να προκύψει το ιδιοδιάνυσμα του αρχικού πίνακα A (ο πίνακας πριν τη αναγωγή σε τριδιαγώνια μορφή – πίνακας A στο παράδειγμα 3.8)

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

Πολλαπλασιάζουμε το $u^{(1)}$ διαδοχικά με του τρεις πίνακες Householder που ορίζονται από τα διανύσματα (5.8), (5.7) και (5.6) (υπάρχουν στο παράδειγμα 5.6). Το αποτέλεσμα είναι το ιδιοδιάνυσμα

$$u = (-0.0249, -0.5952, -0.1920, -0.2885, 0.7246)^T .$$

Χρησιμοποιώντας το διάνυσμα και την ακριβώς υπολογισμένη ιδιοτιμή, μπορούμε να τσεκάρουμε το αποτέλεσμα και βρίσκουμε ότι τα στοιχεία του $Au - \lambda u$ είναι της ίδιας τάξης όπως το στρογγυλοποιημένο σφάλμα. •

9. ΤΟ ΠΗΛΙΚΟ RAYLEIGH

Σε αυτό το κεφάλαιο θα αναπτύξουμε μία απλή τεχνική, η οποία βασίζεται στην έννοια του πηλίκου Rayleigh, για να προκύψει μία ακριβή προσέγγιση σε μία ιδιοτιμή ενός συμμετρικού πίνακα όταν μία λογικά ακριβή προσέγγιση του συσχετιζόμενου ιδιοδιανύσματος είναι ήδη διαθέσιμη.

ΟΡΙΣΜΟΣ 9.1: Δοθέντος ενός διανύσματος $x \in \mathbb{R}_*^n$ και ενός πίνακα $A \in \mathbb{R}_{sym}^{n \times n}$, το συσχετιζόμενο **πηλίκο Rayleigh** $R(x)$ ορίζεται ως ένας πραγματικός αριθμός

$$R(x) = \frac{x^T A x}{x^T x} \quad (9.1)$$

Ολοφάνερα, αν $x \in \mathbb{R}_*^n$ είναι ένα ιδιοδιάνυσμα αντίστοιχο της ιδιοτιμής $\lambda \in \mathbb{R}$ του πίνακα $A \in \mathbb{R}_{sym}^{n \times n}$, τότε $R(x) = \lambda$. Πιο γενικά, αν x είναι ένα μη μηδενικό διάνυσμα του \mathbb{R}^n , τότε ένας αριθμός από πρόσθετες ιδιότητες του πηλίκου Rayleigh είναι άμεσα συμπεράσματα από την επέκταση του x σε αντιστοιχία με τα ιδιοδιανύσματα του πίνακα A .

ΘΕΩΡΗΜΑ 9.2: Υποθέτουμε ότι ο πίνακας $A \in \mathbb{R}_{sym}^{n \times n}$ έχει τα ορθοκανονικά ιδιοδιανύσματα $x^{(j)} \in \mathbb{R}_*^n$, $j=1,2,\dots,n$ και τις αντίστοιχες ιδιοτιμές $\lambda_j \in \mathbb{R}$, $j=1,2,\dots,n$. Αν το διάνυσμα $x \in \mathbb{R}_*^n$ εκφράζεται σε αντιστοιχία με τα ιδιοδιανύσματα $x^{(j)}$, $j=1,2,\dots,n$ ως

$$x = \sum_{j=1}^n a_j x^{(j)} \quad (9.2)$$

τότε

$$R(x) = \frac{\sum_{j=1}^n \lambda_j a_j^2}{\sum_{j=1}^n a_j^2} \quad (9.3)$$

Σημειώνουμε ότι $x^{(i)T} x^{(j)}$ είναι ίσο με το 1 όταν $i=j$ και ίσο με το 0 σε άλλη περίπτωση. Έτσι η σχέση (9.3) προκύπτει εύκολα αν βάλουμε τη σχέση (9.2) στην σχέση (9.1).

ΘΕΩΡΗΜΑ 9.3: Έστω $A \in \mathbb{R}_{sym}^{n \times n}$. Για κάθε διάνυσμα $x \in \mathbb{R}_*^n$,

$$\lambda_{\min} \leq R(x) \leq \lambda_{\max} \quad , \quad (9.4)$$

όπου $\lambda_{\min} \in \mathbb{R}$ και $\lambda_{\max} \in \mathbb{R}$ είναι αντίστοιχα η μικρότερη και η μεγαλύτερη από τις ιδιοτιμές του πίνακα A . Αυτοί οι περιορισμοί πετυχαίνονται όταν x είναι το αντίστοιχο ιδιοδιάνυσμα.

Απόδειξη: Οι ανισότητες προκύπτουν εύκολα από τη σχέση (9.3) σημειώνοντας ότι $\lambda_{\min} \leq \lambda_j \leq \lambda_{\max}$, $j=1,2,\dots,n$.

ΘΕΩΡΗΜΑ 9.4: Υποθέτουμε ότι $x \in \mathbb{R}_*^n$ είναι ένα κανονικοποιημένο διάνυσμα ώστε $\|x\|_2 = 1$. Υποθέτουμε, ακόμα, ότι $x^{(k)} \in \mathbb{R}_*^n$ είναι το k -οστό κανονικοποιημένο ιδιοδιάνυσμα του πίνακα $A \in \mathbb{R}_{sym}^{n \times n}$ ώστε

$$\|x - x^{(k)}\|_2 = O(\varepsilon)$$

για ένα μικρό $\varepsilon \in \mathbb{R}$. Τότε,

$$R(x) = \lambda_k + O(\varepsilon^2) \quad .$$

Απόδειξη: Από τη σχέση (9.2) παίρνουμε ότι $x^T x^{(k)} = a_k$ και συνεπώς

$$\begin{aligned} \|x - x^{(k)}\|_2^2 &= (x - x^{(k)})^T (x - x^{(k)}) \\ &= \|x\|_2^2 - 2x^T x^{(k)} + \|x^{(k)}\|_2^2 \\ &= 2(1 - a_k) \end{aligned}$$

Έτσι, $a_k = 1 + O(\varepsilon^2)$. Επιπλέον,

$$\begin{aligned} 1 = \|x\|_2^2 &= \sum_{j=1}^n a_j^2 \\ &= a_k^2 + \sum_{j \neq k} a_j^2 \\ &= 1 + O(\varepsilon^2) + \sum_{j \neq k} a_j^2 \quad . \end{aligned}$$

Συνεπώς, $a_j = O(\varepsilon)$ για όλα τα $j \neq k$. Το αποτέλεσμα στη συνέχεια ακολουθείται

από τη σχέση (9.3), η οποία (με $\sum_{j=1}^n a_j^2 = \|x\|_2^2 = 1$) δίνει ότι

$$\begin{aligned} R(x) &= \lambda_k a_k^2 + \sum_{j \neq k} \lambda_j a_j^2 \\ &= \lambda_k + O(\varepsilon^2) \quad . \bullet \end{aligned}$$

Αυτό το σημαντικό αποτέλεσμα σημαίνει ότι αν έχουμε μια αρκετά κοντινή προσέγγιση x ενός ιδιοδιανύσματος του πίνακα A , τότε το πηλίκο Rayleigh $R(x)$ δίνει πολύ εύκολα μία πολύ περισσότερο ακριβή προσέγγιση της αντίστοιχης ιδιοτιμής.

