Εθνικό Μετσόβιο Πολυτεχνείο

Σχολή Εφαρμοσμένων Μαθηματικών και Φυσικών Επιστημών

ΑΡΙΘΜΗΤΙΚΕΣ ΜΕΘΟΔΟΙ ΓΙΑ ΤΗΝ ΕΥΡΕΣΗ ΙΔΙΟΤΙΜΩΝ ΠΙΝΑΚΑ

Διπλωματική Εργασία

ΚΑΡΑΝΤΖΙΑ ΑΝΝΑ

Επιβλέπων Καθηγητής: Παναγιώτης Ψαρράκος

ΠΕΡΙΕΧΟΜΕΝΑ

ΠΡΟΛΟΓΟΣ

Στην εργασία αυτή θα ασχοληθούμε με τη μελέτη μεθόδων της Αριθμητικής Γραμμικής Άλγεβρας που σκοπό έχουν να υπολογίσουν τις ιδιοτιμές και τα ιδιοδιανύσματα τυχαίων πινάκων.

Συγκεκριμένα οι μέθοδοι τις οποίες θα μελετήσουμε είναι οι εξής:

- Μέθοδος των Δυνάμεων
- Μέθοδος της Αντίστροφης Επανάληψης
- Μέθοδος του Πηλίκου Rayleigh
- Απλός Αλγόριθμος QR
- "Πρακτικός" Αλγόριθμος QR
- Μέθοδος Ταυτοχρόνων Επαναλήψεων
- Μέθοδος Jacobi
- Μέθοδος Διχοτόμησης
- Μέθοδος Διαίρει και Βασίλευε

Στο πρώτο Κεφάλαιο της εργασίας γίνεται αναφορά σε βασικούς ορισμούς και θεωρήματα της Γραμμικής Άλγεβρας, τα οποία είναι απαραίτητα για τη μελέτη και κατανόηση των επόμενων κεφαλαίων.

Το δεύτερο Κεφάλαιο αναφέρεται στις δυσκολίες που μπορεί κάποιος να συναντήσει όταν προσπαθεί να υπολογίσει με προφανείς τρόπους τις ιδιοτιμές και τα ιδιοδιανύσματα ενός πίνακα. Οι προφανείς αυτοί τρόποι είναι ο υπολογισμός της ιδιοτιμής ως ρίζα του χαρακτηριστικού πολυωνύμου και ο υπολογισμός της ως όριο μιας ακολουθίας. Στο κεφάλαιο αυτό αναφέρεται ακόμη η παραγοντοποίηση Schur, η οποία χρησιμοποιείται από τους περισσότερους γενικού τύπου αλγορίθμους, καθώς επίσης και η 2-φασική προσέγγιση ιδιοτιμών.

Το τρίτο Κεφάλαιο αναφέρεται στην απλοποίηση ενός πίνακα σε μορφή Hessenberg μέσω μιας ακολουθίας ορθοκανονικών μετασχηματισμών Householder, με σκοπό την εισαγωγή μηδενικών κάτω από την πρώτη υποδιαγώνιο. Επίσης θα μελετήσουμε την περίπτωση που ο πίνακας αυτός είναι ερμιτιανός και που όπως θα δούμε μετατρέπεται σε τριδιαγώνια μορφή.

Στο τέλος του κεφαλαίου αυτού γίνεται μία αναφορά στο κόστος σε πράξεις και στην ευστάθεια της απλοποίησης σε μορφή Hessenberg.

Στο τέταρτο Κεφάλαιο παρουσιάζονται μερικοί από τους αλγορίθμους εύρεσης ιδιοτιμών, όπου για την απλοποίηση της θεωρητικής παρουσίασής τους (και των αποδείξεων) θα περιοριστούμε σε πραγματικούς και συμμετρικούς πίνακες, κάτι που εξασφαλίζει πραγματικές ιδιοτιμές και ορθοκανονικά συστήματα ιδιοδιανυσμάτων. Στο κεφάλαιο αυτό λοιπόν θα μελετήσουμε τη Μέθοδο των Δυνάμεων, η οποία βρίσκει μόνο τη κατά μέτρο μεγαλύτερη ιδιοτιμή ενός πίνακα και το αντίστοιχο ιδιοδιάνυσμα. Επίσης θα περιγράψουμε τη Μέθοδο της Αντίστροφης Επανάληψης και την επαναληπτική Μέθοδο του Πηλίκου Rayleigh οι οποίες μπορούν να υπολογίσουν και τις υπόλοιπες ιδιοτιμές του πίνακα.

Στο πέμπτο Κεφάλαιο θα ασχοληθούμε με αριθμητικές μεθόδους οι οποίες δίνουν περισσότερες ιδιοτιμές. Αυτές οι μέθοδοι είναι ο Απλός Αλγόριθμος QR, αλλά και ο " Πρακτικός " Αλγόριθμος QR, ο οποίος προκύπτει από τη τροποποίηση του πρώτου με τη εισαγωγή εναλλαγών. Επίσης θα αναφερθούμε στη Μέθοδο των Ταυτοχρόνων Επαναλήψεων την οποία θα συσχετίσουμε με τον Απλό Αλγόριθμο QR.

Στο έκτο Κεφάλαιο γίνεται μια σύνδεση του Αλγορίθμου QR με εναλλαγές με τις Μεθόδους της Αντίστροφης Επανάληψης, της Εναλλασσόμενης Αντίστροφης Επανάληψης και της Επαναληπτικής Μεθόδου του Πηλίκου Rayleigh. Ακόμη γίνεται αναφορά στις εναλλαγές Wilkinson καθώς και στην ακρίβεια και την ευστάθεια του Αλγορίθμου QR.

Στο έβδομο Κεφάλαιο ολοκληρώνουμε τη θεωρητική μελέτη της εργασίας με μια αναφορά σε τρείς αλγόριθμους για πραγματικά, συμμετρικά προβλήματα ιδιοτιμών. Αυτοί οι αλγόριθμοι είναι ο Αλγόριθμος Jacobi για τους κανονικούς πίνακες, η Διχοτόμηση και οι αλγόριθμοι Διαίρει και Βασίλευε για τους τριδιαγώνιους πίνακες.

Η εργασία ολοκληρώνεται με το όγδοο Κεφάλαιο στο οποίο καταγράφονται κάποια αριθμητικά παραδείγματα τα οποία πραγματοποιήθηκαν με χρήση της Matlab. Τα παραδείγματα αυτά αφορούν τις μεθόδους που περιγράψαμε στο τέταρτο και πέμπτο Κεφάλαιο. Σημειώνεται ότι οι πίνακες που χρησιμοποιούμε στα παραδείγματά μας είναι πραγματικοί, συμμετρικοί αλλά και μη συμμετρικοί, ώστε να διαφανεί η γενικότερη εφαρμογή των μεθόδων.

Στο σημείο αυτό θα ήθελα να ευχαριστήσω θερμά τον επιβλέποντα καθηγητή της διπλωματικής μου, κύριο Παναγιώτη Ψαρράκο, Αναπληρωτή Καθηγητή του Εθνικού Μετσόβιου Πολυτεχνείου, για την καθοδήγηση, τη συνεργασία και τη πολύτιμη βοήθεια του καθ’ όλη τη διάρκεια εκπόνησης της παρούσας εργασίας. Θα ήθελα επίσης να ευχαριστήσω τους γονείς μου για τη διαρκή υποστήριξη που μου παρείχαν έως τώρα.

ΚΕΦΑΛΑΙΟ 1

Προβλήματα Ιδιοτιμών

Το πρόβλημα της εύρεσης των ιδιοτιμών ενός πίνακα είναι ένα ιδιαίτερα ενδιαφέρον πρόβλημα στο πεδίο των επιστημονικών υπολογισμών. Ο λόγος είναι ότι οι καλύτεροι αλγόριθμοι για την εύρεση ιδιοτιμών είναι ισχυροί, κάτι το οποίο θα συζητήσουμε αργότερα. Στο κεφάλαιο αυτό λοιπόν θα αναφερθούμε στα θεωρήματα και τους ορισμούς που αφορούν τις ιδιοτιμές και τα ιδιοδιανύσματα ενός πίνακα, στοιχεία που είναι απαραίτητα για την μελέτη των ακόλουθων κεφαλαίων.

1.1. Ιδιοτιμές και Ιδιοδιανύσματα

Έστω $\mathrm{A} \in \mathbb{C}^{m\times m}$ ένας τετραγωνικός πίνακας. Ένα μη μηδενικό διάνυσμα $x \in \mathbb{C}^m$ καλείται ιδιοδιάνυσμα του A και το $\lambda \in \mathbb{C}$, ιδιοτιμή που αντιστοιχεί στο διάνυσμα αυτό, αν

$$\mathrm{A}x=\lambda x. \tag{1.1}$$

Παρατηρώντας λοιπόν την παραπάνω σχέση, βλέπουμε ότι η δράση ενός πίνακα Α πάνω σε ένα υπόχωρο S του $\mathbb{C}^m$ λειτουργεί μερικές φορές ως βαθμωτό γινόμενο λx. Όταν συμβαίνει αυτό, ο υπόχωρος S, που αποτελείται από τα διανύσματα $x \in \mathbb{C}^m$ τέτοια ώστε $\mathrm{A}x=\lambda x$, ονομάζεται ιδιοχώρος του πίνακα Α που αντιστοιχεί στην ιδιοτιμή λ και κάθε μη μηδενικό διάνυσμα $x \in S$ είναι ιδιοδιάνυσμα του Α που αντιστοιχεί στην ιδιοτιμή λ.

Το σύνολο όλων των ιδιοτιμών ενός πίνακα Α ονομάζεται φάσμα του Α και είναι ένα υποσύνολο του $\mathbb{C}$ το οποίο συμβολίζεται ως Λ(Α).

Τα προβλήματα ιδιοτιμών έχουν ένα πολύ διαφορετικό χαρακτήρα από τα προβλήματα που περιλαμβάνουν τετραγωνικά ή ορθογώνια γραμμικά συστήματα εξισώσεων. Για ένα σύστημα εξισώσεων, το πεδίο ορισμού μπορεί να είναι ένας χώρος και το σύνολο τιμών μπορεί να είναι ένας άλλος. Στο παράδειγμα που ακολουθεί ο πίνακας Α αντιστοιχεί διανύσματα μήκους n με πολυωνυμικούς συντελεστές σε διανύσματα μήκους m απλών πολυωνύμων.

Παράδειγμα 1.1. Πίνακας Vardermonde.

Φτιάχνουμε μια ακολουθία αριθμών $\{ x_1,x_2,\ldots,x_m \}$. Αν p και q είναι πολυώνυμα βαθμού μικρότερου του n και a είναι ένας αριθμός, τότε τα $p+q$ και ap είναι επίσης πολυώνυμα βαθμού μικρότερου του n. Επιπλέον οι τιμές αυτών των πολυωνύμων στα σημεία x_i ικανοποιούν τις ακόλουθες ιδιότητες γραμμικότητας:

$$(p+q)(x_i)=p(x_i)+q(x_i),$$

$$(ap)(x_i) = a(p(x_i)).$$

Έτσι η απεικόνιση από διανύσματα με συντελεστές πολυώνυμα p βαθμού μικρότερου του n σε διανύσματα $\left(p(x_1), p(x_2), \dots, p(x_m)\right)$ απλών πολυωνυμικών τιμών είναι γραμμική. Κάθε γραμμική απεικόνιση μπορεί να εκφραστεί μέσω ενός γινομένου πινάκων. Στην πραγματικότητα εκφράζεται μέσω ενός $m \times n$ Vandermonde πίνακα.

$$\mathrm{A}=\begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^{n-1} \end{bmatrix}.$$

Αν c είναι το διάνυσμα στήλη των συντελεστών του p

$$\mathrm{c} = \begin{bmatrix} c_0 \\ c_1 \\ c_2 \\ \vdots \\ c_{n-1} \end{bmatrix}, \qquad \mathrm{p}(x) = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1} x^{n-1},$$

τότε το γινόμενο Ac δίνει τις απλές πολυωνυμικές τιμές. Αυτό σημαίνει ότι, για κάθε i από το 1 μέχρι το m, έχουμε

$$(\mathrm{A}c)_i = c_0 + c_1 x_i + c_2 {x_i}^2 + \cdots + c_{n-1} {x_i}^{n-1} = p(x_i).$$

Σ' αυτό το παράδειγμα είναι ξεκάθαρο ότι δεν χρειάζεται να σκεφτόμαστε το γινόμενο πίνακας – διάνυσμα, $\mathrm{A}c$, ως m διακριτά βαθμωτά αθροίσματα το καθένα από τα οποία δίνει ένα διαφορετικό γραμμικό συνδυασμό των συνιστωσών του c. Αντιθέτως μπορούμε να θεωρήσουμε τον πίνακα A ως ένα πίνακα στηλών, όπου κάθε μία στήλη δίνει απλές τιμές ενός μονωνύμου,

$$\mathrm{A} = \left[1 \,\middle|\, \quad x \,\middle|\, \quad x^2 \,\middle|\, \quad \dots \quad x^{n-1}\right],$$

και το γινόμενο $\mathrm{A}c$ θα πρέπει να γίνεται αντιληπτό ως ένα μόνο διάνυσμα άθροισης που δίνει απ' ευθείας ένα γραμμικό συνδυασμό αυτών των μονωνύμων.

$$Ac = c_0 + c_1 x + c_2 x^2 + \cdots + c_{n-1} x^{n-1} = p(x). \qquad \square$$

Το να αναζητήσουμε τις ιδιοτιμές ενός τέτοιου πίνακα A δεν θα είχε νόημα. Τα προβλήματα ιδιοτιμών έχουν νόημα μόνο όταν οι χώροι του συνόλου τιμών και του πεδίου ορισμού είναι οι ίδιοι. Αυτό αντανακλά το γεγονός ότι στις εφαρμογές, οι ιδιοτιμές χρησιμοποιούνται γενικά όταν πρόκειται να συνθέσουμε ένα πίνακα αναδρομικά, είτε ως δύναμη A^k είτε σε συναρτησιακή μορφή όπως $e^{t\mathrm{A}}$.

Μιλώντας γενικά οι ιδιοτιμές και τα ιδιοδιανύσματα είναι χρήσιμα για δύο λόγους, έναν αλγοριθμικό και έναν άλλο φυσικό. Αλγοριθμικά, η ανάλυση ιδιοτιμών μπορεί να απλοποιήσει τις λύσεις ορισμένων προβλημάτων μετατρέποντας ένα περίπλοκο σύστημα σε μια συλλογή από βαθμωτά προβλήματα. Για το λόγο που αφορά τη φυσική, η ανάλυση ιδιοτιμών μπορεί να συμβάλλει στην κατανόηση της συμπεριφοράς των συστημάτων που διέπονται από γραμμικές εξισώσεις.

1.2. Διαγωνοποίηση Πίνακα

Η Διαγωνοποίηση ενός τετραγωνικού πίνακα A είναι μια παραγοντοποίηση του (αν αυτή είναι δυνατή) σε

$$\mathrm{A} = \mathrm{X}\Lambda\mathrm{X}^{-1}. \tag{1.2}$$

Εδώ ο X είναι αντιστρέψιμος πίνακας και ο Λ είναι διαγώνιος. Αυτός ο ορισμός μπορεί επίσης να γραφτεί ως εξής:

$$\mathrm{AX} = \mathrm{X}\Lambda\,, \tag{1.3}$$

δηλαδή

$$[A\,]\big[x_1\,\big|\quad x_2\,\big|\quad \dots\quad \big|\,x_m\big] = \big[x_1\,\big|\quad x_2\,\big|\quad \dots\quad \big|\,x_m\big]\begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix}.$$

Αυτό κάνει ξεκάθαρο ότι αν x_j είναι η $j\,-$οστή στήλη του X και λ_j είναι το $j\,-$οστό στοιχείο του Λ τότε $\mathrm{A}x_j = \lambda_j x_j$. Έτσι η $j\,-$οστή στήλη του X είναι ένα ιδιοδιάνυσμα του A και το $j\,-$οστό διαγώνιο στοιχείο του Λ είναι η αντίστοιχη ιδιοτιμή του.

Η διαγωνοποίηση πίνακα εκφράζει μια αλλαγή βάσης στις "συντεταγμένες του ιδιοδιανύσματος". Αν $\mathrm{A}x = b$ και $\mathrm{A} = \mathrm{X}\Lambda\mathrm{X}^{-1}$ έχουμε

$$(\mathrm{X}^{-1}b) = \Lambda(\mathrm{X}^{-1}x). \tag{1.4}$$

Έτσι, για να υπολογίσουμε το $\mathrm{A}x$, μπορούμε να επεκτείνουμε το x με βάση της στήλες του X, να εφαρμόσουμε τον Λ και να ερμηνεύσουμε το αποτέλεσμα σαν ένα διάνυσμα συντελεστών ενός γραμμικού συνδυασμού των στηλών του X.