10 . ΑΝΑΛΥΣΗ ΤΗΣ ΔΙΑΤΑΡΑΧΗΣ

Είναι πολλές φορές απαραίτητο να έχουμε μία εκτίμηση του πόσο οι ιδιοτιμές και τα ιδιοδιανύσματα ενός πίνακα επηρεάζονται από τις αλλαγές στα στοιχεία. Κάποιες διαταραχές μπορεί να προκύψουν, για παράδειγμα, όταν τα στοιχεία του πίνακα λαμβάνονται από φυσικές μετρήσεις οι οποίες είναι μη ακριβείς, ή όταν παίρνω στοιχεία από πεπερασμένες διαφορετικές προσεγγίσεις μιας διαφορικής εξίσωσης. Τα δύο τελευταία θεωρήματα αυτού του κεφαλαίου απευθύνονται σε αυτό το θέμα. Ξεκινάμε με το ακόλουθο αποτέλεσμα.

ΘΕΩΡΗΜΑ 10.1: Έστω ο πίνακας $M \in \mathbb{R}_{sym}^{n \times n}$, με ιδιοτιμές λ_i και αντίστοιχα ορθοκανονικά ιδιοδιανύσματα u_i , $i=1,2,\dots,n$ και υποθέτουμε ότι $u \neq 0$ και w είναι διανύσματα του \mathbb{R}^n και ότι μ είναι ένας πραγματικός αριθμός τέτοιος ώστε

$$(M - \mu I)u = w \quad . \quad (10.1)$$

Τότε, το λιγότερο μία ιδιοτιμή λ_j του πίνακα M ικανοποιεί την σχέση

$$|\lambda_j - \mu| \leq \frac{\|w\|_2}{\|u\|_2} .$$

Απόδειξη: Αν το μ είναι ίσο με μία από τις ιδιοτιμές η απόδειξη είναι άμεση, άρα υποθέτουμε ότι $\mu \neq \lambda_k$, $k=1,2,\dots,n$. Στη συνέχεια γράφουμε τα διανύσματα u και w σαν γραμμικού συνδυασμούς των ιδιοδιανυσμάτων του πίνακα M , ώστε

$$u = \sum_{k=1}^n a_k u_k \quad , \quad w = \sum_{k=1}^n \beta_k u_k \quad .$$

Αντικαθιστώντας στη σχέση (10.1), θα εξισώσουμε τους συντελεστές των γραμμικώς ανεξάρτητων διανυσμάτων u_k , $k=1,2,\dots,n$ και θα συμπεράνουμε ότι

$$(\lambda_k - \mu)a_k = \beta_k \quad , \quad k=1,2,\dots,n \quad .$$

Τώρα, υποθέτουμε ότι λ_j είναι εκείνη η ιδιοτιμή που είναι πιο κοντά στην τιμή του μ . Αυτό σημαίνει ότι

$$|\lambda_j - \mu| \leq |\lambda_k - \mu| \quad , \quad k=1,2,\dots,n \quad .$$

Επειδή τα ιδιοδιανύσματα u_i , $i=1,2,\dots,n$ είναι ορθοκανονικά στον \mathbb{R}^n , έχουμε ότι

$$\sum_{k=1}^n a_k^2 = \|u\|_2^2 \quad , \quad \sum_{k=1}^n \beta_k^2 = \|w\|_2^2$$

Έτσι,

$$\sum_{k=1}^n \frac{\beta_k^2}{(\lambda_k - \mu)^2} = \|u\|_2^2$$

το οποίο δίνει ότι

$$\|w\|_2^2 = \sum_{k=1}^n \beta_k^2 \geq (\lambda_j - \mu)^2 \sum_{k=1}^n \frac{\beta_k^2}{(\lambda_k - \mu)^2} = (\lambda_j - \mu)^2 \|u\|_2^2 \quad ;$$

όπως απαιτείται. Και έτσι η απόδειξη τελείωσε. •

Πρόκειται, τώρα, να χρησιμοποιήσουμε αυτό το αποτέλεσμα για να δείξουμε ότι στην περίπτωση ενός συμμετρικού πίνακα A οι μικρές συμμετρικές διαταραχές στον A οδηγεί σε μικρές αλλαγές στις ιδιοτιμές του.

ΘΕΩΡΗΜΑ 10.2: (Θεώρημα Bauer-Fike (συμμετρική περίπτωση)) Υποθέτουμε ότι $A, E \in \mathbb{R}_{sym}^{n \times n}$ και ότι $B = A - E$. Υποθέτουμε, ακόμα, ότι οι ιδιοτιμές του πίνακα A δηλώνονται από τα λ_j , $j=1,2,\dots,n$ και ότι μ είναι μία ιδιοτιμή του πίνακα B . Τότε, το λιγότερο μία ιδιοτιμή λ_j του πίνακα A ικανοποιεί τη σχέση

$$|\lambda_j - \mu| \leq \|E\|_2 \quad .$$

Απόδειξη: Αυτό είναι μία άμεση συνέπεια του προηγούμενου θεωρήματος. Υποθέτουμε ότι u το κανονικοποιημένο ιδιοδιάνυσμα του B αντίστοιχο της ιδιοτιμής μ , τέτοιο ώστε $Bu = \mu u$. Τότε,

$$(A - \mu I)u = (B + E - \mu I)u = Eu \quad .$$

Στη συνέχεια , έπεται από το Θεώρημα 10.1 ότι υπάρχει μία ιδιοτιμή λ_j του πίνακα A τέτοια ώστε

$$|\lambda_j - \mu| \leq \|Eu\|_2 \leq \|E\|_2 \|u\|_2 = \|E\|_2$$

και έτσι η απόδειξη έχει ολοκληρωθεί. •

ΠΑΡΑΔΕΙΓΜΑ 10.3: Έστω ο 3×3 πίνακας Hilbert

$$A = \begin{pmatrix} 1 & 1/2 & 1/3 \\ 1/2 & 1/3 & 1/4 \\ 1/3 & 1/4 & 1/5 \end{pmatrix}$$

και η διαταραχή του

$$B = \begin{pmatrix} 1.0000 & 0.5000 & 0.3333 \\ 0.5000 & 0.3333 & 0.2500 \\ 0.3333 & 0.2500 & 0.2000 \end{pmatrix}$$

ο οποίος προκύπτει από τη στρογγυλοποίηση κάθε στοιχείου του πίνακα A με 4 δεκαδικά ψηφία.