1.3. Γεωμετρική Πολλαπλότητα

Όπως αναφέρθηκε ανωτέρω, το σύνολο των ιδιοδιανυσμάτων που αντιστοιχούν σε μία και μόνο ιδιοτιμή μαζί με το μηδενικό διάνυσμα, συνθέτουν έναν υπόχωρο του $\mathbb{C}^m$ γνωστό ως ιδιοχώρο. Αν λ είναι μια ιδιοτιμή του A, τότε συμβολίζουμε τον αντίστοιχο ιδιοχώρο με E_λ. Ένας ιδιοχώρος E_λ, είναι ένα παράδειγμα αναλλοίωτου υπόχωρου του A, δηλαδή $\mathrm{AE}_\lambda \subseteq \mathrm{E}_\lambda$.

Η διάσταση του E_λ μπορεί να ερμηνευτεί ως ο μέγιστος αριθμός γραμμικών ανεξαρτήτων ιδιοδιανυσμάτων που μπορούν να βρεθούν και που όλα αντιστοιχούν στην ίδια ιδιοτιμή λ. Αυτός ο αριθμός είναι γνωστός ως η γεωμετρική πολλαπλότητα του λ. Η γεωμετρική πολλαπλότητα μπορεί επίσης να περιγραφεί ως η διάσταση του μηδενοχώρου του $\mathrm{A} - \lambda\mathrm{I}$ από τη στιγμή που αυτός ο μηδενοχώρος είναι ο ίδιος ο E_λ.

1.4. Χαρακτηριστικό Πολυώνυμο

Το χαρακτηριστικό πολυώνυμο του $\mathrm{A} \in \mathbb{C}^{m\times m}$ το οποίο συμβολίζεται με p_A ή απλά με p, είναι το πολυώνυμο βαθμού m που ορίζεται ως :

$$p_\mathrm{A}(z) = det(z\mathrm{I} - \mathrm{A}). \tag{1.5}$$

Λόγω της θέσης του πρόσημου (μείον), το p είναι μονικό που σημαίνει ότι ο συντελεστής του μεγιστοβάθμιου όρου βαθμού m είναι 1.

Θεώρημα 1.1. Το λ είναι μία ιδιοτιμή του A αν και μόνο αν $p_\mathrm{A}(\lambda) = 0$.

Απόδειξη

Αυτό προκύπτει από τον ορισμό της ιδιοτιμής:

Το λ είναι ιδιοτιμή του A $\Leftrightarrow$ Υπάρχει ένα μη μηδενικό διάνυσμα x τέτοιο ώστε $(\lambda I - \mathrm{A}\,)x = 0$

$\Leftrightarrow$ ο πίνακας $\lambda I - \mathrm{A}$ είναι μη αντιστρέψιμος, πράγμα που ισχύει

$\Leftrightarrow$ $\det(\lambda I - \mathrm{A}) = 0$. □

Το Θεώρημα 1.1 έχει μια σημαντική συνέπεια: *Ακόμη και αν ένας πίνακας είναι πραγματικός κάποιες από τις ιδιοτιμές του μπορεί να είναι μιγαδικές*. Στη φυσική αυτό σχετίζεται με το φαινόμενο στο οποίο, πραγματικά δυναμικά συστήματα μπορούν να έχουν κινήσεις που ταλαντεύονται καθώς αυξάνονται ή μειώνονται. Αλγοριθμικά αυτό σημαίνει ότι ακόμη και αν τα δεδομένα εισόδου σε ένα πίνακα προβλήματος ιδιοτιμών είναι πραγματικά, τα δεδομένα εξόδου μπορεί να είναι μιγαδικά.

1.5. Αλγεβρική Πολλαπλότητα

Από το θεμελιώδες θεώρημα της Άλγεβρας, μπορούμε να γράψουμε το

$$p_\mathrm{A}(z) = (z - \lambda_1)(z - \lambda_2) \ldots (z - \lambda_m), \tag{1.6}$$

όπου $\lambda_j \in \mathbb{C}$ οι ρίζες του πολυωνύμου. Σύμφωνα με το Θεώρημα 1.1, κάθε λ_j είναι μια ιδιοτιμή του A. Γενικά μια ιδιοτιμή μπορεί να εμφανιστεί περισσότερες από μια φορές. Ορίζουμε την αλγεβρική πολλαπλότητα μιας Ιδιοτιμής λ του A να είναι η πολλαπλότητα με την οποία εμφανίζεται ως ρίζα του πολυωνύμου p_{A}. Μια ιδιοτιμή είναι απλή αν η αλγεβρική της πολλαπλότητα είναι 1.

Το χαρακτηριστικό πολυώνυμο μας δίνει έναν εύκολο τρόπο να υπολογίσουμε το πλήθος των ιδιοτιμών ενός πίνακα.

Θεώρημα 1.2. Αν $\mathrm{A} \in \mathbb{C}^{m\times m}$, τότε ο A έχει m ιδιοτιμές, μετρημένες με τις αλγεβρικές πολλαπλότητες τους. Ιδίως αν οι ρίζες του p_{A} είναι απλές, τότε ο A έχει m διακριτές ιδιοτιμές.

Σημειώνουμε ότι κάθε πίνακας έχει το λιγότερο μια ιδιοτιμή.

Η αλγεβρική πολλαπλότητα μιας ιδιοτιμής είναι μεγαλύτερη ή ίση από τη γεωμετρική της πολλαπλότητα. Για να το αποδείξουμε αυτό, χρειάζεται να γνωρίζουμε κάτι σχετικό με τους μετασχηματισμούς ομοιότητας.

1.6. Μετασχηματισμοί Ομοιότητας

Αν $X \in \mathbb{C}^{m\times m}$ είναι αντιστρέψιμος, τότε η απεικόνιση $A \longmapsto X^{-1}AX$ καλείται μετασχηματισμός ομοιότητας του A. Λέμε ότι δύο πίνακες A και B είναι όμοιοι αν υπάρχει ένας μετασχηματισμός ομοιότητας που να σχετίζει τον έναν με τον άλλο. Δηλαδή αν υπάρχει ένας αντιστρέψιμος πίνακας $X \in \mathbb{C}^{m\times m}$ τέτοιος ώστε $B = X^{-1}AX$. Όπως περιγράφηκε παραπάνω στην ειδική περίπτωση της διαγωνοποίησης (1.2) κάθε μετασχηματισμός ομοιότητας είναι μια διαδικασία αλλαγής βάσης.

Οι όμοιοι πίνακες A και $X^{-1}AX$ έχουν πολλές κοινές ιδιότητες.

Θεώρημα 1.3. Αν X είναι αντιστρέψιμος, τότε οι A και $X^{-1}AX$ έχουν το ίδιο χαρακτηριστικό πολυώνυμο, τις ίδιες ιδιοτιμές και τις ίδιες αλγεβρικές και γεωμετρικές πολλαπλότητες.

Απόδειξη

Η απόδειξη του ότι τα χαρακτηριστικά πολυώνυμα είναι τα ίδια είναι ο ευθύς υπολογισμός:

$$p_{X^{-1}AX}(z) = det(zI - X^{-1}AX) = det(X^{-1}(zI - A)X)$$

$$= det(X^{-1})det(zI - A)det(X) = det(zI - A) = p_{\mathrm{A}}(z).$$

Εφόσον λοιπόν τα χαρακτηριστικά πολυώνυμα είναι τα ίδια, έπεται ότι οι ιδιοτιμές που αποτελούν τις ρίζες του χαρακτηριστικού πολυωνύμου θα είναι ίσες, αλλά και οι αλγεβρικές πολλαπλότητες είναι ίσες ως οι πολλαπλότητες των ριζών των πολυωνύμων.

Τέλος η απόδειξη του ότι οι γεωμετρικές πολλαπλότητες συμπίπτουν είναι η εξής: Έστω ιδιοτιμή λ του Α με γεωμετρική πολλαπλότητα k και $\{x_1, x_2, \dots, x_k\}$ μια βάση του ιδιοχώρου E_λ . Τότε για κάθε $i = 1,2,\dots,k$, ισχύει $Ax_i = \lambda x_i \Leftrightarrow (X^{-1}AX)(X^{-1}x_i) = \lambda(X^{-1}x_i)$. Άρα $\{X^{-1}x_1, X^{-1}x_2, \dots, X^{-1}x_k\}$ είναι γραμμικώς ανεξάρτητα ιδιοδιανύσματα του λ ως ιδιοτιμή του $X^{-1}AX$. Δηλαδή η γεωμετρική πολλαπλότητα της λ ως ιδιοτιμής του A είναι μικρότερη ή ίση της γεωμετρικής πολλαπλότητας της λ ως ιδιοτιμής του $X^{-1}AX$. Ομοίως αποδεικνύεται και η αντίστροφη ανίσωση, επομένως οι δύο γεωμετρικές πολλαπλότητες είναι ίσες. □

Τώρα μπορούμε να συσχετίσουμε την γεωμετρική πολλαπλότητα με την αλγεβρική.

Θεώρημα 1.4. Η αλγεβρική πολλαπλότητα μιας ιδιοτιμής λ είναι μεγαλύτερη ή ίση της γεωμετρικής πολλαπλότητας.

Απόδειξη

Έστω n η γεωμετρική πολλαπλότητα της ιδιοτιμής λ ενός πίνακα Α. Σχηματίζουμε ένα $m \times n$ πίνακα $\hat{V}$ του οποίου οι n στήλες σχηματίζουν ορθοκανονική βάση του ιδιοχώρου $\{x: Ax = \lambda x\}$. Τότε επεκτείνοντας τον $\hat{V}$ σε ένα τετραγωνικό ορθομοναδιαίο πίνακα V έχουμε ότι ο V^*AV είναι της μορφής.

$$B = V^*AV = \begin{bmatrix} \lambda I & C \\ 0 & D \end{bmatrix}, \tag{1.7}$$

όπου I είναι ο $n \times n$ ταυτοτικός πίνακας, C είναι ένας πίνακας $n \times (m-n)$ και D είναι ένας $(m-n) \times (m-n)$ πίνακας. Από τον ορισμό της ορίζουσας ενός πίνακα προκύπτει $det(zI - B) = det(zI - \lambda I)det(zI - D) = (z-\lambda)^n det(zI - D)$. Άρα η αλγεβρική πολλαπλότητα μιας ιδιοτιμής λ του B είναι τουλάχιστον n. Αφού όμως οι μετασχηματισμοί ομοιότητας διατηρούν τις πολλαπλότητες το ίδιο ισχύει και για τον A. □

1.7. Ιδιόμορφες Ιδιοτιμές και Ιδιόμορφοι Πίνακες.

Παρόλο που ένας τυχαίος πίνακας έχει σχεδόν σίγουρα ίση αλγεβρική και γεωμετρική πολλαπλότητα, αυτό δεν ισχύει για κάθε πίνακα.

Παράδειγμα 1.2. Θεωρούμε τους πίνακες

$$A = \begin{bmatrix} 2 & & \\ & 2 & \\ & & 2 \end{bmatrix}, \qquad B = \begin{bmatrix} 2 & 1 & \\ & 2 & 1 \\ & & 2 \end{bmatrix}.$$

Και ο A και ο B έχουν χαρακτηριστικό πολυώνυμο $(z-2)^3$, έτσι υπάρχει μια μόνο ιδιοτιμή $\lambda = 2$, αλγεβρικής πολλαπλότητας 3. Στην περίπτωση του A μπορούμε να βρούμε 3 γραμμικώς ανεξάρτητα ιδιοδιανύσματα, για παράδειγμα τα e_1, e_2 και e_3 , οπότε η γεωμετρική πολλαπλότητα είναι επίσης 3. Για τον B από την άλλη μεριά μπορούμε να βρούμε μόνο ένα γραμμικά ανεξάρτητο ιδιοδιάνυσμα (ένα πολλαπλάσιο του e_1), έτσι η γεωμετρική πολλαπλότητα της ιδιοτιμής είναι μόνο 1.

Μια ιδιοτιμή της οποίας η αλγεβρική πολλαπλότητα είναι μεγαλύτερη από την γεωμετρική πολλαπλότητα ονομάζεται ιδιόμορφη ιδιοτιμή. Ένας πίνακας ο οποίος έχει μία ή περισσότερες ιδιόμορφες ιδιοτιμές είναι ένας ιδιόμορφος πίνακας.

Κάθε διαγώνιος πίνακας είναι μη ιδιόμορφος. Για έναν τέτοιο πίνακα η αλγεβρική και η γεωμετρική πολλαπλότητα μιας ιδιοτιμής λ είναι ίσες με το πλήθος των εμφανίσεων της ιδιοτιμής στη διαγώνιο.

1.8. Διαγωνοποίηση

Η κλάση των μη ιδιόμορφων πινάκων είναι ακριβώς η κλάση των πινάκων που διαγωνοποιούνται (1.2).

Θεώρημα 1.5. Ένας $m \times m$ πίνακας A είναι μη ιδιόμορφος αν και μόνο αν διαγωνοποιείται $A = X\Lambda X^{-1}$.

Απόδειξη

($\Leftarrow$) Δεδομένης της διαγωνοποίησης $A = X\Lambda X^{-1}$, γνωρίζουμε από το Θεώρημα 1.3 ότι ο Λ είναι όμοιος με τον A, με τις ίδιες ιδιοτιμές και τις ίδιες πολλαπλότητες. Αφού ο Λ είναι ένας διαγώνιος πίνακας, είναι μη ιδιόμορφος και άρα το ίδιο ισχύει και για τον A .

($\Rightarrow$) Ένας μη ιδιόμορφος πίνακας πρέπει να έχει m γραμμικώς ανεξάρτητα ιδιοδιανύσματα, επειδή τα ιδιοδιανύσματα με διαφορετικές ιδιοτιμές πρέπει να είναι γραμμικώς ανεξάρτητα και κάθε ιδιοτιμή μπορεί να έχει τόσα γραμμικώς ανεξάρτητα ιδιοδιανύσματα όσα η πολλαπλότητα της. Αν αυτά τα m ανεξάρτητα διανύσματα τεθούν ως στήλες ενός πίνακα X, τότε ο X είναι αντιστρέψιμος και έχουμε $A = X\Lambda X^{-1}$. □

Παρατηρώντας το προηγούμενο αποτέλεσμα, διαπιστώνουμε ότι ένας άλλος όρος που μπορεί να χρησιμοποιηθεί για τον μη ιδιόμορφο πίνακα, είναι διαγωνοποιήσιμος. Κατόπιν τίθεται το ερώτημα για το κατά πόσο ένας διαγωνοποιήσιμος πίνακας A μπορεί κατά κάποιον τρόπο "να συμπεριφέρεται όπως" ο όμοιος του διαγώνιος Λ. Η απάντηση βασίζεται πάνω στο ποιά πτυχή της συμπεριφοράς μετράμε και στο δείκτη κατάστασης του X, του πίνακα των ιδιοδιανυσμάτων. Αν ο X είναι κακής κατάστασης, τότε ένας μεγάλος αριθμός πληροφοριών μπορεί να απορροφηθεί κατά το πέρασμα από το A στο Λ.

1.9. Ορίζουσα και Ίχνος

Το ίχνος του πίνακα $A \in \mathbb{C}^{m\times m}$ είναι το άθροισμα των διαγώνιων στοιχείων του: $tr(A) = \sum_{j=1}^{m} a_{jj}$. Και το ίχνος και η ορίζουσα συνδέονται απλά με τις ιδιοτιμές.

Θεώρημα 1.6. Η ορίζουσα $det(A)$ και το ίχνος $tr(A)$ είναι ίσα με το γινόμενο και το άθροισμα των ιδιοτιμών του A αντίστοιχα, μετρημένων με τις αλγεβρικές τους πολλαπλότητες.

$$det(A) = \prod_{j=1}^{m} \lambda_j , \qquad tr(A) = \sum_{j=1}^{m} \lambda_j . \tag{1.8}$$

Απόδειξη

Από την (1.5) και την (1.6) υπολογίζουμε

$$det(A) = (-1)^m det(-A) = (-1)^m p_A(0) = \prod_{J=1}^{m} \lambda_j.$$

Αυτό αποδεικνύει τον πρώτο τύπο. Όσο για τον δεύτερο από την (1.5) συνεπάγεται ότι ο συντελεστής του όρου z^{m-1} του p_A είναι το αντίθετο του αθροίσματος των διαγωνίων στοιχείων του A, δηλαδή το $-tr(A)$. Από την άλλη μεριά, από την (1.6) αυτός ο συντελεστής είναι επίσης ίσος με το $-\sum_{j=1}^{m} \lambda_j$. Έτσι $tr(A) = \sum_{j=1}^{m} \lambda_j$. □

1.10. Ορθοκανονική Διαγωνοποίηση

Μερικές φορές συμβαίνει ο $m \times m$ πίνακας A να έχει όχι μόνο m γραμμικώς ανεξάρτητα ιδιοδιανύσματα αλλά αυτά να μπορούν να επιλεγούν ώστε να είναι ορθογώνια. Σε αυτή την περίπτωση ο A είναι ορθοκανονικά διαγωνοποιήσιμος που σημαίνει ότι υπάρχει ένας ορθομοναδιαίος πίνακας Q τέτοιος ώστε

$$A = Q\Lambda Q^*. \tag{1.9}$$

Έχουμε ήδη δει μια κλάση πινάκων που είναι ορθοκανονικά διαγωνοποιήσιμοι, αυτοί είναι οι ερμιτιανοί πίνακες. Το ακόλουθο αποτέλεσμα προκύπτει από το Θεώρημα 1.9.