Σε αυτήν την περίπτωση , $E = A - B$ και $\|E\|_2 = 3.3 \times 10^{-5}$. Έστω μ είναι μία ιδιοτιμή του πίνακα B . Στη συνέχεια σύμφωνα με το θεώρημα 10.2 τουλάχιστον μία από τις ιδιοτιμές $\lambda_1, \lambda_2, \lambda_3$ του πίνακα A θα ικανοποιεί την ανισότητα

$$|\lambda_j - \mu| \leq 3.3 \times 10^{-5} \quad (10.2)$$

Πράγματι , οι πραγματικές ιδιοτιμές του A και B αντίστοιχα είναι

$$\lambda_1 = 0.002687338072 \quad , \quad \lambda_2 = 0.1223270673 \quad , \quad \lambda_3 = 1.408318925$$

και

$$\mu_1 = 0.002664493933 \quad , \quad \mu_2 = 0.1223414532 \quad , \quad \mu_3 = 1.408294053 \quad .$$

Συμπεώς ,

$$\lambda_1 - \mu_1 = 2.29 \times 10^{-5} \quad , \quad \lambda_2 - \mu_2 = -1.44 \times 10^{-5} \quad , \quad \lambda_3 - \mu_3 = 2.49 \times 10^{-5} \quad ,$$

οι οποίες τιμές έρχονται σε συμφωνία με την ανισότητα στη σχέση (10.2). •

Στον τέλος αυτού του κεφαλαίου αναφέρουμε ότι το Θεώρημα 10.2 που είδαμε παραπάνω είναι μία ειδική περίπτωση του θεωρήματος Bauer-Fike , το οποίο παραθέτουμε παρακάτω.

ΘΕΩΡΗΜΑ 10.4: Υποθέτουμε ότι ο $A \in \mathbb{C}^{n \times n}$ είναι διαγωνοποιήσιμος , δηλαδή υπάρχει ένας ομαλός πίνακας $X \in \mathbb{C}^{n \times n}$ τέτοιος ώστε $X^{-1}AX = \Lambda$, όπου Λ διαγώνιος πίνακας του οποίου τα διαγώνια στοιχεία λ_j , $j=1, \dots, n$ είναι οι ιδιοτιμές του πίνακα A . Υποθέτουμε , ακόμα , ότι $E \in \mathbb{C}^{n \times n}$, $B = A - E$ και μ είναι μία ιδιοτιμή του πίνακα B . Τότε , τουλάχιστον μία ιδιοτιμή λ_j του πίνακα A ικανοποιεί τη σχέση

$$|\lambda_j - \mu| \leq \kappa_2(X) \|E\|_2 \quad ,$$

όπου $\kappa_2(X) = \|X\|_2 \|X^{-1}\|_2$ είναι ο αριθμός κατάστασης του πίνακα X στην 2-νόρμα $\|\cdot\|_2$ στον $\mathbb{C}^{n \times n}$.

Στην ειδική περίπτωση , όταν δηλαδή $A, E \in \mathbb{R}_{sym}^{n \times n}$, ο πίνακας X διαλέγεται να είναι ορθογώνιος δηλαδή $X^{-1} = X^T$. Συνεπώς , $\|X\|_2 = \|X^{-1}\|_2 = 1$ και έτσι $\kappa_2(X) = 1$, σύμφωνα με την ανισότητα που διατυπώνεται στο Θεώρημα 10.2 . Τα Θεωρήματα 10.2 και 10.4 εκτιμούν πόσο οι ιδιοτιμές του πίνακα A διαταράσσονται από τις αλλαγές στα στοιχεία του A . Η απάντηση την ερώτηση πόσο μεγάλες αλλαγές μπορεί να γίνουν στα ιδιοδιανύσματα , είναι πολύ δύσκολη και δεν την παραθέτουμε . (Αυτό το θέμα εξηγείται με λεπτομέρειες –βλέπε βιβλιογραφία το [5]).

11. ΚΩΔΙΚΕΣ ΜΑΤΛΑΒ ΚΑΙ ΠΑΡΑΔΕΙΓΜΑΤΑ

Σε αυτό το κεφάλαιο θα παραθέσουμε κάποια προγράμματα σε Matlab για κάποιες από τις μεθόδους που έχουμε αναφέρει μέχρι τώρα. Επίσης για όλα αυτά τα προγράμματα θα παραθέσουμε και παραδείγματα για να δούμε πως χρησιμοποιούνται στην πράξη. (Για τους παρακάτω κώδικες χρησιμοποιήσαμε το βιβλίο [11] σε συνδυασμό με την χρήση του Internet μέσω της μηχανής αναζήτησης Google).

ΚΛΑΣΣΙΚΗ ΜΕΘΟΔΟΣ JACOBI

Το πρώτο πρόγραμμα που θα παραθέσουμε είναι για την κλασική μέθοδο Jacobi του κεφαλαίου 3. Με την μέθοδο αυτή μπορούμε να υπολογίσουμε τις ιδιοτιμές ενός συμμετρικού πίνακα τάξης n . Η μέθοδος αυτή χρησιμοποιεί διαδοχικούς ορθογώνιους μετασχηματισμούς και παράγει μία ακολουθία από πίνακες, η οποία στο όριο προσεγγίζει ένα διαγώνιο πίνακα. Τα στοιχεία αυτού του διαγώνιου πίνακα αποτελούν και τις ζητούμενες ιδιοτιμές.

Παρακάτω υπάρχει το πρόγραμμα για την μέθοδο Jacobi σε Matlab:

```
function [D,sweep,psi]=cycjacobi(A,tol,nmax)
%CYCJACOBI Cyclic Jacobi method
%[D,SWEEP,PSI]=CYCJACOBI(A,TOL) computes the eigenvalues D of
the symmetric
%matrix A. TOL specifies the tolerance of the method.
PSI=PSINORM(D) and
%SWEEP is the number of sweeps. NMAX specifies the maximum
number of
%iterations
[n,m]=size(A);
if n~=m, error('Only square matrices'); end
D=A;
psi=norm(A,'fro');
epsi=tol*psi;
psi=psinorm(D);
sweep=0;
iter=0;
while psi>epsi & iter<=nmax
    iter=iter+1;
    sweep=sweep+1;
    for p=1:n-1
        for q=p+1:n
            [c,s]=symschur(D,p,q);
            [D]=gacol(D,c,s,1,n,p,q);
            [D]=garow(D,c,s,p,q,1,n);
        end
    end
end
```

```

    end
    psi=psinorm(D);
end
return

```

Παρατηρούμε ότι το παραπάνω πρόγραμμα καλεί κάποια άλλα προγράμματα. Οι υπορουτίνες αυτές παραθέτονται παρακάτω:

```

function [psi]=psinorm(A)
%PSINORM Evaluation of Psi(A)
[n,m]=size(A);
if n~=m ,error('only square matrices');end
psi=0;
for i=1:n-1
    j=[i+1:n];
    psi=psi+sum(A(i,j).^2+A(j,i).^2');
end
psi=sqrt(psi);
return

```

```

function [c,s]=symschur(A,p,q)
%SYMSCHUR Evaluation of parameters c and s
if A(p,q)==0
    c=1;
    s=0;
else
    eta=(A(q,q)-A(p,p))/(2*A(p,q));
    if eta>=0
        t=1/(eta+sqrt(1+eta^2));
    else
        t=-1/(-eta+sqrt(1+eta^2));
    end
    c=1/sqrt(1+t^2);
    s=c*t;
end
return

```

```

function [M]=garow(M,c,s,i,k,j1,j2)
% GAROW Product of the transpose of Givens rotation matrix
with M
for j=j1:j2
    t1=M(i,j);
    t2=M(k,j);
    M(i,j)=c*t1-s*t2;
    M(k,j)=s*t1+c*t2;
end
return

```

```
function [M]=gacol (M,c,s,j1,j2,i,k)
% GACOL Product of M with a Givens rotation matrix
for j=j1:j2
    t1=M(j,i);
    t2=M(j,k);
    M(j,i)=c*t1-s*t2;
    M(j,k)=s*t1+c*t2;
end
return
```

Το παραπάνω πρόγραμμα χρησιμοποιεί σαν είσοδο ένα συμμετρικό πίνακα A τάξης n καθώς και ένα όριο ανοχής tol . Σαν έξοδο το πρόγραμμα μας δίνει ένα πίνακα D διαγώνιο (του οποίου τα στοιχεία αποτελούν τις ιδιοτιμές του αρχικού πίνακα A). Επίσης μας δίνει το $sweep$ που είναι ο αριθμός των σαρώσεων καθώς και την ποσότητα psi που είναι η νόρμα του πίνακα D (όπως υπολογίζεται στην υπορουτίνα `psinorm`). Για να γίνει πιο κατανοητό παραθέτουμε ένα παράδειγμα. Θεωρούμε τον πίνακα A που χρησιμοποιήσαμε στο κεφάλαιο 3 –Παράδειγμα 3.8.

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

Χρησιμοποιώντας αυτόν τον πίνακα εκτελούμε στο περιβάλλον του *Matlab* την παρακάτω εντολή:

```
>> [D,sweep,psi]=cycjacobi(A,0.001,10)
```

Το αποτέλεσμα που λαμβάνουμε είναι το παρακάτω:

$D =$

```
8.0937 -0.0000 0.0000 0.0000 -0.0000
-0.0000 1.6896 0.0000 -0.0000 -0.0000
0.0000 0.0000 -0.6711 0.0000 -0.0000
0.0000 -0.0000 0.0000 7.1703 -0.0000
-0.0000 -0.0000 -0.0000 -0.0000 -3.2824
```

sweep =

3

psi =

5.3318e-006

Ο πίνακας *D* μας δίνει τον ζητούμενο διαγώνιο πίνακα. Το *sweep* είναι ίσο με 3 δηλαδή στην τρίτη σάρωση τα μη διαγώνια στοιχεία είναι όλα μηδέν και τα διαγώνια μας δίνουν τις ζητούμενες ιδιοτιμές. Παρατηρούμε ότι τα παραπάνω αποτελέσματα είναι πολύ κοντά με τα αποτελέσματα του παραδείγματος 3.8, τα οποία χρησιμοποιούν την σειριακή μέθοδο *Jacobi*. •

ΜΕΘΟΔΟΣ HOUSEHOLDER

Το δεύτερο πρόγραμμα που θα παραθέσουμε είναι για την μέθοδο *Householder* που αναλύσαμε στο κεφάλαιο 5. Η μέθοδος αυτή, χρησιμοποιώντας πίνακες *Householder* και ορθογώνιους μετασχηματισμούς, μετασχηματίζει έναν συμμετρικό πίνακα σε τριδιαγώνια μορφή.

Παρακάτω παραθέτουμε τον πρόγραμμα για την μέθοδο *Householder* σε *Matlab*:

```
function [A, V] = Householder(A)
% Reduction of the square matrix A to the upper
% Hessenberg form using Householder reflectors.
% Matrix A is overwritten with its
% upper Hessenberg form.
[m,n] =size(A);
if A == triu(A,-1)
V = eye(m);
return
end
V = [];
for k=1:m-2
x = A(k+1:m,k);
v = Housv(x);
A(k+1:m,k:m) = A(k+1:m,k:m) - 2*v*(v'*A(k+1:m,k:m));
A(1:m,k+1:m) = A(1:m,k+1:m) - 2*(A(1:m,k+1:m)*v)*v';
v = [zeros(k,1);v];
V = [V v];
end
```


Το παραπάνω πρόγραμμα χρησιμοποιεί ένα υποπρόγραμμα (*Housv*) που είναι το παρακάτω:

```
function u = Housv(x)
% Householder reflection unit vector u from the vector x.
m = max(abs(x));
u = x/m;
if u(1) == 0
su = 1;
else
su = sign(u(1));
end
u(1) = u(1)+su*norm(u);
u = u/norm(u);
u = u(:);
```

Το πρόγραμμα που επιλύει την μέθοδο *Householder* δέχεται σαν είσοδο ένα συμμετρικό πίνακα *A* και δίνει σαν έξοδο τον αρχικό πίνακα στην τριδιαγώνια μορφή του. Στη συνέχεια η τριδιαγώνια μορφή του πίνακα θα χρησιμοποιηθεί για να δώσει τις ιδιοτιμές του πίνακα μέσα από διάφορες μεθόδους όπως τη μέθοδο της διχοτόμησης (κεφάλαιο 6) και τον αλγόριθμο *QR* (κεφάλαιο 7).

Χρησιμοποιώντας τον συμμετρικό πίνακα *A* του παραδείγματος 3.8 θα χρησιμοποιήσουμε το παραπάνω πρόγραμμα για την μέθοδο *Householder* ώστε να τον μετασχηματίσουμε σε τριδιαγώνια μορφή. Έχουμε

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

Χρησιμοποιώντας την παρακάτω εντολή σε περιβάλλον *Matlab*

```
>> Householder(A)
```

παίρνουμε το παρακάτω αποτέλεσμα:

```
ans =
```

```

4.0000 -3.1623 -0.0000 -0.0000 0.0000
-3.1623 5.3000 -1.3077 -0.0000 0.0000
0.0000 -1.3077 0.0567 2.3059 -0.0000
0.0000 0 2.3059 5.2085 -3.4113
0.0000 0.0000 0.0000 -3.4113 -1.5652
    
```

Ο παραπάνω πίνακας είναι ο ζητούμενος τριδιαγώνιος πίνακας. (Μπορούμε να δούμε και τη διαδικασία λίγο πιο αναλυτικά στο παράδειγμα 5.6 το οποίο δίνει το ίδιο αποτέλεσμα).•

ΑΛΓΟΡΙΘΜΟΣ QR

Εδώ θα παραθέσουμε το πρόγραμμα του αλγορίθμου *QR* που αναλύθηκε στο κεφάλαιο 7. Αυτόν τον αλγόριθμο τον χρησιμοποιούμε για να βρούμε τις ιδιοτιμές ενός πίνακα συμμετρικού και τριδιαγώνιου (βέβαια ο αλγόριθμος αυτός μπορεί να χρησιμοποιηθεί και για πλήρεις πίνακες απλά είναι πιο αποδοτικό να χρησιμοποιήσουμε πρώτα την μέθοδο *Householder* για να τον μετασχηματίσουμε σε τριδιαγώνια μορφή). Ο αλγόριθμος αυτός δέχεται σαν είσοδο ένα πίνακα A συμμετρικό και τριδιαγώνιο και υπολογίζει μια ακολουθία πινάκων $A^{(k)}$ (θέτοντας $A^{(0)} = A$ και επιλέγοντας μία μετατόπιση σ_k (ή αλλιώς τη συμβολίζω με μ_k) τέτοια ώστε όλοι οι πίνακες της να έχουν τις ίδιες ιδιοτιμές με τον πίνακα A . Επίσης όλοι οι πίνακες της ακολουθίας $A^{(k)}$ είναι συμμετρικοί και τριδιαγώνιοι. Τέλος ο αλγόριθμος θα μας δώσει σαν αποτέλεσμα ένα διαγώνιο πίνακα του οποίου τα διαγώνια στοιχεία θα είναι οι ιδιοτιμές του αρχικού συμμετρικού τριδιαγώνιου πίνακα A .