Θεώρημα 1.7. Ένας ερμιτιανός πίνακας είναι ορθοκανονικά διαγωνοποιήσιμος, και οι ιδιοτιμές του είναι πραγματικές.

Οι ερμιτιανοί πίνακες δεν είναι οι μοναδικοί πίνακες που είναι ορθοκανονικά διαγωνοποιήσιμοι. Άλλα παραδείγματα ορθοκανονικά διαγωνοποιήσιμων πινάκων, είναι οι αντιερμιτιανοί, οι ορθομοναδιαίοι, οι κυκλικοί και κάθε άλλος πίνακας που αποτελεί το άθροισμα ενός από τους παραπάνω πίνακες συν ενός πολλαπλασίου του ταυτοτικού πίνακα.

Γενικώς, η κλάση των πινάκων που είναι ορθοκανονικά διαγωνοποιήσιμοι έχει ένα κομψό χαρακτηρισμό. Εξ' ορισμού, λέμε ότι ένας πίνακας A είναι κανονικός αν $A^*A = AA^*$.

Το ακόλουθο αποτέλεσμα είναι πολύ γνωστό.

Θεώρημα 1.8. Ένας πίνακας A είναι ορθοκανονικά διαγωνοποιήσιμος αν και μόνο αν είναι κανονικός.

1.11. Κανονική Μορφή του Schur

Μια άλλη παραγοντοποίηση πινάκων η οποία είναι πολύ χρήσιμη στην αριθμητική ανάλυση επειδή όλοι οι πίνακες, συμπεριλαμβανομένου και των ιδιόμορφων πινάκων, μπορούν να παραγοντοποιηθούν χρησιμοποιώντας την, είναι η παραγοντοποίηση Schur. Μια παραγοντοποίηση Schur ενός πίνακα A είναι η παραγοντοποίηση :

$$A = QTQ^*, \tag{1.10}$$

όπου Q είναι ένας ορθομοναδιαίος πίνακας και T ένας άνω τριγωνικός. Σημειώνουμε ότι εφόσον οι πίνακες A και T είναι όμοιοι, οι ιδιοτιμές του A εμφανίζονται αναγκαστικά στη διαγώνιο του T.

Θεώρημα 1.9. : Κάθε τετραγωνικός πίνακας A επιδέχεται μια Schur παραγοντοποίηση.

Απόδειξη

Εφαρμόζουμε επαγωγή στη διάσταση m του πίνακα A. Η περίπτωση $m = 1$ είναι τετριμμένη, έτσι υποθέτουμε ότι $m \geq 2$. Έστω x ένα οποιοδήποτε ιδιοδιάνυσμα του A, με αντίστοιχη ιδιοτιμή λ. Παίρνουμε το x να είναι κανονικοποιημένο και έστω ότι είναι η πρώτη στήλη ενός ορθομοναδιαίου πίνακα U. Τότε, όπως στην (1.7) είναι εύκολο να δούμε ότι το γινόμενο U^*AU έχει τη μορφή:

$$U^*AU = \begin{bmatrix} \lambda & B \\ 0 & C \end{bmatrix}.$$

Από την επαγωγική υπόθεση υπάρχει μια Schur παραγοντοποίηση VTV^* του C.

Γράφουμε τώρα

$$Q = U \begin{bmatrix} 1 & 0 \\ 0 & V \end{bmatrix}.$$

Αυτός είναι ένας ορθομοναδιαίος πίνακας, και έχουμε

$$Q^*AQ = \begin{bmatrix} \lambda & BV \\ 0 & T \end{bmatrix}.$$

Αυτή είναι η παραγοντοποίηση Schur που ψάχνουμε. □

1.12. Παραγοντοποιήσεις Πινάκων για την Εύρεση Ιδιοτιμών

Στις προηγούμενες σελίδες περιγράψαμε τρία παραδείγματα παραγοντοποιήσεων ενός πίνακα οι οποίες τον μετατρέπουν σε μια άλλη μορφή, όπου οι ιδιοτιμές είναι εμφανείς. Μπορούμε να τις συνοψίσουμε ως εξής:

- Μία διαγωνοποίηση $A = X\Lambda X^{-1}$ υπάρχει αν και μόνο ο A είναι μη ιδιόμορφος.
- Μία ορθοκανονική διαγωνοποίηση $A = Q\Lambda Q^{*}$ υπάρχει αν και μόνο αν ο A είναι κανονικός.
- Μια ορθοκανονική τριγωνοποίηση (παραγοντοποίηση Schur) $A = QTQ^{*}$ πάντα ισχύει.

Για να υπολογίσουμε ιδιοτιμές, θα πρέπει λοιπόν να κατασκευάσουμε μία από αυτές τις παραγοντοποιήσεις. Συνήθως αυτή είναι η Schur παραγοντοποίηση, δεδομένου ότι εφαρμόζεται σε όλους τους πίνακες χωρίς περιορισμούς. Εξάλλου, αφού εμπλέκονται ορθοκανονικοί μετασχηματισμοί, οι αλγόριθμοι που προκύπτουν, τείνουν να είναι αριθμητικά σταθεροί. Αν A είναι κανονικός πίνακας, τότε η μορφή του A που προκύπτει από τη παραγοντοποίηση Schur είναι διαγώνια. Ειδικότερα, αν ο A είναι ερμιτιανός, τότε μπορούμε να εκμεταλλευτούμε το πλεονέκτημα αυτής της συμμετρίας κατά τη διάρκεια του υπολογισμού, και να μετατρέψουμε τον A σε διαγώνιο με τη μισή εργασία ή λιγότερη από αυτή που απαιτείται για τη μετατροπή ενός γενικού πίνακα.

ΚΕΦΑΛΑΙΟ 2

Γενική Περιγραφή Αλγορίθμων Εύρεσης Ιδιοτιμών

Το συγκεκριμένο κεφάλαιο καθώς και τα επόμενα πέντε που θα ακολουθήσουν περιγράφουν μερικούς από τους άμεσους αλγορίθμους για τον υπολογισμό ιδιοτιμών και ιδιοδιανυσμάτων, καθώς και μερικές σύγχρονες παραλλαγές τους. Οι περισσότεροι από αυτούς τους αλγορίθμους λειτουργούν σε δύο φάσεις. Η πρώτη φάση είναι μια προκαταρκτική μεταβολή από γενική σε συγκεκριμένα δομημένη μορφή και η δεύτερη μια αναδρομική διαδικασία για την τελική σύγκλιση. Αυτό το κεφάλαιο δίνει τη 2-φασική προσέγγιση και εξηγεί γιατί είναι επωφελής.

2.1. Μειονεκτήματα Άμεσων Αλγορίθμων

Παρόλο που οι ιδιοτιμές και τα ιδιοδιανύσματα έχουν απλούς ορισμούς και κομψούς χαρακτηρισμούς, οι καλύτεροι τρόποι για να τα υπολογίσουμε δεν είναι προφανείς.

Ίσως η πρώτη μέθοδος που θα μπορούσε κανείς να σκεφτεί είναι να υπολογίσει τους συντελεστές του χαρακτηριστικού πολυωνύμου και να χρησιμοποιήσει μεθόδους εύρεσης ριζών για να υπολογίσει τις ιδιοτιμές. Δυστυχώς, αυτή η ιδέα είναι κακή, επειδή οι αλγόριθμοι εύρεσης ριζών πολυωνύμων δημιουργούν γενικά, προβλήματα κακής κατάστασης, ακόμα και αν το υποκείμενο πρόβλημα ιδιοτιμών είναι καλά ορισμένο. (Στην πραγματικότητα οι αλγόριθμοι εύρεσης ριζών πολυωνύμων είναι χωρίς αμφιβολία ένα γενικό θέμα στην επιστήμη των υπολογιστών-ακριβώς επειδή είναι ο καλύτερος τρόπος για την επίλυση εφαρμοσμένων προβλημάτων.)

Μια απλή ιδέα θα ήταν να επωφεληθούμε από το γεγονός ότι η ακολουθία

$$\frac{x}{\|x\|}, \frac{Ax}{\|Ax\|}, \frac{A^2x}{\|A^2x\|}, \frac{A^3x}{\|A^3x\|}, \cdots$$

συγκλίνει υπό ορισμένες προϋποθέσεις σε ένα ιδιοδιάνυσμα που αντιστοιχεί στη μεγαλύτερη κατά μέτρο ιδιοτιμή του A. Αυτή η μέθοδος εύρεσης ιδιοδιανύσματος λέγεται «Μέθοδος Δυνάμεων». Δυστυχώς, παρόλο που η «Μέθοδος Δυνάμεων» είναι δημοφιλής, δεν αποτελεί ένα αποτελεσματικό εργαλείο για γενική χρήση. Εκτός από μερικούς ειδικούς πίνακες, η μέθοδος αυτή είναι πολύ αργή.

Οι καλύτεροι γενικού τύπου αλγόριθμοι ιδιοτιμών δεν βασίζονται στις ιδέες που προαναφέραμε, αλλά σε μια διαφορετική αρχή: τον υπολογισμό μιας παραγοντοποίησης του πίνακα A όπου οι ιδιοτιμές εμφανίζονται ως στοιχεία ενός από τους πίνακες της παραγοντοποίησης αυτής. Είδαμε τρείς τέτοιες παραγοντοποιήσεις εύρεσης ιδιοτιμών στο προηγούμενο κεφάλαιο: τη διαγωνοποίηση, την ορθοκανονική διαγωνοποίηση και την ορθοκανονική τριγωνοποίηση (Schur παραγοντοποίηση). Στην πράξη, οι ιδιοτιμές υπολογίζονται συνήθως, κατασκευάζοντας μια από αυτές τις παραγοντοποιήσεις. Για να το καταφέρουμε αυτό εφαρμόζουμε μια σειρά από μετασχηματισμούς ομοιότητας του A οι οποίοι εισάγουν μηδενικά στις κατάλληλες θέσεις. Έτσι, βλέπουμε ότι η εύρεση ιδιοτιμών καταλήγει μάλλον με τρόπο παρόμοιο με την επίλυση συστημάτων εξισώσεων ή τουλάχιστον τετραγωνικών προβλημάτων. Οι αλγόριθμοι της αριθμητικής γραμμικής άλγεβρας βασίζονται κυρίως σε μια τεχνική που χρησιμοποιείται ξανά και ξανά: Τοποθέτηση μηδενικών μέσα στους πίνακες.

2.2.Μια Θεμελιώδης Δυσκολία

Για να δούμε τη δυσκολία, σημειώνουμε ότι όπως τα προβλήματα ιδιοτιμών μπορούν να μετατραπούν σε προβλήματα εύρεσης ριζών πολυωνύμων, έτσι και αντίστροφα, κάθε πρόβλημα εύρεσης ριζών πολυωνύμων μπορεί να θεωρηθεί ως ένα πρόβλημα ιδιοτιμών. Υποθέτουμε ότι έχουμε το μονικό πολυώνυμο.

$$p(z) = z^m + a_{m-1}z^{m-1} + \cdots + a_1 z + a_0. \tag{2.1}$$

Με την επέκταση σε υποπίνακες, δεν είναι δύσκολο να εξακριβωθεί ότι το $p(z)$ είναι ίσο με $(-1)^m$ φορές την ορίζουσα του $m \times m$ πίνακα.

$$\begin{bmatrix} -z & & & & & -a_0 \\ 1 & -z & & & & -a_1 \\ & 1 & -z & & & -a_2 \\ & & 1 & \ddots & & \vdots \\ & & & \ddots & -z & -a_{m-2} \\ & & & & 1 & (-z - a_{m-1}) \end{bmatrix}. \tag{2.2}$$

Αυτό σημαίνει ότι οι ρίζες του p είναι ίσες με τις ιδιοτιμές του πίνακα.

$$A=\begin{bmatrix} 0 & & & & & -a_0 \\ 1 & 0 & & & & -a_1 \\ & 1 & 0 & & & -a_2 \\ & & 1 & \ddots & & \vdots \\ & & & \ddots & 0 & -a_{m-2} \\ & & & & 1 & -a_{m-1} \end{bmatrix}. \tag{2.3}$$

(μπορούμε επίσης να πάμε στην (2.3) κατευθείαν, χωρίς να περάσουμε από την (2.2), παρατηρώντας ότι αν z είναι ρίζα του p, τότε από την (2.1) προκύπτει ότι το $(1, z, z^2, \dots, z^{m-1})$ είναι ένα αριστερό ιδιοδιάνυσμα του A με ιδιοτιμή z.) Ο A ονομάζεται οδηγός πίνακα που αντιστοιχεί στο p.

Τώρα η δυσκολία είναι προφανής. Είναι γνωστό ότι δεν υπάρχει τύπος που να δίνει τις ρίζες ενός τυχαίου πολυωνύμου όταν οι συντελεστές του είναι γνωστοί. Αυτή λοιπόν η αδυναμία έκφρασης των ριζών ενός τυχαίου πολυωνύμου, είναι ένα από τα επιτεύγματα των μαθηματικών εργασιών των Abel, Galois και άλλων μαθηματικών το 19 αιώνα. Ο Abel απέδειξε το 1824 ότι δεν υπάρχει τύπος ανάλογος με αυτόν για τις ρίζες ενός πολυωνύμου δευτέρου βαθμού για πολυώνυμα βαθμού 5 ή μεγαλύτερου.

Θεώρημα 2.1. Για κάθε $m \geq 5$, υπάρχει ένα πολυώνυμο $p(z)$ βαθμού m με ρητούς συντελεστές που έχει πραγματική ρίζα $p(r) = 0$ με την ιδιότητα ότι η r δεν μπορεί να γραφτεί χρησιμοποιώντας οποιαδήποτε έκφραση που περιέχει μόνο ρητούς αριθμούς, πρόσθεση, αφαίρεση, πολλαπλασιασμό, διαίρεση και κ-οστές ρίζες.

Αυτό το θεώρημα σημαίνει ότι ακόμη και αν μπορούσαμε να χρησιμοποιήσουμε ακριβή αριθμητική, δεν θα μπορούσε να υπάρχει υπολογιστικό πρόγραμμα το οποίο να βρίσκει ακριβώς τις ρίζες ενός τυχαίου πολυωνύμου σε πεπερασμένο αριθμό βημάτων. Συνεπάγεται ότι το ίδιο ισχύει στο γενικότερο πρόβλημα υπολογισμού ιδιοτιμών πίνακα.

Αυτό δεν σημαίνει ότι δεν μπορούμε να γράψουμε ένα καλό πρόγραμμα εύρεσης ιδιοτιμών. Απλά ένα τέτοιο πρόγραμμα δεν μπορεί να βασιστεί στις ίδιες τεχνικές που χρησιμοποιούμε για να λύσουμε γραμμικά συστήματα. Μέθοδοι όπως οι ανακλάσεις Householder και η απαλοιφή Gauss θα μπορούσαν να λύσουν γραμμικά συστήματα εξισώσεων ακριβώς σε πεπερασμένο αριθμό βημάτων, αν μπορούσαν να εφαρμοστούν σε ακριβή αριθμητική. Αντίθετα, κάθε μέθοδος εύρεσης ιδιοτιμών πρέπει να είναι αναδρομική.

Σκοπός μιας μεθόδου εύρεσης ιδιοτιμών είναι να δημιουργήσει ακολουθίες αριθμών οι οποίες συγκλίνουν γρήγορα στις ιδιοτιμές. Από την άποψη αυτή οι υπολογισμοί ιδιοτιμών είναι πιο αντιπροσωπευτικοί του επιστημονικού υπολογισμού από τις λύσεις των γραμμικών συστημάτων εξισώσεων.

Η ανάγκη για αναδρομή μπορεί να φανεί αποθαρρυντική στην αρχή αλλά οι αλγόριθμοι που είναι διαθέσιμοι σε αυτή την περιοχή συγκλίνουν πάρα πολύ γρήγορα. Στις περισσότερες περιπτώσεις είναι δυνατόν να υπολογίσουμε ακολουθίες αριθμών που διπλασιάζουν ή τριπλασιάζουν τον αριθμό των ψηφίων ακρίβειας σε κάθε βήμα. Έτσι, μολονότι ο υπολογισμός ιδιοτιμών είναι ένα άλυτο πρόβλημα γενικά, πρακτικά διαφέρει από τη λύση γραμμικών συστημάτων κατά ένα μόνο μικρό σταθερό παράγοντα, τυπικά πιο κοντά στο 1 παρότι στο 10.

2.3. Παραγοντοποίηση Schur και Διαγωνοποίηση

Οι περισσότεροι από τους γενικού τύπου αλγορίθμους ιδιοτιμών που χρησιμοποιούνται σήμερα, χρησιμοποιούν την παραγοντοποίηση Schur. Υπολογίζουμε μια Schur παραγοντοποίηση $A = QTQ^*$ μετασχηματίζοντας τον A με μια ακολουθία στοιχειωδών ορθοκανονικών μετασχηματισμών ομοιότητας $X \to Q_j^* X Q_j$. Έτσι το γινόμενο

$$Q_j^* \dots Q_2^* Q_1^* A Q_1 Q_2 \dots Q_j, \tag{2.4}$$

συγκλίνει σε ένα άνω τριγωνικό πίνακα T καθώς $j \to \infty$.

Αν ο A είναι πραγματικός αλλά όχι συμμετρικός τότε γενικά μπορεί να έχει μιγαδικές ιδιοτιμές σε ζεύγη συζυγών. Σε αυτή την περίπτωση η Schur μορφή θα είναι μιγαδική. Έτσι ένας αλγόριθμος ο οποίος υπολογίζει την παραγοντοποίηση Schur θα πρέπει να είναι ικανός να έχει μιγαδικά αποτελέσματα από πραγματικά δεδομένα εισόδου. Αυτό μπορεί να γίνει. Εξάλλου, οι μέθοδοι εύρεσης ριζών πολυωνύμων με πραγματικούς συντελεστές έχουν την ίδια ιδιότητα. Εναλλακτικά, μπορούμε να κάνουμε ολόκληρο τον υπολογισμό σε αριθμητική πραγματικών αριθμών αν χρησιμοποιήσουμε αυτό που είναι γνωστό ως πραγματική παραγοντοποίηση Schur. Εδώ ο T επιτρέπεται να έχει 2×2 σύνθετους πίνακες στην διαγώνιο έναν για κάθε ζεύγος συζυγών μιγαδικών ιδιοτιμών. Αυτή η περίπτωση είναι σημαντική στην πράξη και περιέχεται σε όλες τις βιβλιοθήκες software.

Από την άλλη πλευρά υποθέτουμε ότι ο A είναι ερμιτιανός πίνακας. Τότε το γινόμενο $Q_j^* \dots Q_2^* Q_1^* A Q_1 Q_2 \dots Q_j$ είναι επίσης ερμιτιανός και έτσι το όριο της συγκλίνουσας ακολουθίας είναι τριγωνικός και ερμιτιανός πίνακας, δηλαδή διαγώνιος. Αυτό συνεπάγεται ότι οι ίδιοι οι αλγόριθμοι που υπολογίζουν μια ορθοκανονική τριγωνοποίηση ενός γενικού πίνακα, υπολογίζουν επίσης ορθοκανονική διαγωνοποίηση ενός ερμιτιανού πίνακα. Πρακτικά αυτό δείχνει το πώς αντιμετωπίζουμε την περίπτωση του ερμιτιανού πίνακα, αν και υπάρχουν διάφορες τροποποιήσεις που χρησιμοποιούν το πλεονέκτημα της ερμιτιανής δομής σε κάθε βήμα.

2.4. Δύο Φάσεις στον Υπολογισμό Ιδιοτιμών

Είτε είναι ο A ερμιτιανός είτε δεν είναι, η ακολουθία (2.4) χωρίζεται συνήθως σε δύο φάσεις. Στην πρώτη φάση εφαρμόζουμε μια άμεση μέθοδο προκειμένου να δημιουργήσουμε ένα άνω-Hessenberg πίνακα H, του οποίου τα στοιχεία κάτω από την πρώτη υποδιαγώνιο είναι μηδενικά. Στη δεύτερη φάση, εφαρμόζουμε μια επαναληπτική μέθοδο με σκοπό να δημιουργήσουμε μια τυπικά άπειρη ακολουθία πινάκων μορφής Hessenberg, η οποία συγκλίνει σε ένα τριγωνικό πίνακα. Σχηματικά η διαδικασία που μόλις περιγράψαμε είναι η εξής:

$$\begin{bmatrix} X & X & X & X & X \\ X & X & X & X & X \\ X & X & X & X & X \\ X & X & X & X & X \\ X & X & X & X & X \end{bmatrix} \xrightarrow{\Phi\acute{\alpha}\sigma\eta\ 1} \begin{bmatrix} X & X & X & X & X \\ X & X & X & X & X \\ & X & X & X & X \\ & & X & X & X \\ & & & X & X \end{bmatrix} \xrightarrow{\Phi\acute{\alpha}\sigma\eta 2} \begin{bmatrix} X & X & X & X & X \\ & X & X & X & X \\ & & X & X & X \\ & & & X & X \\ & & & & X \end{bmatrix}.$$

$$A \neq A^* \qquad\qquad\qquad H \qquad\qquad\qquad T$$

Η πρώτη φάση, απαιτεί $O(m^3)$ πράξεις. Η δεύτερη, αναδρομική φάση δεν τερματίζει ποτέ και αν την αφήναμε να τρέχει για πάντα θα απαιτούσε ένα άπειρο αριθμό πράξεων. Ωστόσο, η σύγκλιση με ακρίβεια μηχανής επιτυγχάνεται σε $O(m)$ αναδρομές. Κάθε αναδρομή απαιτεί $O(m^2)$ πράξεις και έτσι συνολικά στη φάση αυτή απαιτούνται $O(m^3)$ πράξεις. Χωρίς αυτό το προκαταρκτικό βήμα, όπου δημιουργείται ο άνω-Hessenberg πίνακας H, κάθε αναδρομή της φάσης 2 θα περιείχε ένα ολόκληρο πίνακα απαιτώντας $O(m^3)$ πράξεις, δηλαδή συνολικά $O(m^4)$ πράξεις ή περισσότερες καθώς η σύγκλιση μπορεί μερικές φορές να χρειάζεται περισσότερες από $O(m)$ αναδρομές.

Σε περίπτωση που ο A είναι ερμιτιανός, η διφασική προσέγγιση γίνεται ακόμα πιο γρήγορη. Ο ενδιάμεσος πίνακας είναι τώρα ένας ερμιτιανός Hessenberg πίνακας, δηλαδή τριδιαγώνιος. Το τελικό αποτέλεσμα είναι ένας ερμιτιανός τριγωνικός πίνακας, δηλαδή διαγώνιος όπως αναφέρθηκε προηγουμένως. Σχηματικά η περίπτωση του ερμιτιανού πίνακα περιγράφεται ως εξής:

$$\begin{bmatrix} X & X & X & X & X \\ X & X & X & X & X \\ X & X & X & X & X \\ X & X & X & X & X \\ X & X & X & X & X \end{bmatrix} \xrightarrow{\Phi\acute{\alpha}\sigma\eta\ 1} \begin{bmatrix} X & X & & & \\ X & X & X & & \\ & X & X & X & \\ & & X & X & X \\ & & & X & X \end{bmatrix} \xrightarrow{\Phi\acute{\alpha}\sigma\eta\ 2} \begin{bmatrix} X & & & & \\ & X & & & \\ & & X & & \\ & & & X & \\ & & & & X \end{bmatrix}.$$

$$A = A^* \qquad\qquad\qquad T \qquad\qquad\qquad D$$

Στην περίπτωση του ερμιτιανού πίνακα θα πρέπει να παρατηρήσουμε ότι αν απαιτούνται μόνο οι ιδιοτιμές (όχι τα ιδιοδιανύσματα), τότε κάθε βήμα της φάσης 2 μπορεί να εκτελεστεί με μόνο $O(m)$ πράξεις ανεβάζοντας τη συνολική εκτίμηση της εργασίας για τη φάση αυτή σε $O(m^2)$ πράξεις.

Έτσι στα προβλήματα ιδιοτιμών ερμιτιανών πινάκων είμαστε στην παράδοξη περίπτωση, όπου το "άπειρο" μέρος του αλγορίθμου είναι στην πράξη γρηγορότερο από το πεπερασμένο τμήμα του, κατά μια τάξη μεγέθους.

ΚΕΦΑΛΑΙΟ 3

Απλοποίηση σε Hessenberg ή Τριδιαγώνια Μορφή.

Στο κεφάλαιο αυτό θα περιγράψουμε την πρώτη από τις δύο φάσεις υπολογισμού που αναφέρθηκαν στο προηγούμενο κεφάλαιο. Θα ασχοληθούμε λοιπόν με την απλοποίηση ενός πίνακα σε μορφή Hessenberg μέσω μιας ακολουθίας ορθοκανονικών μετασχηματισμών ομοιότητας. Αν ο αρχικός πίνακας είναι ερμιτιανός, το αποτέλεσμα είναι τριδιαγώνιος πίνακας.

3.1. Πίνακες Householder

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

$$P = I^n - \frac{2}{u^*u} uu^*,$$

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

$$Pw = w - \frac{2}{u^*u} uu^*w = w - \frac{2u^*w}{u^*u} u,$$

είναι η αντανάκλαση του w ως προς το υπερεπίπεδο $span\{u\}^{\perp}$ (δηλαδή το ορθογώνιο συμπλήρωμα του $span\{u\}$). Ο πίνακας Householder P είναι ερμιτιανός και ορθομοναδιαίος (δηλαδή $P^*P = PP^* = I_n$). Επίσης αποτελεί μια διαταραχή του μοναδιαίου πίνακα I_n, βαθμού

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

Υποθέτουμε τώρα πως θέλουμε το Pw να είναι πολλαπλάσιο του διανύσματος e_1 της κανονικής βάσης. Τότε $Pw \in span\{e_1\}$ και $u \in span\{w, e_1\}$. Γράφουμε αυθαίρετα $u = w + ae_1$ $(a \in \mathbb{C})$ και παρατηρούμε ότι

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

όπου w_1 είναι το πρώτο στοιχείο του w το οποίο (χωρίς βλάβη της γενικότητας) μπορεί να θεωρηθεί πραγματικός αριθμός, ενώ $uu^* = w^*w + w_1 2Re(a) + |a|^2$. Επομένως,

$$Pw = \left(I_n - \frac{2}{u^*u} uu^*\right) w = \left(1 - 2\frac{w^*w + \overline{a}w_1}{w^*w + 2w_1 Re(a) + |a|^2}\right) w - 2a\frac{u^*w}{u^*u} e_1.$$

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

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

Δηλαδή, για δεδομένο διάνυσμα w, κατασκευάσαμε το διάνυσμα Householder u και τον αντίστοιχο πίνακα Householder $P = I_n - \frac{2}{u^*u}uu^*$, έτσι ώστε το διάνυσμα Pw να ανήκει στο $span\{e_1\}$.

Παράδειγμα 3.1.

Αν $w = \begin{bmatrix} 2 \\ 6 \\ 4 \\ 5 \end{bmatrix}$ με $\|w\|_2 = 9$ και $w_1 = 2$.

Τότε για το διάνυσμα Householder

$$u = w + e^{i0}\|w\|_2 e_1 = w + \|w\|_2 e_1 = \begin{bmatrix} 11 \\ 6 \\ 4 \\ 5 \end{bmatrix}.$$

Ο αντίστοιχος πίνακας Householder:

$$P = I_n - 2\frac{uu^T}{u^Tu} = I_n - \frac{2}{198}\begin{bmatrix} 121 & 66 & 44 & 55 \\ 66 & 36 & 24 & 30 \\ 44 & 24 & 16 & 20 \\ 55 & 30 & 20 & 25 \end{bmatrix} = \frac{1}{99}\begin{bmatrix} -22 & -66 & -44 & -55 \\ -66 & 63 & -24 & -30 \\ -44 & -24 & 83 & -20 \\ -55 & -30 & -20 & 74 \end{bmatrix},$$

ικανοποιεί τη σχέση

$$Pw = \begin{bmatrix} -9 \\ 0 \\ 0 \\ 0 \end{bmatrix} \in span\{e_1\}.$$

□

3.2. Μια Κακή Ιδέα

Για να υπολογίσουμε τη Schur παραγοντοποίηση $A = QTQ^*$, θα θέλαμε να εφαρμόσουμε ορθοκανονικούς μετασχηματισμούς ομοιότητας στον A έτσι ώστε να εισάγουμε μηδενικά κάτω από τη διαγώνιο. Μια πρώτη ιδέα για να επιτύχουμε ευθεία διαγωνοποίηση θα μπορούσε να είναι η χρησιμοποίηση μετασχηματισμών Householder με σκοπό να εισάγουμε αυτά τα μηδενικά το ένα μετά το άλλο.

Ο πρώτος μετασχηματισμός Householder Q_1^* πολλαπλασιασμένος εξ' αριστερών του A θα εισάγει μηδενικά κάτω από τη διαγώνιο στην πρώτη στήλη του πίνακα. Με τη διαδικασία αυτή αλλάζουν όλες οι γραμμές του A.

Σ’ αυτό και στα ακόλουθα διαγράμματα, τα στοιχεία τα οποία αλλάζουν σε κάθε βήμα είναι γραμμένα με έντονα γράμματα.

$$\underset{A}{\begin{bmatrix} X & X & X & X & X \\ X & X & X & X & X \\ X & X & X & X & X \\ X & X & X & X & X \\ X & X & X & X & X \end{bmatrix}} \xrightarrow{Q_1^* \cdot} \underset{Q_1^* A}{\begin{bmatrix} \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} \\ \mathbf{0} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} \\ \mathbf{0} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} \\ \mathbf{0} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} \\ \mathbf{0} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} \end{bmatrix}}.$$

Δυστυχώς, για να ολοκληρώσουμε το μετασχηματισμό ομοιότητας θα πρέπει να πολλαπλασιάσουμε με Q_1 εκ των δεξιών του A.

$$\underset{Q_1^* A}{\begin{bmatrix} X & X & X & X & X \\ & X & X & X & X \\ & X & X & X & X \\ & X & X & X & X \\ & X & X & X & X \end{bmatrix}} \xrightarrow{\cdot Q_1} \underset{Q_1^* A Q_1}{\begin{bmatrix} \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} \\ \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} \\ \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} \\ \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} \\ \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} & \boldsymbol{X} \end{bmatrix}}.$$

Σε αυτή την περίπτωση κάθε στήλη του πίνακα, αντικαθίσταται από ένα γραμμικό συνδυασμό όλων των στηλών. Το αποτέλεσμα είναι, τα μηδενικά τα οποία εισήχθησαν προηγουμένως κάτω από την διαγώνιο στην πρώτη στήλη του πίνακα, να καταστραφούν και τελικά να καταλήξουμε σε μια κατάσταση παρόμοια με την αρχική. Φυσικά γνωρίζουμε ότι αυτή η ιδέα δεν θα πετύχαινε εξαιτίας της "βασικής δυσκολίας" που περιγράφηκε στο προηγούμενο κεφάλαιο. Καμία πεπερασμένη διαδικασία δεν μπορεί να υπολογίσει ακριβώς τις ιδιοτιμές ενός πίνακα A.

Ωστόσο, αυτή η τόσο απλή στρατηγική η οποία μοιάζει μάταια καθώς τη συζητάμε, μπορεί να μειώνει το μέγεθος των στοιχείων κάτω από τη διαγώνιο ακόμη και αν δεν τα εξισώνει με το μηδέν.

3.3. Μια Καλή Ιδέα

Η σωστή στρατηγική για να δημιουργήσουμε μηδενικά στη φάση 1 είναι να επιχειρήσουμε σε λιγότερα στοιχεία του πίνακα.