Παρακάτω υπάρχει το πρόγραμμα για τον αλγόριθμο *QR* σε *Matlab*:

```

function [D] = shifted_qr_iteration(A,M)
%
% function [D] = shifted_qr_iteration(A,M)
%
% Given a square matrix A, this function performs M iterations
% of the shifted QR algorithm to produce a diagonal matrix D,
% with A's eigenvalues on the main diagonal.
%
%
[m,n]=size(A);
    
```

```

if m~=n
error('Matrix must be square!');
else
D = A;
for i=1:M
sigma = D(end,end);
[Q,R]=qr(D-sigma*eye(size(D))); % built-in Matlab command
% for QR factorization
D = R*Q+sigma*eye(size(D));
end
end

```

Το παραπάνω πρόγραμμα δέχεται σαν είσοδο ένα συμμετρικό τριδιαγώνιο πίνακα A τάξης n καθώς και ένα αριθμό M που είναι ο αριθμός των επαναλήψεων που θα εκτελεστεί το πρόγραμμα. Σαν έξοδο θα πάρουμε ένα διαγώνιο πίνακα D του οποίου τα διαγώνια στοιχεία θα είναι οι ζητούμενες ιδιοτιμές του πίνακα A . Για να γίνει πιο κατανοητό παραθέτουμε ένα παράδειγμα χρησιμοποιώντας σαν αρχικό πίνακα A του παραδείγματος 3.8 που είναι ο παρακάτω:

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

Αυτός ο πίνακας είναι συμμετρικός αλλά είναι σε πλήρη μορφή και όχι τριδιαγώνιος. Χρησιμοποιώντας την μέθοδο Householder (παράδειγμα 5.6) τον μετατρέπουμε σε τριδιαγώνια μορφή και παίρνουμε τον παρακάτω πίνακα:

$$A = \begin{pmatrix} 4.000 & -3.162 & 0.000 & 0.000 & 0.000 \\ -3.162 & 5.300 & -1.308 & 0.000 & 0.000 \\ 0.000 & -1.308 & 0.057 & 2.306 & 0.000 \\ 0.000 & 0.000 & 2.306 & 5.208 & -3.411 \\ 0.000 & 0.000 & 0.000 & -3.411 & -1.565 \end{pmatrix}$$

Χρησιμοποιώντας τον παραπάνω συμμετρικό και τριδιαγώνιο πίνακα θα εκτελέσουμε στο περιβάλλον του Matlab την παραπάνω εντολή που θα υλοποιήσει τον αλγόριθμο QR .

`>> [D]=shifted_qr_iteration(A,20)`

Το αποτέλεσμα που θα πάρουμε είναι ο παρακάτω πίνακας:

$D =$

```

8.0886 -0.0670 0.0000 0.0000 -0.0000
-0.0670 7.1746 -0.0000 0.0000 0.0000
0 -0.0000 1.6899 0.0000 -0.0000
0 0 0.0000 -0.6711 0.0000
0 0 0 0 -3.2821
    
```

Τα διαγώνια στοιχεία του πίνακα D είναι οι ιδιοτιμές του τριδιαγώνιου πίνακα A . Παρατηρούμε ότι αυτές οι ιδιοτιμές που βρήκαμε συμπίπτουν με τα αποτελέσματα του παραδείγματος 3.8 που χρησιμοποιεί την μέθοδο *Jacobi*. •

ΜΕΘΟΔΟΣ ΤΩΝ ΔΥΝΑΜΕΩΝ ΚΑΙ ΑΝΤΙΣΤΡΟΦΗ ΕΠΑΝΑΛΗΨΗ

Σε αυτό το σημείο θα παραθέσουμε του κώδικες για την μέθοδο των δυνάμεων καθώς και την μέθοδο της αντίστροφης επανάληψης που αναλύσαμε στο κεφάλαιο 8. Για να δούμε και πρακτικά πως λειτουργούν αυτά τα προγράμματα θα παραθέσουμε και ένα παράδειγμα.

Παρακάτω παραθέτουμε τα προγράμματα για τις δύο μεθόδους σε *Matlab*. Πρώτα το πρόγραμμα για την μέθοδο των δυνάμεων και στη συνέχεια το πρόγραμμα για την μέθοδο της αντίστροφης επανάληψης.

```

function [lambda,x,iter,relres]=powerm(A,z0,tol,nmax)
%POWERM Power mmethod
%[LAMBDA,X,ITER,RELRES]=POWERM(A,Z0,TOL,NMAX) computes the
    
```

```

%eigenvalue LAMBDA of largest module of the matrix A and the
corresponding
%eigenvector X of unit norm. TOL specifies the tolerance of
the method.
%NMAX specifies the maximum number of iterations. Z0 specifies
the initial
%guess. ITER is the iteration number at which X is computed.
q=z0/norm(z0); q2=q;
relres=tol+1;iter=0;z=A*q;
while relres(end)>=tol & iter<=nmax
    q=z/norm(z);z=A*q;
    lambda=q'*z;x=q;
    z2=q2'*A;q2=z2/norm(z2);q2=q2';
    y1=q2;costheta=abs(y1'*x);
    if costheta>=5e-2
        iter=iter+1;
        temp=norm(z-lambda*q)/costheta;
        relres=[relres;temp];
    else
        fprintf('Multiple eigenvalue');break;
    end
end
return

```

```

function [sigma,x,iter,relres]=invpower(A,z0,mu,tol,nmax)
%INVPOWER inverse power method
%[SIGMA,X,ITER,RELRES]=INVPOWER(A,Z0,MU,TOL,NMAX) computes the
%eigenvalue LAMBDA of smallest module of the matrix A and the
%corresponding eigenvector X of unit norm. TOL specifies
%the tolerance of the method. NMAX specifies the maximum
number
%of iterations. Z0 specifies the initial guess . MU is the
shift.
%ITER is the iteration number at which X is computed.
M=A-mu*eye(size(A));[L,U,P]=lu(M);
q=z0/norm(z0);q2=q';sigma=[];
relres=tol+1;iter=0;
while relres(end)>=tol&iter<=nmax
    iter=iter+1;
    b=P*q;
    y=L\b;z=U\y;
    q=z/norm(z);z=A*q;sigma=q'*z;
    b=q2';y=U\b;w=L\y;
    q2=w'*P;q2=q2/norm(q2);costheta=abs(q2*q);
    if costheta>=5e-2
        temp=norm(z-sigma*q)/costheta;relres=[relres,temp];
    else
        fprintf('Multiple eigenvalue');break;
    end
    x=q;
end
return

```

Το πρώτο πρόγραμμα για την μέθοδο των δυνάμεων δέχεται σαν είσοδο ένα πίνακα A , ένα αρχικό διάνυσμα z_0 , την ανοχή tol της μεθόδου καθώς και τον μέγιστο αριθμό επαναλήψεων n_{max} . Σαν έξοδο παίρνουμε το $lambda$ που είναι η μεγαλύτερη ιδιοτιμή (σε απόλυτη τιμή) του πίνακα A , το x που είναι το αντίστοιχο ιδιοδιάνυσμα καθώς και την τιμή $iter$ που είναι ο αριθμός της επανάληψης στην οποία υπολογίστηκε το ιδιοδιάνυσμα x .