Στο πρώτο βήμα, θα επιλέξουμε ένα μετασχηματισμό Householder Q_1^* ο οποίος όταν πολλαπλασιάζεται εκ των αριστερών του A, αφήνει την πρώτη σειρά αναλλοίωτη και στη συνέχεια σχηματίζει γραμμικούς συνδυασμούς των γραμμών μόνο $2, \dots, m$ για να εισάγει μηδενικά στις σειρές $3, \dots, m$ της πρώτης στήλης. Κατόπιν όταν το Q_1 πολλαπλασιαστεί με τον $Q_1^* A$ από τα δεξιά, αφήνει την πρώτη στήλη αναλλοίωτη, σχηματίζει γραμμικούς συνδυασμούς των στηλών $2, \dots, m$ και αφήνει ως είναι τα μηδενικά που δημιουργήθηκαν.

$$\begin{bmatrix} X & X & X & X & X \\ X & X & X & X & X \\ X & X & X & X & X \\ X & X & X & X & X \\ X & X & X & X & X \end{bmatrix} \xrightarrow{Q_1^*\cdot} \begin{bmatrix} X & X & X & X & X \\ \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ \mathbf{0} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ \mathbf{0} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ \mathbf{0} & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} \end{bmatrix} \xrightarrow{\cdot Q_1} \begin{bmatrix} X & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ X & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} \end{bmatrix}.$$

$$A \qquad\qquad\qquad Q_1^* A \qquad\qquad\qquad Q_1^* A Q_1$$

Αυτή η ιδέα επαναλαμβάνεται για να εισάγει μηδενικά στις ακόλουθες στήλες. Για παράδειγμα ο δεύτερος μετασχηματισμός Householder Q_2, αφήνει την πρώτη και τη δεύτερη σειρά και στήλη ανεπηρέαστες.

$$\begin{bmatrix} X & X & X & X & X \\ X & X & X & X & X \\ & X & X & X & X \\ & X & X & X & X \\ & X & X & X & X \end{bmatrix} \xrightarrow{Q_2^*\cdot} \begin{bmatrix} X & X & X & X & X \\ X & X & X & X & X \\ & \mathbf{X} & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ & \mathbf{0} & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ & \mathbf{0} & \mathbf{X} & \mathbf{X} & \mathbf{X} \end{bmatrix} \xrightarrow{\cdot Q_2} \begin{bmatrix} X & X & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ X & X & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ & X & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ & & \mathbf{X} & \mathbf{X} & \mathbf{X} \\ & & \mathbf{X} & \mathbf{X} & \mathbf{X} \end{bmatrix}.$$

$$Q_1^* A Q_1 \qquad\qquad\qquad Q_2^* Q_1^* A Q_1 \qquad\qquad\qquad Q_2^* Q_1^* A Q_1 Q_2$$

Μετά από επανάληψη αυτής της διαδικασίας $m-2$ φορές έχουμε ως αποτέλεσμα ένα πίνακα μορφής Hessenberg

$$\begin{bmatrix} X & X & X & X & X \\ X & X & X & X & X \\ & X & X & X & X \\ & & X & X & X \\ & & & X & X \end{bmatrix}$$

$$Q_{m-2}^* \dots Q_2^* Q_1^* A Q_1 Q_2 \dots Q_{m-2} = H.$$

όπου $Q_{m-2}^* \dots Q_2^* Q_1^* = Q^*$ και $Q_1 Q_2 \dots Q_{m-2} = Q$.

Ο αλγόριθμος διατυπώνεται παρακάτω:

Αλγόριθμος 3.1. Αναγωγή τύπου Householder ενός πίνακα A, σε μορφή Hessenberg.

$$\begin{cases} \text{για } k = 1 \text{ μέχρι } m-2 \\ x = A_{k+1:m,k} \\ u_k = sign(x_1)\|x\|_2 e_1 + x \\ u_k = u_k / \|u_k\|_2 \\ A_{k+1:m,k:m} = A_{k+1:m,k:m} - 2u_k\big(u_k^* A_{k+1:m,k:m}\big) \\ A_{1:m,k+1:m} = A_{1:m,k+1:m} - 2\big(A_{1:m,k+1:m} u_k\big)u_k^* \end{cases}$$

Εδώ ο πίνακας $Q = \prod_{k=1}^{m-2} Q_k$ δεν σχηματίζεται ρητά. Τα διανύσματα των ανακλάσεων u_k σώζονται αντί για αυτόν και μπορούν να χρησιμοποιηθούν για να πολλαπλασιάσουμε με τον Q ή για να ανακατασκευάσουμε τον Q αργότερα εάν χρειαστεί.

3.4. Μέτρηση Πράξεων

Ο αριθμός των πράξεων που απαιτούνται από τον Αλγόριθμο 3.1 μπορεί να υπολογιστεί με την ίδια γεωμετρική λογική που χρησιμοποιήσαμε πριν.

Η μέθοδος αυτή διαχωρίζει τον πίνακα A σε δύο υποπίνακες. Στον πρώτο βρόγχο επανάληψης εφαρμόζεται ένας ανακλαστής Householder στα αριστερά του πίνακα. Ο κ-οστός τέτοιος ανακλαστής επηρεάζει τις τελευταίες $m-k$ γραμμές. Από τη στιγμή που ο ανακλαστής εφαρμόζεται, αυτές οι γραμμές έχουν μηδενικά στις $k-1$ στήλες και πρέπει να εκτελέσουμε μόνο τις πράξεις που αφορούν τα τελευταία $m-k+1$ στοιχεία κάθε γραμμής.

Με τέσσερις πράξεις για κάθε στοιχείο, το σύνολο των πράξεων σε αυτό το βρόγχο επανάληψης είναι $\sim\frac{4}{3}m^3$.

Στο δεύτερο εσωτερικό βρόγχο εφαρμόζεται ένας ανακλαστής Householder στα δεξιά του πίνακα. Στο κ-οστό βήμα ο ανακλαστής διαμορφώνει γραμμικούς συνδυασμούς των τελευταίων $m-k$ στηλών. Σ' αυτό τον βρόγχο επανάληψης απαιτείται περισσότερη δουλειά απ' ότι στο πρώτο, καθώς δεν υπάρχουν μηδενικά που μπορούν να αγνοηθούν. Πρέπει λοιπόν να υπολογίσουμε και τα m στοιχεία κάθε στήλης από αυτές που αλλάζουν, δηλαδή συνολικά $m(m-k)$ στοιχεία για κάθε τιμή του k.

Έτσι για τέσσερις πράξεις ανά στοιχείο, αυτός ο δεύτερος βρόχος απαιτεί $\sim 2m^3$ πράξεις.
Συνολικά, η αναγωγή ενός $m \times m$ πίνακα σε μορφή Hessenberg απαιτεί $\sim\frac{10}{3}m^3$πράξεις. (3.1)

3.5. Η Ερμιτιανή Περίπτωση: Μετατροπή σε Τριδιαγώνια Μορφή

Αν ο A είναι ερμιτιανός ο αλγόριθμος που μόλις περιγράφηκε θα μετατρέψει τον A σε τριδιαγώνια μορφή (τουλάχιστον κατά την απουσία σφαλμάτων στρογγυλοποίησης). Αυτό είναι εύκολο να το αντιληφθούμε: Από τη στιγμή που ο A είναι ερμιτιανός, ο πίνακας Q^*AQ είναι επίσης ερμιτιανός αλλά και Hessenberg πίνακας δηλαδή τριδιαγώνιος.

Καθώς τώρα εισάγονται μηδενικά στις γραμμές όπως και στις στήλες, μπορούμε αγνοώντας αυτά τα επιπρόσθετα μηδενικά να αποφύγουμε τις επιπλέον αριθμητικές πράξεις. Με αυτή τη βελτιστοποίηση εφαρμόζοντας ένα ανακλαστή Householder στα δεξιά ενός πίνακα, το κόστος του σε πράξεις είναι εξίσου χαμηλό με το αν εφαρμοστεί στα αριστερά του ίδιου πίνακα όπου το συνολικό κόστος και για τις δύο περιπτώσεις μειώνεται από $2m^3$ σε $\frac{4}{3}m^3$ πράξεις. Συνεπώς η συνολική ποσότητα των πράξεων που απαιτείται για την μετατροπή είναι $\frac{8}{3}m^3$ πράξεις.

Αυτό το αποτέλεσμα ωστόσο βασίζεται στην αραιότητα του νέου πίνακα και όχι στη συμμετρικότητα του. Στην πραγματικότητα το γεγονός ότι σε κάθε βήμα του υπολογισμού, ο πίνακας είναι ερμιτιανός, μπορεί να ληφθεί ως πλεονέκτημα, καθώς σε αυτή την περίπτωση αρκεί να υπολογίσουμε τα μισά στοιχεία του πίνακα και επομένως το συνολικό κόστος εργασίας μειώνεται στο μισό, αφού:

Εργασία για μετατροπή σε τριδιαγώνια μορφή : $\sim\frac{4}{3}m^3$ πράξεις. (3.2)

3.6. Ευστάθεια

Ο αλγόριθμος που μόλις περιγράφηκε είναι backward ευσταθής.

Έστω $\widetilde{H}$, ο πραγματικός Hessenberg πίνακας υπολογισμένος με αριθμητική κινητής υποδιαστολής και έστω $\tilde{Q}$ ορθομοναδιαίος πίνακας που αντιστοιχεί στα διανύσματα-ανακλαστές $\tilde{u}_k$ υπολογισμένα επίσης με αριθμητική κινητής υποδιαστολής.

Θεώρημα 3.1. Έστω η απλοποίηση Hessenberg $A = QHQ^*$ ενός πίνακα $A \in \mathbb{C}^{m\times m}$ υπολογισμένη από τον Αλγόριθμο 3.1 σε έναν υπολογιστή που ικανοποιεί τα αξιώματα (3.1) και (3.2), και έστω οι πίνακες $\tilde{Q}$ και $\widetilde{H}$ οι οποίοι ορίζονται όπως περιγράφεται παραπάνω. Τότε έχουμε:

$$\tilde{Q}\widetilde{H}\tilde{Q}^* = A + \delta A, \qquad \frac{\|\delta A\|}{\|A\|} = O(\epsilon_{μηχανής}) \tag{3.3}$$

για κάποιο $\delta A \in \mathbb{C}^{m\times m}$.

Αξίωμα 3.1. Για κάθε $x \in \mathbb{R}$ υπάρχει ϵ με $|\epsilon| \le \epsilon_{μηχανής}$ τέτοιο ώστε $fl(x) = x(1+\epsilon)$ όπου $fl : \mathbb{R} \to \mathrm{F}$ μια απεικόνιση που αντιστοιχεί ένα πραγματικό αριθμό στο πιο κοντινό του ρητό.

Αξίωμα 3.2. Για κάθε $x, y \in \mathrm{F}$, υπάρχει ϵ με $|\epsilon| \le \epsilon_{μηχανής}$ τέτοιο ώστε $x \circledast y = (x * y)(1 + \epsilon)$ όπου $fl : \mathbb{R} \to \mathrm{F}$ μια απεικόνιση ίδια με την προηγούμενη (αξίωμα 3.2).

ΚΕΦΑΛΑΙΟ 4

Πηλίκο Rayleigh, Αντίστροφη Επανάληψη

Σε αυτό το κεφάλαιο θα παρουσιάσουμε μερικούς κλασικούς αλγορίθμους εύρεσης ιδιοτιμών. Αυτοί είναι πολύ χρήσιμοι σε συγκεκριμένες περιπτώσεις, ειδικά στην Αντίστροφη Επανάληψη η οποία είναι μια καθιερωμένη μέθοδος για τον προσδιορισμό ενός ιδιοδιανύσματος όταν η αντίστοιχη ιδιοτιμή του είναι γνωστή.

4.1.Περιορισμός σε Πραγματικούς Συμμετρικούς Πίνακες

Στην αριθμητική γραμμική άλγεβρα, οι περισσότερες αλγοριθμικές ιδέες εφαρμόζονται είτε σε γενικούς πίνακες είτε σε ερμιτιανούς. Αυτό είναι κάτι που θα ισχύει, τουλάχιστον εν μέρει, για τις αριθμητικές μεθόδους με τις οποίες θα ασχοληθούμε σε αυτό, αλλά και στα επόμενα τρία κεφάλαια. Ωστόσο μερικές από τις διαφορές μεταξύ των γενικών και των ερμιτιανών περιπτώσεων είναι μάλλον μεγάλες. Άρα σε αυτά τα τέσσερα κεφάλαια για την απλοποίηση της παρουσίασης αυτών των αλγορίθμων θα περιοριστούμε σε πίνακες που είναι πραγματικοί και συμμετρικοί. Θα υποθέσουμε επίσης ότι $\|\cdot\| = \|\cdot\|_2$ που σημαίνει ότι η νόρμα στο χώρο $\mathbb{R}^m$ ισούται με την ευκλείδεια νόρμα στον ίδιο χώρο.

Επίσης για αυτά τα τέσσερα κεφάλαια θα έχουμε :

$$A = A^T \in \mathbb{R}^{m\times m},\ x \in \mathbb{R}^m,\ x^* = x^T,\ \|x\| = \sqrt{x^T x}\,.$$

Πιο συγκεκριμένα αυτό σημαίνει ότι ο A έχει πραγματικές ιδιοτιμές και ένα πλήρες σύνολο ορθογώνιων ιδιοδιανυσμάτων. Θα χρησιμοποιήσουμε τον ακόλουθο συμβολισμό :

Πραγματικές ιδιοτιμές: $\lambda_1, \dots, \lambda_m$,

Ορθοκανονικά ιδιοδιανύσματα : $q_1, \dots, q_m$.

Τα ιδιοδιανύσματα θεωρούνται κανονικοποιημένα δηλαδή: $\|q_j\| = 1$ και η διάταξη των ιδιοτιμών θα καθορίζεται όπως χρειάζεται σε κάθε περίπτωση.

Οι περισσότερες από τις ιδέες που περιγράφονται στα επόμενα κεφάλαια ανήκουν στη δεύτερη από τις δύο φάσεις που περιγράφηκαν στο Κεφάλαιο 2. Αυτό σημαίνει ότι από τη στιγμή που εφαρμόζουμε αυτές τις ιδέες, ο A δεν θα είναι απλώς ένας πραγματικός και συμμετρικός πίνακας, αλλά θα είναι και τριδιαγώνιος. Αυτή η τριδιαγώνια μορφή του έχει μαθηματική σημασία, για παράδειγμα στην επιλογή εναλλαγών στον αλγόριθμο QR, και επίσης έχει αλγοριθμική σημασία καθώς μειώνει τις πράξεις σε κάθε βήμα από $O(m^3)$ σε $O(m)$.

4.2. Το Πηλίκο Rayleigh

Το Πηλίκο Rayleigh ενός διανύσματος $x \in \mathbb{R}^m$ είναι ο αριθμός

$$r(x) = \frac{x^T A x}{x^T x} \; . \qquad (4.1)$$

Σημειώνουμε ότι αν ο x είναι ένα ιδιοδιάνυσμα τότε η $r(x) = \lambda$ είναι η αντίστοιχη ιδιοτιμή του. Ένας τρόπος για να χρησιμοποιήσουμε αυτό τον τύπο είναι να αναρωτηθούμε δεδομένου του x, ποιος αριθμός a «συμπεριφέρεται περισσότερο ως ιδιοτιμή» του x με την έννοια ότι ελαχιστοποιεί την $\|Ax - ax\|$. Αυτό είναι ένα $m \times 1$ πρόβλημα ελαχίστων τετραγώνων της μορφής $xa \approx Ax$ (όπου ο x είναι ο πίνακας, a ένα άγνωστο διάνυσμα, Ax το δεξί μέλος). Γράφοντας την κανονική εξίσωση $x^T x a = x^T A x$ για αυτό το σύστημα παίρνουμε την απάντηση : $a = r(x)$. Έτσι το $r(x)$ μπορεί να θεωρηθεί ότι είναι μια φυσιολογική εκτίμηση της ιδιοτιμής αν το x προσεγγίζει αλλά δεν είναι απαραίτητα ίδιο με ιδιοδιάνυσμα.

Για να κάνουμε αυτές τις ιδέες πιο ποσοτικές, είναι σκόπιμο να δούμε το $x \in \mathbb{R}^m$ ως μια μεταβλητή, έτσι ώστε το r να είναι μια συνάρτηση $\mathbb{R}^m \to \mathbb{R}$. Μας ενδιαφέρει η συμπεριφορά του $r(x)$ τοπικά όταν το x είναι κοντά σε ένα ιδιοδιάνυσμα. Ένας τρόπος για να το προσεγγίσουμε αυτό είναι να υπολογίσουμε τις μερικές παραγώγους του $r(x)$ ως προς τις συνιστώσες x_j .

$$\frac{\partial r(x)}{\partial x_j} = \frac{\frac{\partial}{\partial x_j}(x^T A x)}{x^T x} - \frac{(x^T A x)\frac{\partial}{\partial x_j}(x^T x)}{(x^T x)^2} = \frac{2(Ax)j}{x^T x} - \frac{(x^T A x) 2 x_j}{(x^T x)^2} = \frac{2}{x^T x}(Ax - r(x)x)_j.$$

Εάν συλλέξουμε αυτές τις μερικές παραγώγους σε ένα m-διάνυσμα βρίσκουμε ότι έχουμε υπολογίσει την κλίση του $r(x)$, που συμβολίζεται με $\nabla r(x)$. Έχουμε δείξει λοιπόν ότι:

$$\nabla r(x) = \frac{2}{x^T x}(Ax - r(x)x). \qquad (4.2)$$

Από αυτόν τον τύπο βλέπουμε ότι για κάθε ιδιοδιάνυσμα x του A, η κλίση του $r(x)$ είναι το μηδενικό διάνυσμα (πράγματι αντικαθιστώντας στη σχέση (4.2) όπου $r(x) = \lambda$ και $Ax=\lambda x$ προκύπτει $\nabla r(x) = \frac{2}{x^T x}(\lambda x - \lambda x) = 0$). Αντιστρόφως, αν $\nabla r(x) = 0$ με $x \neq 0$, τότε το x είναι ένα ιδιοδιάνυσμα και $r(x)$ είναι η αντίστοιχη ιδιοτιμή του.

Από γεωμετρική άποψη, τα ιδιοδιανύσματα του A είναι τα τοπικά ακρότατα της συνάρτησης $r(x)$ και οι ιδιοτιμές του A είναι οι τιμές της $r(x)$ σ' αυτά τα τοπικά ακρότατα. Τα ιδιοδιανύσματα λοιπόν του A είναι σημεία του $\mathbb{R}^m$ τα οποία ανήκουν σε ευθείες οι οποίες διέρχονται από την αρχή των αξόνων. Αν τα κανονικοποιήσουμε, τα σημεία αυτά θα βρίσκονται στην επιφάνεια της μοναδιαίας σφαίρας (υποθέτοντας ότι οι ιδιοτιμές του A είναι απλές).

Έστω q_J ένα από τα ιδιοδιανύσματα του A. Από το γεγονός ότι $\nabla r(q_J) = 0$ και από την ομαλότητα της συνάρτησης $r(x)$ (παντού εκτός από το σημείο $x = 0$). Εξάγουμε ένα σημαντικό αποτέλεσμα:

$$r(x) - r(q_J) = O\left(\|x - q_J\|^2\right) \text{ καθώς } x \to q_J. \quad (4.3)$$

Έτσι το Πηλίκο Rayleigh είναι μια τετραγωνικά ακριβής εκτίμηση μιας ιδιοτιμής. Από αυτό προκύπτει η δύναμη του.

Ένας πιο σαφής τρόπος για να προκύψει η (4.3) είναι να επεκτείνουμε το x ως γραμμικό συνδυασμό των ιδιοδιανυσμάτων $q_1, \dots, q_m$ του A. Αν $x = \sum_{j=1}^{m} a_j q_j$, τότε $r(x) = \sum_{j=1}^{m} {a_j}^2 \lambda_j / \sum_{j=1}^{m} {a_j}^2$. Έτσι το $r(x)$ είναι ο σταθμικός μέσος των ιδιοτιμών του A με βάρη ίσα με τα τετράγωνα των συντεταγμένων του x στη βάση των ιδιοδιανυσμάτων. Εξαιτίας αυτού του τετραγωνισμού των συντεταγμένων δεν είναι δύσκολο να δούμε ότι εάν $|a_j/a_J| \leq \epsilon$ για κάθε $j \neq J$, τότε $r(x) - r(q_J) = O(\epsilon^2)$.

4.3. Μέθοδος των Δυνάμεων

Τώρα αλλάζουμε τακτική. Υποθέτουμε ότι $u^{(0)}$ είναι ένα διάνυσμα με $\|u^{(0)}\| = 1$. Η διαδικασία που ακολουθεί, η Μέθοδος των Δυνάμεων, είχε αναφερθεί στις αρχές του δεύτερου Κεφαλαίου ως μια όχι και τόσο καλή ιδέα εξαιτίας του ότι είναι πολύ αργή. Μπορεί να περιμένει κανείς από αυτή τη μέθοδο να παραχθεί μια ακολουθία $u^{(i)}$ η οποία συγκλίνει σε ένα ιδιοδιάνυσμα το οποίο αντιστοιχεί στη μεγαλύτερη ιδιοτιμή του πίνακα A .

Αλγόριθμος 4.1. Μέθοδος των Δυνάμεων

$$\begin{aligned}
&u^{(0)} = \text{κάποιο διάνυσμα με } \|u^{(0)}\| = 1 \\
&\text{για } k = 1,2,.. \\
&\qquad w = Au^{(k-1)} && \text{εφαρμόζει τον A} \\
&\qquad u^{(k)} = w/\|w\| && \text{κανονικοποιεί} \\
&\qquad \lambda^{(k)} = \left(u^{(k)}\right)^T A u^{(k)} && \text{Πηλίκο Rayleigh}
\end{aligned}$$

Ο αλγόριθμος επομένως της Μεθόδου των Δυνάμεων παράγει επαναληπτικά την ακολουθία $u^{(k)}$, για $k = 1,2, \dots$, και θεωρεί στην k επανάληψη το Πηλίκο Rayleigh $\lambda^{(k)} = \left(u^{(k)}\right)^T A u^{(k)}$ ως την προσέγγιση της ιδιοτιμής λ, και το διάνυσμα $u^{(k)}$ ως την προσέγγιση του ιδιοδιανύσματος που αντιστοιχεί στην ιδιοτιμή λ.

Σε αυτόν τον αλγόριθμο και στον επόμενο που ακολουθεί δεν δίνουμε σημασία στις συνθήκες τερματισμού, περιγράφουμε το βρόχο μόνο με την έκφραση "$\text{για } k = 1,2,..$". Φυσικά στην πράξη, οι συνθήκες τερματισμού είναι πολύ σημαντικές.

Μπορούμε να αναλύσουμε τη Μέθοδο των Δυνάμεων εύκολα. Γράφουμε το $u^{(0)}$ ως ένα γραμμικό συνδυασμό των ορθοκανονικών ιδιοδιανυσμάτων q_i:

$$u^{(0)} = a_1 q_1 + a_2 q_2 + \cdots + a_m q_m.$$

Αφού το $u^{(k)}$ είναι πολλαπλάσιο του $A^k u^{(0)}$ έχουμε για κάποια σταθερά c_k

$$\begin{aligned} u^{(k)} &= c_k A^k u^{(0)} \\ &= c_k\left(a_1 \lambda_1^k q_1 + a_2 \lambda_2^k q_2 + \cdots + a_m \lambda_m^k q_m\right) \\ &= c_k \lambda_1^k (a_1 q_1 + a_2 (\lambda_2/\lambda_1)^k q_2 + \cdots + a_m (\lambda_m/\lambda_1)^k q_m). \end{aligned} \tag{4.4}$$

Από εδώ παίρνουμε το ακόλουθο συμπέρασμα.

Θεώρημα 4.1. Υποθέτουμε ότι $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_m| \geq 0$ και $q_1^T u^{(0)} \neq 0$. Τότε οι επαναλήψεις του Αλγορίθμου 4.1 ικανοποιούν:

$$\left\|u^{(k)} - (\pm q_1)\right\| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right), \qquad \left|\lambda^{(k)} - \lambda_1\right| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right) \tag{4.5}$$

καθώς $k \to \infty$. Το πρόσημο $\pm$ σημαίνει ότι σε κάθε βήμα k, οποιοδήποτε από το τις δύο επιλογές πρόσημου μπορεί να γίνει και το φράγμα να ισχύει.

Απόδειξη

Η πρώτη εξίσωση προκύπτει από τη σχέση (4.4). Υποθέτουμε ότι $a_1 = q_1^T u^{(0)} \neq 0$.
Έστω λοιπόν η τυχαία αρχική προσέγγιση ιδιοδιανύσματος

$$u^{(0)} = a_1 q_1 + a_2 q_2 + \cdots + a_m q_m.$$

Από τη σχέση (4.4) έχουμε:

$$\begin{aligned} u^{(k)} &= c_k \lambda_1^k (a_1 q_1 + a_2 (\lambda_2/\lambda_1)^k q_2 + \cdots + a_m (\lambda_m/\lambda_1)^k q_m) \\ &= c_k \lambda_1^k a_1 q_1 + O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right). \end{aligned}$$

Από υπόθεση ισχύει $a_1 = q_1^T u^{(0)} \neq 0$. Άρα η προηγούμενη σχέση γίνεται

$$u^{(k)} = c_k \lambda_1^k q_1^T u^{(0)} q_1 + O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right).$$

Στη συνέχεια πολλαπλασιάζουμε τον όρο $O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$ με $(q_1^T u^{(0)})^{-1} q_1^T u^{(0)}$ από δεξιά (αφού $q_1^T u^{(0)} \neq 0$) και προκύπτει:

$$u^{(k)} = c_k \lambda_1^k q_1^T u^{(0)} q_1 + O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)(q_1^T u^{(0)})^{-1} q_1^T u^{(0)} = (c_k \lambda_1^k q_1 + O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)) q_1^T u^{(0)}.$$

Καθώς $k \to \infty$, λαμβάνουμε προσεγγιστικά ότι $u^{(k)} \approx c_k \lambda_1^k q_1 \approx q_1$ άρα πράγματι ισχύει $\left\|u^{(k)} - (\pm q_1)\right\| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$.

Η απόδειξη της δεύτερης σχέσης προκύπτει άμεσα από τις (4.4) και (4.3). Αν $\lambda_1 > 0$, τότε τα πρόσημα $\pm$ είναι όλα $+$ ή $-$, ενώ αν $\lambda_1 < 0$, τα πρόσημα εναλλάσσονται. □

Όσων αφορά τη Μέθοδο των Δυνάμεων είναι περιορισμένης χρήσης, για διάφορους λόγους. Πρώτον, προσεγγίζει μόνο το ιδιοδιάνυσμα που αντιστοιχεί στη μεγαλύτερη κατά μέτρο ιδιοτιμή. Δεύτερον, η σύγκλιση είναι γραμμική μειώνοντας το σφάλμα μόνο κατά ένα σταθερό όρο $\approx |\lambda_2/\lambda_1|$ σε κάθε επανάληψη. Τελικά ο όρος αυτός εξαρτάται από τη μέγιστη ιδιοτιμή και το πόσο μεγαλύτερη είναι από τις άλλες. Αν οι δύο μεγαλύτερες ιδιοτιμές είναι πολύ κοντά σε μέγεθος μεταξύ τους, τότε η σύγκλιση θα είναι πολύ αργή.

Ευτυχώς υπάρχει τρόπος να ενισχύσουμε τις διαφορές μεταξύ των ιδιοτιμών.

4.4. Αντίστροφη Επανάληψη

Για κάθε μ∈ $\mathbb{R}$ που δεν είναι ιδιοτιμή του A, τα ιδιοδιανύσματα του $(A-\mu I)^{-1}$ είναι τα ίδια τα ιδιοδιανύσματα του A, και οι αντίστοιχες ιδιοτιμές είναι $\left\{(\lambda_j-\mu)^{-1}\right\}$, όπου $\{\lambda_j\}$ είναι οι ιδιοτιμές του A. Αυτό προτείνει μια ιδέα: υποθέτουμε ότι το μ είναι κοντά σε μια ιδιοτιμή λ_J του A. Τότε το $(\lambda_J-\mu)^{-1}$μπορεί να είναι πολύ μεγαλύτερο από το $(\lambda_j-\mu)^{-1}$για όλα τα $j \neq J$. Έτσι αν εφαρμόσουμε τη Μέθοδο των Δυνάμεων στον $(A-\mu I)^{-1}$, η διαδικασία θα συγκλίνει γρήγορα στο q_J. Αυτή η ιδέα καλείται Αντίστροφη Επανάληψη.

Αλγόριθμος 4.2. Αντίστροφη Επανάληψη

$u^{(0)} =$ κάποιο διάνυσμα με $\left\|u^{(0)}\right\| = 1$
για $k = 1,2, \dots$

λύνουμε το σύστημα $(A-\mu I)w = u^{(k-1)}$ως προς w	εφαρμόζει τον $(A-\mu I)^{-1}$
$u^{(k)} = w/\|w\|$	κανονικοποιεί
$\lambda^{(k)} = \left(u^{(k)}\right)^T A u^{(k)}$	Πηλίκο Rayleigh

Σε αυτό το σημείο τίθεται το ερώτημα για το τι θα γίνει σε περίπτωση που ο μ είναι μια ιδιοτιμή του A, οπότε ο $A-\mu I$ είναι μη αντιστρέψιμος, ή σε περίπτωση που ο μ είναι σχεδόν ιδιοτιμή του A, οπότε δεν μπορεί να βρεθεί μια ακριβής λύση του $(A-\mu I)w = u^{(k-1)}$. Αυτές οι προφανείς παγίδες της Αντίστροφης Επανάληψης δεν θα προκαλέσουν κανένα πρόβλημα.
Όπως η Μέθοδος των Δυνάμεων, η Αντίστροφη Επανάληψη συγκλίνει μόνο γραμμικά. Αντίθετα όμως με την Μέθοδο των Δυνάμεων, μπορούμε να επιλέξουμε ένα ιδιοδιάνυσμα το οποίο βρίσκουμε χρησιμοποιώντας μια εκτίμηση μ της αντίστοιχης ιδιοτιμής. Επιπλέον ο ρυθμός της γραμμικής σύγκλισης της μεθόδου μπορεί να ελεγχθεί καθώς εξαρτάται από το μ.

Αν το μ είναι περισσότερο κοντά σε μια ιδιοτιμή του A απ' ότι στις άλλες, τότε η μέγιστη ιδιοτιμή του $(A-\mu I)^{-1}$ θα είναι πολύ μεγαλύτερη από τις υπόλοιπες. Χρησιμοποιώντας την ίδια αιτιολογία με την Μέθοδο των Δυνάμεων, διατυπώνουμε το ακόλουθο θεώρημα.

Θεώρημα 4.2. Υποθέτουμε ότι η λ_J είναι η πιο κοντινή ιδιοτιμή στο μ και η λ_K είναι η δεύτερη πιο κοντινή. Δηλαδή $|\mu-\lambda_J| < |\mu-\lambda_K| \leq |\mu-\lambda_j|$ για κάθε $j \neq J$. Επιπλέον υποθέτουμε $q_J^T u^{(0)} \neq 0$. Τότε οι επαναλήψεις του Αλγορίθμου 4.2 ικανοποιούν:

$$\left\|u^{(k)} - (\pm q_J)\right\| = O\left(\left|\frac{\mu-\lambda_J}{\mu-\lambda_K}\right|^{k}\right), \qquad \left|\lambda^{(k)} - \lambda_J\right| = O\left(\left|\frac{\mu-\lambda_J}{\mu-\lambda_K}\right|^{2k}\right)$$

καθώς $k \to \infty$, όπου τα πρόσημα $\pm$ έχουν την ίδια σημασία με το Θεώρημα 4.1.

Η Αντίστροφη Επανάληψη είναι ένα από τα πιο πολύτιμα εργαλεία της αριθμητικής γραμμικής άλγεβρας, για το λόγω αυτό αποτελεί τη κλασσική μέθοδο υπολογισμού ενός ή περισσοτέρων ιδιοδιανυσμάτων ενός πίνακα όταν οι ιδιοτιμές είναι ήδη γνωστές. Τότε ο Αλγόριθμος 4.2 εφαρμόζεται όπως είναι, χωρίς όμως να υπολογίζεται το Πηλίκο του Rayleigh.

4.5. Επαναληπτική Μέθοδος Πηλίκου Rayleigh.

Μέχρι στιγμής σε αυτό το κεφάλαιο έχουμε παρουσιάσει μια μέθοδο εύρεσης προσεγγίσεων ιδιοτιμών από εκτιμήσεις ιδιοδιανυσμάτων (το Πηλίκο Rayleigh), και μια άλλη μέθοδο για την εύρεση προσεγγίσεων ιδιοδιανυσμάτων από εκτιμήσεις ιδιοτιμών (Αντίστροφη Επανάληψη).

Ο τρόπος με τον οποίο συνδυάζονται οι δύο αυτές μέθοδοι προκειμένου να βρούμε εκτιμήσεις των ιδιοτιμών και ιδιοδιανυσμάτων είναι ο ακόλουθος : Από μια εκτίμηση μιας ιδιοτιμής λ_J μπορούμε με ένα βήμα της Μεθόδου της Αντίστροφης Επανάληψης να πάρουμε μια εκτίμηση για ένα ιδιοδιάνυσμα q_J, και από αυτό το ιδιοδιάνυσμα (με το Πηλίκο Rayleigh) να βρούμε μια νέα εκτίμηση της ιδιοτιμής λ_J. Προϋπόθεση για να γίνουν όλα αυτά είναι να υπάρχει μια προκαταρτική εκτίμηση του διανύσματος q_J.

Ο αλγόριθμος που ακολουθεί καλείται Μέθοδος Πηλίκου Rayleigh.