Θεωρούμε τον παρακάτω πίνακα:

$$A = \begin{pmatrix} 10 & 6 \\ 2 & 5 \end{pmatrix} \quad (**)$$

Αρχικά με μια απλή εντολή μπορούμε να βρούμε τις ιδιοτιμές και τα αντίστοιχα ιδιοδιανύσματα του πίνακα για να μπορέσουμε να τα συγκρίνουμε με τα αποτελέσματα που θα πάρουμε εκτελώντας τις μεθόδους. Έτσι έχουμε:

```
>> [X,D] = eig(A)
```

```
X = 0.9590 -0.6632
     0.2832 0.7485
```

```
D = 11.7720 0
     0 3.2280
```

Ο πίνακας D είναι ένας διαγώνιος πίνακας του οποίου τα διαγώνια στοιχεία είναι οι ιδιοτιμές του πίνακα A και οι στήλες του πίνακα X είναι τα αντίστοιχα ιδιοδιανύσματα.

Χρησιμοποιώντας τον παραπάνω πίνακα A με αρχικό διάνυσμα $z_0 = (1,1)^T$ και $tol = 1e-06$, $n_{max} = 100$ εκτελούμε στο περιβάλλον Matlab την παρακάτω εντολή που εκτελεί την μέθοδο των δυνάμεων

```
>> [lambda,x,iter]=powerm(A,[1;1],1e-06,100)
```

και παίρνουμε τα παρακάτω αποτελέσματα :

lambda =

11.7720

x =

0.9590

0.2832

iter =

12

Παρατηρούμε ότι το πρόγραμμα μας επιστρέφει την μεγαλύτερη ιδιοτιμή $lambda$ καθώς και το αντίστοιχο ιδιοδιάνυσμα x , το οποίο ανακτήθηκε μετά από 12 επαναλήψεις. Παρατηρούμε ότι τα αποτελέσματα συμβαδίζουν με τα παραπάνω που πήραμε από την εντολή `eig`.

Στη συνέχεια θα χρησιμοποιήσουμε τον ίδιο πίνακα A (**), για να δούμε πως λειτουργεί η μέθοδος της αντίστροφης επανάληψης. Το πρόγραμμα αυτό έχει τις ίδιες εισόδους με το προηγούμενο απλά χρειάζεται και άλλη μία. Δέχεται σαν είσοδο το μ που είναι μία προσέγγιση της ιδιοτιμής. Σαν έξοδο μας επιστρέφει την ιδιοτιμή (που είναι πιο κοντά στην προσέγγιση που δώσαμε) καθώς και το αντίστοιχο ιδιοδιάνυσμα της μαζί με τον αριθμό των επαναλήψεων που χρησιμοποιήθηκαν για να ανακτηθεί. Έτσι, όμοια με πριν εκτελούμε την παρακάτω εντολή

```
>> [lambda,x,iter]=invpower(A,[1;1],3.23,1e-06,100)
```

και παίρνουμε τα παρακάτω αποτελέσματα

lambda =

3.2280

$x =$

-0.6632

0.7485

$iter =$

2

Παρατηρούμε ότι το πρόγραμμα μας επιστρέφει την ιδιοτιμή που είναι πιο κοντά στην προσέγγιση 3.23 που δώσαμε καθώς και το αντίστοιχο ιδιοδιάνυσμα, το οποίο υπολογίστηκε μετά από 2 επαναλήψεις. Τα αποτελέσματα συμβαδίζουν πάλι με αυτά που δίνει η εντολή eig.

Παρατήρηση: Όπως αναφέραμε και στο κεφάλαιο 8, υπάρχουν δύο δρόμοι που μπορούμε να χρησιμοποιήσουμε την αντίστροφη επανάληψη. Ο πρώτος είναι να χρησιμοποιήσουμε ένα πλήρη πίνακα. Σε αυτήν την περίπτωση χρησιμοποιούμε το παραπάνω πρόγραμμα το οποίο στηρίζεται στην LU παραγοντοποίηση. Ο δεύτερος χρησιμοποιεί πρώτα την μέθοδο Householder για να μετασχηματίσει τον πίνακα σε τριδιαγώνια μορφή και στη συνέχεια τον αλγόριθμο του Thomas (βλέπε παράδειγμα 8.2).

ΠΗΛΙΚΟ RAYLEIGH

Σαν τελευταίο πρόγραμμα σε αυτό το κεφάλαιο θα παραθέσουμε την τεχνική που βασίζεται στο πηλίκo Rayleigh όπου δοθέντος μιας προσέγγισης σε ένα ιδιοδιάνυσμα προκύπτει μια ακριβής προσέγγιση της αντίστοιχης ιδιοτιμής του. Το πρόγραμμα που εκτελεί αυτή τη μέθοδο είναι το παρακάτω:

```
function [lambda,x] = rq_method(A,x0,tol)
% function [lambda,x] = rq_method(A,x0,tol)
%
% This function finds the dominant eigenvalue and
corresponding
% (normalized) eigenvector of a matrix A, using the Rayleigh-
% Quotient iteration method.
%
% x0 is an initial guess for the eigenvector and tol is the
```



```

% tolerance within which we want two successive approximation
to
% be.
x0 = x0/norm(x0,2);
if norm(A*x0) <= eps
error('Bad choice for initial guess')
disp('')
return;
end
[n,m] = size(A);
if n~=m
error('Matrix must be square')
disp('')
return;
end
NORM = 1;
iter = 0;
while NORM > tol
rj = x0'*A*x0;
xhj = (A-rj*eye(n))\x0;
2
xj = xhj/norm(xhj,2);
NORM = norm(x0-xj)/norm(xj);
x0 = xj;
iter = iter + 1;
end
lambda=rj;
x=xj/norm(xj,2);
disp('Number of Iterations :')
disp(iter)

```

Το παραπάνω πρόγραμμα δέχεται σαν είσοδο έναν πίνακα A , ένα αρχικό διάνυσμα x_0 (προσέγγιση ιδιοδιανύσματος) καθώς και την ανοχή tol της μεθόδου. Σαν έξοδο παίρνουμε την ιδιοτιμή $lambda$ και το αντίστοιχο κανονικοποιημένο ιδιοδιάνυσμα x . Χρησιμοποιώντας το πίνακα A (**) του παραδείγματος της μεθόδου των δυνάμεων εκτελούμε την παρακάτω εντολή χρησιμοποιώντας σαν αρχική προσέγγιση ιδιοδιανύσματος το $x_0 = (1,1)^T$

```
>> [lambda,x] = rq_method(A,[1;1],1e-06)
```

και παίρνουμε τα παρακάτω αποτελέσματα

Number of Iterations :

6

lambda =

11.7720

x =

0.9590

0.2832

Παρατηρούμε ότι τα αποτελέσματα ταιριάζουν με εκείνα που πήραμε στα προηγούμενα παραδείγματα για την μέθοδο των δυνάμεων παραπάνω. Για επαλήθευση εκτελούμε την εντολή

>> [X,D] = eig(A)

X = 0.9590 -0.6632

0.2832 0.7485

D = 11.7720 0

0 3.2280

Τα αποτελέσματα συμπίπτουν.