Αλγόριθμος 4.3. Επαναληπτική Μέθοδος Πηλίκου Rayleigh

$u^{(0)} =$ κάποιο διάνυσμα με $\|u^{(0)}\| = 1$
$\lambda^{(0)} = \left(u^{(0)}\right)^T A u^{(0)} =$ το αντίστοιχο στο $u^{(0)}$ Πηλίκο Rayleigh
για $k = 1,2, \ldots$
λύνουμε το σύστημα $\left(A - \lambda^{(k-1)} I\right) w = u^{(k-1)}$ως προς w εφαρμόζει τον $\left(A - \lambda^{(k-1)} I\right)^{-1}$
$u^{(k)} = w/\|w\|$ κανονικοποιεί
$\lambda^{(k)} = \left(u^{(k)}\right)^T A u^{(k)}$ Πηλίκο Rayleigh

Η σύγκλιση σε αυτόν τον αλγόριθμο είναι πολύ γρήγορη: Σε κάθε επανάληψη τριπλασιάζει τον αριθμό των ψηφίων ακρίβειας.

Θεώρημα 4.3. Η Επαναληπτική Μέθοδος Πηλίκου Rayleigh συγκλίνει σε ένα ζεύγος ιδιοτιμής/ιδιοδιανύσματος για κάθε αρχικό διάνυσμα $u^{(0)}$ εκτός από ένα σύνολο τέτοιων διανυσμάτων τάξης 0. Στις περιπτώσεις όπου η μέθοδος συγκλίνει, η σύγκλιση είναι κυβική με την έννοια ότι αν λ_J είναι μια ιδιοτιμή του A και το $u^{(0)}$ είναι αρκετά κοντά σε ένα ιδιοδιάνυσμα q_J τότε:

$$\left\|u^{(k+1)} - \left(\pm q_J\right)\right\| = O\left(\left\| u^{(k)} - \left(\pm q_J\right)\right\|^3\right) \tag{4.6}$$

και

$$\left|\lambda^{(k+1)} - \lambda_J\right| = O\left(\left|\lambda^{(k)} - \lambda_J\right|^3\right) \tag{4.7}$$

όταν $k \to \infty$. Το σύμβολο $\pm$ δεν είναι αναγκαστικά το ίδιο και στα δύο μέλη της (4.6).

Απόδειξη

Δεν θα αποδείξουμε τον ισχυρισμό σχετικά με τη σύγκλιση όλων σχεδόν των αρχικών διανυσμάτων, αλλά θα δείξουμε ότι αν η μέθοδος συγκλίνει, τότε συγκλίνει κυβικά. Για να κάνουμε πιο απλή την απόδειξη, υποθέτουμε ότι η ιδιοτιμή λ_J είναι απλή. Από τη σχέση (4.3), αν $\left\|u^{(k)} - q_J\right\| \leq \epsilon$ για αρκετά μικρό ϵ, τότε το Πηλίκο του Rayleigh δίνει μία εκτίμηση της ιδιοτιμής $\lambda^{(k)}$ με $\left|\lambda^{(k)} - \lambda_J\right| = O(\epsilon^2)$. Χρησιμοποιώντας παρόμοια επιχειρήματα με αυτά που χρησιμοποιήσαμε για να αποδείξουμε τα προηγούμενα, αν πάρουμε ένα βήμα της μεθόδου της Αντίστροφης Επανάληψης για να βρούμε μια νέα εκτίμηση ιδιοδιανύσματος $u^{(k+1)}$από τα $u^{(k)}$ και $\lambda^{(k)}$, τότε

$$\left\|u^{(k+1)} - q_J\right\| = O\left(\left|\lambda^{(k)} - \lambda_J\right| \left\|u^{(k)} - q_J\right\|\right) = O(\epsilon^3).$$

Ακόμα οι σταθεροί όροι που βρίσκονται στα σύμβολα O είναι ομοιόμορφοι για αρκετά μικρές γειτονιές του λ_J και q_J. Έτσι έχουμε σύγκλιση στο ακόλουθο σχήμα:

$$\begin{array}{ccc} \|u^{(k)} - (\pm q_J)\| & & |\lambda^{(k)} - \lambda_J| \\ \epsilon & \rightarrow & O(\epsilon^2) \\ \downarrow & \swarrow & \\ O(\epsilon^3) & \rightarrow & O(\epsilon^6) \\ \downarrow & \swarrow & \\ O(\epsilon^9) & \rightarrow & O(\epsilon^{18}) \\ \vdots & & \vdots \end{array}$$

Σύμφωνα λοιπόν με το παραπάνω σχήμα έχουμε ότι αν το $\|u^{(k)} - (\pm q_J)\|$ είναι της τάξης του ϵ, τότε το $|\lambda^{(k)} - \lambda_J|$ θα είναι της τάξης του ϵ^2. Στο επόμενο βήμα παρατηρούμε ότι αν το $\|u^{(k+1)} - (\pm q_J)\|$ είναι της τάξης του ϵ^3, τότε το $|\lambda^{(k+1)} - \lambda_J|$ θα είναι της τάξης του ϵ^6 κ.ο.κ.

Οι εκτιμήσεις (4.6)-(4.7) απορρέουν από την ομοιομορφία που μόλις αναφέραμε.

Παράδειγμα 4.1. Η κυβική σύγκλιση είναι τόσο γρήγορη που θα έπρεπε να δώσουμε ένα αριθμητικό παράδειγμα ώστε να δούμε την ταχύτητα της. Θεωρούμε τον συμμετρικό πίνακα

$$A = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 4 \end{bmatrix},$$

και έστω $u^{(0)} = (1,1,1)^T/\sqrt{3}$ το αρχικό εκτιμούμενο ιδιοδιάνυσμα. Όταν εφαρμόζουμε την Επαναληπτική Μέθοδο του Πηλίκου του Rayleigh στον πίνακα A, οι ακόλουθες τιμές τις ιδιοτιμής $\lambda^{(k)}$ υπολογίζονται στις τρεις πρώτες επαναλήψεις:

$$\lambda^{(0)} = 5, \quad \lambda^{(1)} = 5{,}2131 \ldots, \quad \lambda^{(2)} = 5.21439743184 \ldots.$$

Η πραγματική τιμή της ιδιοτιμής που αντιστοιχεί στο ιδιοδιάνυσμα που είναι πιο κοντά στο $u^{(0)}$ είναι $\lambda = 5.214319743377$. Μετά από τρείς μόνο επαναλήψεις, η Επαναληπτική Μέθοδος του Πηλίκου Rayleigh έχει δώσει αποτέλεσμα με ακρίβεια δέκα ψηφίων. Τρείς ακόμα επαναλήψεις της μεθόδου θα μπορούσαν να αυξήσουν την τιμή αυτή σε περίπου 270 ψηφία, αν η ακρίβεια της μηχανής μας ήταν αρκετά υψηλή.

4.6. Μέτρηση Πράξεων

Θα ολοκληρώσουμε αυτό το κεφάλαιο με μια αναφορά στο πλήθος των πράξεων που απαιτούνται για να εκτελέσουμε κάθε βήμα των τριών αριθμητικών μεθόδων που περιγράψαμε.

Αρχικά υποθέτουμε ότι ο $A \in \mathbb{R}^{m\times m}$ είναι ένας πυκνός πίνακας. Τότε κάθε βήμα της Επαναληπτικής Μεθόδου των Δυνάμεων περιέχει έναν πολλαπλασιασμό πίνακα-διανύσματος, που απαιτεί $O(m^2)$ πράξεις. Κάθε βήμα της Μεθόδου της Αντίστροφης Επανάληψης περιλαμβάνει τη λύση ενός γραμμικού συστήματος, η οποία απαιτεί $O(m^3)$ πράξεις, αλλά αυτός ο αριθμός των πράξεων μειώνεται σε $O(m^2)$ αν ο πίνακας υποβάλλεται σε παραγοντοποίηση LU ή QR ή έχει χρησιμοποιηθεί κάποια άλλη μέθοδος. Στην περίπτωση της Επαναληπτικής Μεθόδου του Πηλίκου Rayleigh, ο πίνακας που πρέπει να αντιστραφεί αλλάζει σε κάθε βήμα, καθώς αλλάζει η εκτίμηση της ιδιοτιμής. Συνεπώς σε αυτή τη μέθοδο δεν είναι εύκολο να μειώσουμε τις $O(m^3)$ πράξεις που απαιτεί .

Αυτός ο αριθμός των πράξεων βελτιώνεται σημαντικά αν ο A είναι τριδιαγώνιος. Σε αυτή την περίπτωση, και οι τρείς μέθοδοι απαιτούν μόλις $O(m)$ πράξεις σε κάθε βήμα. Ακόμη, αν έχουμε μη συμμετρικούς πίνακες, πρέπει να ασχοληθούμε με πίνακες Hessenberg αντί με τριδιαγώνιες δομές, οπότε ο απαιτούμενος αριθμός πράξεων σε κάθε βήμα αυξάνεται σε $O(m^2)$.

ΚΕΦΑΛΑΙΟ 5

Αλγόριθμος QR Χωρίς Εναλλαγές

Ο Αλγόριθμος QR, ο οποίος χρονολογείται από τις αρχές του 1960 είναι ένας από τους πιο σημαντικούς αλγορίθμους της Αριθμητικής Ανάλυσης. Εδώ θα δείξουμε ότι στην απλούστερη μορφή του αυτός ο αλγόριθμος μπορεί να θεωρηθεί ως μια σταθερή διαδικασία για τον υπολογισμό παραγοντοποιήσεων QR των δυνάμεων ενός πίνακα $A, A^2, A^3, \ldots$

5.1. Ο Αλγόριθμος QR

Η πιο βασική μορφή του Αλγορίθμου QR φαίνεται ιδιαίτερα απλή.

Αλγόριθμος 5.1. Απλός αλγόριθμος QR

$A^{(0)} = A$
για $k = 1,2, \ldots$
  $Q^{(k)}R^{(k)} = A^{(k-1)}$ παραγοντοποίηση QR του $A^{(k-1)}$
  $A^{(k)} = R^{(k)}Q^{(k)}$ αναδιατάσσει τους παράγοντες $Q^{(k)}$και $R^{(k)}$με αντίστροφη σειρά

Το μόνο που κάνουμε είναι να βρούμε μια παραγοντοποίηση QR ενός πίνακα και να πολλαπλασιάσουμε, τους υπολογισμένους με την αντίστροφη διάταξη, παράγοντες Q και R. Κατόπιν επαναλαμβάνουμε τη διαδικασία αυτή. Κάτω από κατάλληλες προϋποθέσεις, αυτός ο απλός αλγόριθμος συγκλίνει σε ένα πίνακα A μορφής Schur (άνω τριγωνικός αν ο A είναι τυχαίος και διαγώνιος αν ο A είναι ερμιτιανός). Στο σημείο αυτό για να εξακολουθήσουμε να μιλάμε με απλό τρόπο, θα πρέπει να συνεχίσουμε να υποθέτουμε ότι ο A είναι πραγματικός και συμμετρικός πίνακας, με πραγματικές ιδιοτιμές λ_j και ορθοκανονικά ιδιοδιανύσματα q_j. Έτσι το ενδιαφέρον μας επικεντρώνεται στη σύγκλιση της ακολουθίας πινάκων $A^{(k)}$σε διαγώνια μορφή.

Για να είναι χρήσιμη η σύγκλιση σε διαγώνια μορφή, δηλαδή να μας δίνει τις ιδιοτιμές του αρχικού πίνακα, οι διαδικασίες που πραγματοποιούμε πρέπει να είναι αποτέλεσμα μετασχηματισμών ομοιότητας. Αυτό μπορεί εύκολα να αποδειχθεί: Από τον τύπο $Q^{(k)}R^{(k)} = A^{(k-1)}$, πολλαπλασιάζοντας από τα αριστερά με τον $\left(Q^{(k)}\right)^T$, (Αφού ο $Q^{(k)}$ είναι ορθοκανονικός) προκύπτει $R^{(k)} = \left(Q^{(k)}\right)^T A^{(k-1)}$. Στη συνέχεια, πολλαπλασιάζοντας από τα δεξιά με τον όρο $Q^{(k)}$, παίρνουμε τη σχέση $A^{(k)} = \left(Q^{(k)}\right)^T A^{(k-1)} Q^{(k)}$ (η οποία είναι μετασχηματισμός ομοιότητας), καθώς σύμφωνα με τον αλγόριθμο ισχύει $A^{(k)} = R^{(k)}Q^{(k)}$.

Στην πραγματικότητα έχουμε δει ξανά αυτόν τον μετασχηματισμό ομοιότητας στο Κεφάλαιο 3 όπου είχε αναφερθεί ως η κακή ιδέα, και παρόλο που αυτός ο μετασχηματισμός είναι πράγματι κακή ιδέα όταν προσπαθούμε να μετατρέψουμε τον A σε μια τριγωνική μορφή με ένα μόνο βήμα, αποδεικνύεται ότι είναι αρκετά ισχυρός, όπως η βάση μιας επανάληψης.

Όπως και η Επαναληπτική Μέθοδος του Πηλίκου Rayleigh, έτσι και ο Αλγόριθμος QR για πραγματικούς και συμμετρικούς πίνακες συγκλίνει κυβικά. Για να επιτευχθεί αυτό όμως, ο αλγόριθμος που παρουσιάστηκε παραπάνω, θα πρέπει να τροποποιηθεί με την εισαγωγή εναλλαγών σε κάθε βήμα. Η χρήση των εναλλαγών είναι μια από τις τρείς τροποποιήσεις του Αλγορίθμου 5.1 που είναι απαραίτητες προκειμένου να τον μετατρέψουμε σε ένα πρακτικό αλγόριθμο:

1. Πριν ξεκινήσει η επανάληψη, ο A τριδιαγωνιοποιείται όπως αναφέρθηκε στο προηγούμενο κεφάλαιο.
2. Αντί του $A^{(k)}$, ένας άλλος πίνακας $A^{(k)} - \mu^{(k)}I$ παραγοντοποιείται σε κάθε βήμα όπου το $\mu^{(k)}$ είναι μια προσέγγιση ιδιοτιμής.
3. Όποτε είναι δυνατόν και ιδίως όταν έχει βρεθεί μια ιδιοτιμή, το πρόβλημα διαχωρίζεται σε μικρότερα προβλήματα σπάζοντας τον $A^{(k)}$ σε υποπίνακες.

Ένας Αλγόριθμος QR

Αλγόριθμος 5.2. "Πρακτικός" Αλγόριθμος QR

$(Q^0)^T A^{(0)} Q^{(0)} = A$ Το $A^{(0)}$είναι μια τριδιαγωνιοποίηση του A

για $k = 1,2,\ldots$

    επέλεξε ένα $\mu^{(k)}$ π.χ επέλεξε $\mu^{(k)} = A_{mm}^{(k-1)}$

    $Q^{(k)}R^{(k)} = A^{(k-1)} - \mu^{(k)}I$ παραγοντοποίηση QR του $A^{(k-1)} - \mu^{(k)}I$

    $A^{(k)} = R^{(k)}Q^{(k)} + \mu^{(k)}I$ αναδιατάσσει παράγοντες με αντίστροφη σειρά

    Αν κάποιο μη διαγώνιο στοιχείο $A_{j,j+1}^{(k)}$ είναι αρκετά κοντά στο μηδέν,

θέτουμε $A_{j,j+1} = A_{j+1,j} = 0$ ώστε να πάρουμε

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} = A^{(k)}$$

Εφαρμόζουμε τον αλγόριθμο QR για το A_1 και A_2

Η τριδιαγωνιοποίηση έχει αναφερθεί σε προηγούμενο κεφάλαιο (Κεφ. 3), και οι εναλλαγές θα αναφερθούν στο επόμενο. Τώρα θα επικεντρώσουμε την προσοχή μας στον Απλό Αλγόριθμο QR και θα εξηγήσουμε το πώς αυτός βρίσκει τις ιδιοτιμές του πίνακα.

5.2. Μη Κανονικοποιημένες Ταυτόχρονες Επαναλήψεις.

Η προσέγγισή μας γίνεται με σκοπό να συσχετίσουμε τον Αλγόριθμο QR με μία άλλη μέθοδο, τη Μέθοδο των Ταυτόχρονων Επαναλήψεων, η συμπεριφορά της οποίας είναι πιο προφανής.