(Περισσότερες λεπτομέρειες για το μαθηματικό λογισμικό Matlab βλέπε βιβλιογραφία το [10]).

ΠΑΡΑΡΤΗΜΑ

1 . ΒΑΣΙΚΟΙ ΟΡΙΣΜΟΙ

Σε αυτό το κομμάτι θα αναφέρουμε κάποιους ορισμούς από όρους που χρησιμοποιήθηκαν σε αυτό το βιβλίο. Θα ορίσουμε κάποιες βασικές έννοιες για πίνακες , νόρμες πινάκων ,διανύσματα και σφάλματα.

Τετραγωνικοί πίνακες : Πίνακες που έχουν ίσο αριθμό γραμμών και στηλών . Θα λέμε τετραγωνικό πίνακα τάξης n .

Διαγώνιος πίνακας (diagonal matrix): Ένας πίνακας ονομάζεται διαγώνιος όταν τα μόνα μη μηδενικά στοιχεία του βρίσκονται πάνω στην κύρια διαγώνιό του.

Άνω-τριγωνικός πίνακας (upper triangular): Ένας πίνακας που έχει όλα τα στοιχεία κάτω από την κύρια διαγώνιό του ίσα με το μηδέν ονομάζεται άνω-τριγωνικός.

Κάτω-τριγωνικός πίνακας (lower triangular): Ένας πίνακας που έχει όλα τα στοιχεία πάνω από την κύρια διαγώνιό του ίσα με το μηδέν ονομάζεται κάτω-τριγωνικός.

Μοναδιαίος πίνακας: Ένας διαγώνιος πίνακας I ,του οποίου όλα τα διαγώνια στοιχεία είναι ίσα με τη μονάδα.

Τριδιαγώνιος πίνακας (tridiagonal): Ένας τετραγωνικός πίνακας ο οποίος έχει μη μηδενικά στοιχεία στην κύρια διαγώνιό του και στις γειτονικές άνω και κάτω διαγωνίους ως προς αυτήν , ενώ τα υπόλοιπα στοιχεία είναι ίσα με το μηδέν , ονομάζεται τριδιαγώνιος .

Ανάστροφος πίνακας (transpose): Έστω ένας πίνακας A $n \times m$. Τότε , ο ανάστροφος πίνακας A^T είναι ένας πίνακας $m \times n$ του οποίου οι στήλες είναι οι γραμμές του A .

Συμμετρικός πίνακας (symmetric): Ένας τετραγωνικός πίνακας A ονομάζεται συμμετρικός αν ισχύει $A^T = A$.

Ομαλός ή μη ιδιάζων πίνακας (nonsingular): Ένας τετραγωνικός πίνακας A λέγεται ομαλός αν υπάρχει πίνακας B με την ιδιότητα $AB = BA = I$,όπου I μοναδιαίος. Αυτός ο πίνακας B ,αν υπάρχει , θα είναι μοναδικός και θα συμβολίζεται με A^{-1} και θα τον λέμε αντίστροφο του πίνακα A . Σε άλλη περίπτωση ο πίνακας A θα λέγεται μη ομαλός ή ιδιάζων (singular).

Ορθογώνιος πίνακας (orthogonal) : Ένας πίνακας A για τον οποίο ισχύει η σχέση $A^T A = AA^T = I$ (ή $A^{-1} = A^T$) ονομάζεται ορθογώνιος.

Σύνθετος πίνακας: Ένας πίνακας λέγεται σύνθετος αν τα στοιχεία του είναι επίσης πίνακες και έχουν την ιδιότητα: Τα στοιχεία-πίνακες που βρίσκονται στην ίδια γραμμή να έχουν τον ίδιο αριθμό γραμμών και τα στοιχεία πίνακες στην ίδια στήλη να έχουν τον ίδιο αριθμό στηλών.

Διαγωνοποιήσιμος πίνακας: Ένας πίνακας A καλείται διαγωνοποιήσιμος αν υπάρχει αντιστρέψιμος πίνακας P τέτοιος ώστε ο πίνακας $P^{-1}AP$ να είναι διαγώνιος.

Νόρμες διανύσματος (norms) : Έστω V_n το σύνολο όλων των n -διάστατων διανυσμάτων και \mathbb{R} το σύνολο των πραγματικών αριθμών . Μία νόρμα $\|\cdot\|$ στο V_n είναι μία συνάρτηση με πεδίο ορισμού το V_n και τιμές στο \mathbb{R} η οποία πληρεί τις παρακάτω ιδιότητες:

1. $\|x\| \geq 0$, $x \in V_n$
2. $\|x\| = 0$, αν και μόνο αν $x = 0$
3. $\|ax\| = |a| \|x\|$, $a \in \mathbb{R}$, $x \in V_n$
4. $\|x + y\| \leq \|x\| + \|y\|$, $x, y \in V_n$

Στον χώρο αυτόν ορίζονται οι παρακάτω νόρμες:

1. $\|x\|_\infty = \max_{1 \leq i \leq n} |x_i|$ (άπειρη νόρμα)
2. $\|x\|_1 = \sum_{i=1}^n |x_i|$ (ένα νόρμα)
3. $\|x\|_2 = \left(\sum_{i=1}^n x_i^2 \right)^{1/2}$ (ευκλείδεια νόρμα)
4. $\|x\|_p = \left(\sum_{i=1}^n x_i^p \right)^{1/p}$

Νόρμες πίνακα (norms) : Η νόρμα ενός $n \times n$ πίνακα A με στοιχεία a_{ij} , $i, j = 1, 2, \dots, n$ συμβολίζεται με $\|A\|$ και είναι ένας μη αρνητικός αριθμός.

Έστω $M_n = \mathbb{R}^{n \times n}$ το σύνολο των πινάκων $n \times n$. Μία νόρμα στο M_n είναι μία συνάρτηση με πεδίο ορισμού το M_n και σύνολο τιμών το \mathbb{R} , που έχει τις παρακάτω ιδιότητες:

1. $\|A\| \geq 0, A \in M_n$
2. $\|A\| = 0$, αν και μόνο αν $A = 0$ (μηδενικός πίνακας)
3. $\|aA\| = |a|\|A\|$, $a \in \mathbb{R}$, $A \in M_n$
4. $\|A+B\| \leq \|A\| + \|B\|$, $A, B \in M_n$
5. $\|AB\| \leq \|A\|\|B\|$, $A, B \in M_n$

Οι αμέσως παρακάτω σχέσεις ορίζουν νόρμες πινάκων:

1. $\|A\|_{\infty} = \max_{1 \leq i \leq n} \sum_{j=1}^n |a_{ij}|$
2. $\|A\|_1 = \max_{1 \leq j \leq n} \sum_{i=1}^n |a_{ij}|$
3. $\|A\|_2 = \left(\sum_{1 \leq i, j \leq n} a_{ij}^2 \right)^{1/2}$

Ορθογώνια διανύσματα: Δύο μη μηδενικά διανύσματα x, y ονομάζονται ορθογώνια όταν το εσωτερικό τους γινόμενο είναι ίσο με το μηδέν και συμβολίζω $x \perp y \Leftrightarrow xy = 0$.

Ορθοκανονικοποιημένα διανύσματα: Τα ορθογώνια διανύσματα που έχουν και μέτρο ίσο με τη μονάδα ονομάζονται ορθοκανονικοποιημένα.

Γραμμικώς εξαρτημένα διανύσματα: Τα διανύσματα x_1, x_2, \dots, x_k θα λέγονται γραμμικώς εξαρτημένα αν και μόνο αν υπάρχουν συντελεστές $\lambda_1, \lambda_2, \dots, \lambda_k$, όχι όλοι μηδέν, τέτοιοι ώστε να ισχύει $\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_k x_k = 0$.

Γραμμικώς ανεξάρτητα διανύσματα: Τα διανύσματα x_1, x_2, \dots, x_k θα λέγονται γραμμικώς ανεξάρτητα αν και μόνο αν η σχέση $\lambda_1 x_1 + \lambda_2 x_2 + \dots + \lambda_k x_k = 0$ ισχύει μόνο όταν $\lambda_1 = \lambda_2 = \dots = \lambda_k = 0$.

Συνεπίπεδα διανύσματα: Θα λέμε ότι τρία διανύσματα είναι συνεπίπεδα αν και μόνο αν οι διευθύνσεις τους είναι παράλληλες προς το ίδιο επίπεδο. Πιο απλά, δύο μη παράλληλα διανύσματα p, q τοποθετημένα με την ίδια αρχή ορίζουν ένα επίπεδο. Ένα τρίτο διάνυσμα r είναι συνεπίπεδο με αυτά αν τοποθετημένο με την

αρχή βρίσκεται στο ίδιο επίπεδο. Πρακτικά ένα διάνυσμα r μπορεί να είναι συνεπίπεδο με δύο μη παράλληλα διανύσματα p, q είτε αν μπορεί να γραφεί σαν γραμμικός συνδυασμός των p, q δηλαδή $r = \alpha p + \beta q$ με $\alpha, \beta \in \mathbb{R}$ είτε αν μηδενίζεται η ορίζουσα των συντεταγμένων των τριών διανυσμάτων.

Σφάλματα: Είναι βασικό να αναφέρουμε ότι στην εφαρμογή της Αριθμητικής Ανάλυσης πρέπει να συνειδητοποιήσουμε ότι οι υπολογιστικές λύσεις δεν είναι ακριβείς και συνήθως παρατηρούμε διαφόρων ειδών σφάλματα. Ένα από αυτά είναι το σφάλμα στρογγύλευσης. Ο προγραμματισμός μιας μεθόδου και η χρήση υπολογιστή για υπολογισμό αριθμητικών λύσεων, συνεπάγεται αποθήκευση δεδομένων στον υπολογιστή. Έτσι όταν ένας αριθμός αποθηκευτεί στον υπολογιστή θα αντικατασταθεί με κάποιον άλλο που είναι μία προσέγγιση του. Αυτή, λοιπόν, η διαδικασία ορίζει ένα σφάλμα το οποίο ονομάζεται σφάλμα στρογγύλευσης (round error) και εξαρτάται κυρίως με τον υπολογιστή.

(Για τους ορισμούς αυτούς αλλά και για περισσότερες λεπτομέρειες σε θέματα πινάκων, διανυσμάτων, σφαλμάτων κα βλέπε βιβλιογραφία το [8]-(κεφάλαια 1,5,8) και [9]-(κεφάλαιο3)

ΒΙΒΛΙΟΓΡΑΦΙΑ

- [1]. *AN INTRODUCTION TO NUMERICAL ANALYSIS*
 ENDRE SULI AND DAVID MAYERS
 CAMBRIDGE UNIVERSITY PRESS , 2006
- [2]. *BASIC LINEAR ALGEBRA*
 T.S. BLYTH AND F.F ROBERTSON
- [3]. *ΣΤΟΙΧΕΙΩΔΗΣ ΔΙΑΦΟΡΙΚΕΣ ΕΙΣΩΣΕΙΣ ΚΑΙ ΠΡΟΒΛΗΜΑΤΑ ΣΥΝΟΡΙΑΚΩΝ ΤΙΜΩΝ*
 WILLIAM E. BOYCE AND RICHARD C. DIPRIMA
- [4]. *ΓΡΑΜΜΙΚΗ ΑΛΓΕΒΡΑ*
 ΙΩΑΝΝΗΣ Β. ΜΑΡΟΥΛΑΣ
- [5]. *THE ALGEBRAIC EIGENVALUE PROBLEM*
 J.H. WILKINSON
 CLARENDON PRESS , OXFORD UNIVERSITY PRESS , NEW YORK ,1988
- [6]. *THE SYMMETRIC EIGENVALUE PROBLEM*
 B. PARLETT
 PRENTICE-HALL , ENGLEWOOD CLIFFS , NJ ,1980
- [7]. *MATRIX COMPUTATIONS*
 G.H GOLUB AND C.F. VAN LOAN
 (THIRD EDITION , JOHN HOPKINS UNIVERSITY PRESS , BALTIMORE , 1996)
- [8]. *ΓΡΑΜΜΙΚΗ ΑΛΓΕΒΡΑ ΑΝΑΛΥΤΙΚΗ ΓΕΩΜΕΤΡΙΑ ΚΑΙ ΕΦΑΡΜΟΓΕΣ*
 Ν. ΚΑΔΙΑΝΑΚΗΣ ΚΑΙ Σ. ΚΑΡΑΝΑΣΙΟΣ
 ΕΜΠ ΑΘΗΝΑ 2003 ΕΚΔΟΣΗ 3^η
- [9]. *ΑΡΙΘΜΗΤΙΚΗ ΑΝΑΛΥΣΗ*
 Γ.Σ. ΠΑΠΑΓΕΩΡΓΙΟΥ ΚΑΙ Χ.Γ. ΤΣΙΤΟΥΡΑΣ
 ΕΚΔΟΣΕΙΣ ΣΥΜΕΩΝ , ΑΘΗΝΑ 2004
- [10]. *ΣΥΓΧΡΟΝΟ ΜΑΘΗΜΑΤΙΚΟ ΛΟΓΙΣΜΙΚΟ: MATLAB-MATHEMATICA*
 Γ.Σ ΠΑΠΑΓΕΩΡΓΙΟΥ ,Χ.Γ. ΤΣΙΤΟΥΡΑΣ ΚΑΙ Ι.Θ. ΦΑΜΕΛΗΣ
 ΕΚΔΟΣΕΙΣ ΣΥΜΕΩΝ , ΑΘΗΝΑ 2004
- [11]. *NUMERICAL MATHEMATICS*
 A. QUARTERONI , R. SACCO AND F. SALERI
 SPRINGER - VERLAY , 2007