Η ιδέα της Μεθόδου των Ταυτόχρονων Επαναλήψεων είναι να εφαρμόσει τη Μέθοδο των Δυνάμεων σε περισσότερα από ένα διανύσματα σε κάθε βήμα (Ένας ισοδύναμος όρος είναι η Μέθοδος των Δυνάμεων σε block). Υποθέτουμε ότι ξεκινάμε με ένα σύνολο n γραμμικών ανεξαρτήτων διανυσμάτων $u_1^{(0)}, \dots, u_n^{(0)}$. Φαίνεται λογικό ότι όπως το $A^k u_1^{(0)}$ συγκλίνει καθώς $k \to \infty$ (κάτω από τις κατάλληλες προϋποθέσεις) σε ένα ιδιοδιάνυσμα το οποίο αντιστοιχεί στη μεγαλύτερη κατά μέτρο τιμή ιδιοτιμή του πίνακα A , έτσι και ο χώρος $< A^k u_1^{(0)}, \dots, A^k u_n^{(0)} >$ θα έπρεπε να συγκλίνει (ξανά υπό τις κατάλληλες προϋποθέσεις) στο χώρο $< q_1, \dots, q_n >$ που παράγεται από τα ιδιοδιανύσματα $q_1, \dots, q_n$ του A που αντιστοιχούν στις n μεγαλύτερες κατά μέτρο ιδιοτιμές του A.

Σε συμβολισμό πινάκων, θα μπορούσαμε να το προσεγγίσουμε ως εξής:

Ορίζουμε $V^{(0)}$ να είναι ο $m \times n$ αρχικός πίνακας.

$$V^{(0)} = \left[u_1^{(0)} \mid \dots \mid \quad u_n^{(0)}\right], \tag{5.1}$$

και ορίζουμε $V^{(k)}$ να είναι το αποτέλεσμα μετά από k εφαρμογές του A στον $V^{(0)}$.

$$V^{(k)} = A^k V^{(0)} = \left[u_1^{(k)} \mid \dots \mid \quad u_n^{(k)}\right]. \tag{5.2}$$

Καθώς το ενδιαφέρον μας είναι στο χώρο στηλών του $V^{(k)}$, θα εξάγουμε μια βάση αυτού του χώρου υπολογίζοντας μια παραγοντοποίηση QR του $V^{(k)}$.

$$\hat{Q}^{(k)} \hat{R}^{(k)} = V^{(k)}. \tag{5.3}$$

Εδώ οι $\hat{Q}^{(k)}$ και $\hat{R}^{(k)}$ έχουν διαστάσεις $m \times n$ και $n \times n$, αντίστοιχα. Φαίνεται εύλογο ότι καθώς $k \to \infty$, υπό κατάλληλες προϋποθέσεις, οι διαδοχικές στήλες του $\hat{Q}^{(k)}$ συγκλίνουν στα ιδιοδιανύσματα $\pm q_1, \pm q_2, \dots, \pm q_n$.

Αυτό μπορεί να πραγματοποιηθεί με μια ανάλογη ανάλυση όπως και στο προηγούμενο κεφάλαιο. Αν εκφράσουμε τα $u_j^{(0)}$ και $u_j^{(k)}$ ως γραμμικούς συνδυασμούς των ιδιοδιανυσμάτων του A, τότε έχουμε:

$$u_j^{(0)} = a_{1j} q_1 + \cdots + a_{mj} q_m,$$

$$u_j^{(k)} = \lambda_1^k a_{1j} q_1 + \cdots + \lambda_m^k a_{mj} q_m.$$

Όπως και πριν έχουμε απλή σύγκλιση αν ισχύουν οι προϋποθέσεις. Η πρώτη προϋπόθεση είναι ότι οι $n+1$ μεγαλύτερες ιδιοτιμές, έχουν διακριτές απόλυτες τιμές

$$|\lambda_1| > |\lambda_2| > \cdots > |\lambda_n| > |\lambda_{n+1}| \geq |\lambda_{n+2}| \geq \cdots \geq |\lambda_m| \,. \tag{5.4}$$

Η δεύτερη προϋπόθεση είναι ότι ο πίνακας που συνθέτει το σύνολο των συντελεστών a_{ij} είναι αντιστρέψιμος. Ορίζουμε $\hat{Q}$ να είναι ένας $m \times n$ πίνακας του οποίου οι στήλες είναι τα ιδιοδιανύσματα $q_1, q_2, \dots, q_n$ (δηλαδή ο $\hat{Q}$, πίνακας των ιδιοδιανυσμάτων, είναι εντελώς διαφορετικός από τους $\hat{Q}^{(k)}$, που παράγονται από την παραγοντοποίηση QR). Υποθέτουμε τα ακόλουθα:

$$\text{Όλοι οι κύριοι υποπίνακες του } \hat{Q}^T V^{(0)} \text{ είναι αντιστρέψιμοι.} \tag{5.5}$$

Με τον όρο κύριοι υποπίνακες του $\hat{Q}^T V^{(0)}$, εννοούμε τους πάνω-αριστερούς τετραγωνικούς υποπίνακες διαστάσεων $1 \times 1, 2 \times 2, \dots, n \times n$ (η συνθήκη (5.5) είναι ισοδύναμη με την συνθήκη ο $\hat{Q}^T V^{(0)}$ να δέχεται παραγοντοποίηση LU).

Θεώρημα 5.1. Υποθέτουμε ότι η επανάληψη που περιγράφεται από τις σχέσεις (5.1)-(5.3) εκτελείται, και ότι οι προϋποθέσεις (5.4) και (5.5) ικανοποιούνται. Τότε καθώς $k \to \infty$, οι στήλες των πινάκων $\hat{Q}^{(k)}$ συγκλίνουν γραμμικά στα ιδιοδιανύσματα του A:

$$\left\| q_j^{(k)} - (\pm q_j) \right\| = O(C^k) \tag{5.6}$$

για κάθε j με $1 \leq j \leq n$, όπου το $C < 1$ είναι η σταθερά $max_{1 \leq k \leq n} |\lambda_{k+1}|/|\lambda_k|$. Όπως και στα θεωρήματα του προηγούμενου κεφαλαίου, το πρόσημο $\pm$ σημαίνει ότι σε κάθε βήμα k μπορούμε να επιλέξουμε το ένα ή το άλλο πρόσημο και τα φράγματα να ισχύουν.

Απόδειξη

Επεκτείνουμε τον $\hat{Q}$ σε έναν γεμάτο $m \times m$ ορθογώνιο πίνακα Q ιδιοδιανυσμάτων του πίνακα A και έστω Λ ο αντίστοιχος διαγώνιος πίνακας των ιδιοτιμών. Έτσι $A = Q\Lambda Q^T$. Όπως ο $\hat{Q}$ είναι το κύριο $m \times n$ τμήμα του Q, ορίζουμε τον $\hat{\Lambda}$ (επίσης διαγώνιο) να είναι το κύριο $n \times n$ τμήμα του Λ. Τότε έχουμε

$$V^{(k)} = A^k V^{(0)} = Q\Lambda^k Q^T V^{(0)} = \hat{Q}\hat{\Lambda}^k \hat{Q}^T V^{(0)} + O(|\lambda_{n+1}|^k)$$

καθώς $k \to \infty$. Εάν η (5.5) ισχύει, τότε ειδικότερα ο $\hat{Q}^T V^{(0)}$ είναι αντιστρέψιμος, έτσι μπορούμε να πολλαπλασιάσουμε τον όρο $O(|\lambda_{n+1}|^k)$ από τα δεξιά με $\left(\hat{Q}^T V^{(0)}\right)^{-1} \hat{Q}^T V^{(0)}$ και να μετατρέψουμε την εξίσωση σε

$$V^{(k)} = \left[\hat{Q}\hat{\Lambda}^k + O(|\lambda_{n+1}|^k)\left(\hat{Q}^T V^{(0)}\right)^{-1}\right] \hat{Q}^T V^{(0)} = \left(\hat{Q}\hat{\Lambda}^k + O(|\lambda_{n+1}|^k)\right) \hat{Q}^T V^{(0)}.$$

Αφού ο $\hat{Q}^T V^{(0)}$ είναι αντιστρέψιμος, ο χώρος των στηλών αυτού του πίνακα είναι ο ίδιος με τον χώρο στηλών του πίνακα

$$\hat{Q}\hat{\Lambda}^k + O(|\lambda_{n+1}|^k).$$

Από τη μορφή του $\hat{Q}\hat{\Lambda}^k$ και την προϋπόθεση (5.4) είναι ξεκάθαρο ότι ο χώρος των στηλών του συγκλίνει γραμμικά στο χώρο στηλών του $\hat{Q}$.

Στην πραγματικότητα, έχουμε υποθέσει ότι όχι μόνο ο πίνακας $\hat{Q}^T V^{(0)}$ είναι αντιστρέψιμος, αλλά ότι όλοι οι κύριοι υποπίνακές του είναι αντιστρέψιμοι. Επίσης το ίδιο ισχύει και για τα κύρια υποσύνολα στηλών του $V^{(k)}$ και του $\hat{Q}$: οι πρώτες στήλες, οι πρώτες και οι δεύτερες στήλες, οι πρώτες, οι δεύτερες και οι τρίτες στήλες κ.ο.κ. Σε κάθε περίπτωση καταλήγουμε στο ότι ο χώρος που παράγεται από τις συγκεκριμένες στήλες του $V^{(k)}$ συγκλίνει γραμμικά στο χώρο που παράγεται από τις αντίστοιχες στήλες του $\hat{Q}$. Από τη σύγκλιση όλων των διαδοχικών χώρων στηλών, μαζί με τον ορισμό της παραγοντοποίησης QR (5.3), (5.6) προκύπτει το ζητούμενο. □

5.3. Ταυτόχρονες Επαναλήψεις

Καθώς $k \to \infty$, τα διανύσματα $u_1^{(k)}, \dots, u_n^{(k)}$ του αλγορίθμου (5.1)-(5.3) συγκλίνουν όλα σε πολλαπλάσια του ιδιοδιανύσματος q_1 , το οποίο αντιστοιχεί στη μεγαλύτερη κατά μέτρο ιδιοτιμή του πίνακα A. Έτσι, αν και ο χώρος που παράγουν κάποια από αυτά $< u_1^{(k)}, \dots, u_j^{(k)} >$ συγκλίνει σε κάτι χρήσιμο, τα διανύσματα αυτά αποτελούν μια βάση του χώρου, κακής κατάστασης. Αν όντως πραγματοποιηθεί μια ταυτόχρονη επανάληψη με αριθμητική κινητής υποδιαστολής, η επιθυμητή πληροφορία θα χαθεί γρήγορα στα λάθη στρογγυλοποίησης.

Η λύση είναι απλή, συγκεκριμένα πρέπει να ορθοκανονικοποιούμε σε κάθε βήμα της επανάληψης και όχι μια φορά για όλα τα βήματα. Έτσι δεν πρέπει να κατασκευάσουμε τους $V^{(k)}$ όπως ορίστηκαν παραπάνω, αλλά θα κατασκευάσουμε μια διαφορετική ακολουθία πινάκων $Z^{(k)}$ οι οποίοι θα έχουν τον ίδιο χώρο στηλών.

Αλγόριθμος 5.3. Μέθοδος Ταυτόχρονων Επαναλήψεων

Πάρε $\hat{Q}^{(0)} \in \mathbb{R}^{m\times n}$ με ορθοκανονικές στήλες.

για $k = 1,2, \dots$

$Z = A\hat{Q}^{(k-1)}$

$\hat{Q}^{(k)}\hat{R}^{(k)} = Z$ γίνεται παραγοντοποίηση QR του Z

Από τη μορφή αυτού του αλγορίθμου, είναι φανερό ότι οι χώροι στηλών του $\hat{Q}^{(k)}$ και του $Z^{(k)}$ είναι οι ίδιοι, καθώς και οι δύο είναι ίσοι με το χώρο στηλών του $A^k\hat{Q}^{(0)}$. Έτσι, μιλώντας μαθηματικά, αυτή η νέα διατύπωση της Μεθόδου των Ταυτόχρονων Επαναλήψεων (ο νέος αλγόριθμος) συγκλίνει υπό τις ίδιες συνθήκες όπως και ο προηγούμενος.

Θεώρημα 5.2. Ο Αλγόριθμος 5.3 δημιουργεί τους ίδιους πίνακες $\hat{Q}^{(k)}$ όπως η επαναληπτική μέθοδος που περιγράφουν οι εξισώσεις (5.1)-(5.3) όπως στο Θεώρημα 5.1 (προϋποθέτοντας ότι οι αρχικοί πίνακες $\hat{Q}^{(0)}$ είναι οι ίδιοι) και κάτω από τις ίδιες υποθέσεις (5.4) και (5.5) συγκλίνει όπως περιγράφηκε σε εκείνο το θεώρημα.

5.4. Ισοδυναμία Μεθόδου Ταυτόχρονων Επαναλήψεων και Αλγορίθμου QR

Τώρα μπορούμε να εξηγήσουμε τον Αλγόριθμο QR. Αυτός είναι ισοδύναμος με τη Μέθοδο των Ταυτόχρονων Επαναλήψεων για ένα σύνολο $n = m$ αρχικών διανυσμάτων, που συνθέτουν ένα ταυτοτικό πίνακα $\hat{Q}^{(0)} = I$. Αφού οι πίνακες $\hat{Q}^{(k)}$ είναι τώρα τετραγωνικοί, έχουμε να κάνουμε με πλήρεις QR παραγοντοποιήσεις, οπότε μπορούμε να γράψουμε $Q^{(k)}$ και $R^{(k)}$ αντί για $\hat{Q}^{(k)}$ και $\hat{R}^{(k)}$. Τώρα όμως θα αντικαταστήσουμε τον συμβολισμό του πίνακα $\hat{R}^{(k)}$ με $R^{(k)}$, αλλά τον πίνακα $\hat{Q}^{(k)}$ θα τον συμβολίζουμε πλέον με $\underline{Q}^{(k)}$ προκειμένου να διακρίνουμε τους πινάκες Q που προκύπτουν από τη Μέθοδο των Ταυτοχρόνων Επαναλήψεων από αυτούς που προκύπτουν από τον Αλγόριθμο QR.

Εδώ υπάρχουν τρείς εξισώσεις που ορίζουν τη Μέθοδο των Ταυτόχρονων Επαναλήψεων με $\underline{Q}^{(0)} = I$, οι οποίες ακολουθούνται από μια τέταρτη εξίσωση την οποία θα χρησιμοποιήσουμε ως ορισμό του $m \times m$ πίνακα $A^{(k)}$:

Ταυτόχρονες επαναλήψεις:

$$\underline{Q}^{(0)} = I, \tag{5.7}$$

$$Z = A\underline{Q}^{(k-1)}, \tag{5.8}$$

$$Z = \underline{Q}^{(k)}R^{(k)}, \tag{5.9}$$

$$A^{(k)} = \left(\underline{Q}^{(k)}\right)^T A\underline{Q}^{(k)}. \tag{5.10}$$

Και εδώ υπάρχουν οι τρείς εξισώσεις που ορίζουν τον Απλό Αλγόριθμο QR και οι οποίες ακολουθούνται από μια τέταρτη εξίσωση, την οποία θα χρησιμοποιήσουμε ως ορισμό του $m \times m$ πίνακα $\underline{Q}^{(k)}$:

Απλός Αλγόριθμος QR χωρίς επαναλήψεις:

$$A^{(0)} = A, \tag{5.11}$$

$$A^{(k-1)} = Q^{(k)}R^{(k)}, \tag{5.12}$$

$$A^{(k)} = R^{(k)}Q^{(k)}, \tag{5.13}$$

$$\underline{Q}^{(k)} = Q^{(1)}Q^{(2)} \dots Q^{(k)}. \tag{5.14}$$

Επιπρόσθετα και για τους δυο αλγορίθμους, ορίζουμε επιπλέον ένα $m \times m$ πίνακα $\underline{R}^{(k)}$,

$$\underline{R}^{(k)} = R^{(k)}R^{(k-1)} \dots R^{(1)}. \tag{5.15}$$

Τώρα μπορούμε να δούμε την ισοδυναμία αυτών των δύο αλγορίθμων.

Θεώρημα 5.3. Οι μέθοδοι που περιγράφονται στις σχέσεις (5.7)-(5.10) και (5.11)-(5.14) δημιουργούν πανομοιότυπες ακολουθίες πινάκων $\underline{R}^{(k)}$, $\underline{Q}^{(k)}$ και $A^{(k)}$, δηλαδή αυτές ορίζονται από την παραγοντοποίηση QR της κ-οστής δύναμης του A.

$$A^{(k)} = \underline{Q}^{(k)}\underline{R}^{(k)}\,, \tag{5.16}$$

μαζί με τον υπολογισμό

$$A^{(k)} = \left(\underline{Q}^{(k)}\right)^T A\underline{Q}^{(k)}. \tag{5.17}$$

Απόδειξη

Προχωρούμε με επαγωγή ως προς τον k. Η βασική περίπτωση $k = 0$ είναι τετριμμένη. Για τη Μέθοδο των Ταυτόχρονων Επαναλήψεων και για τον QR Αλγόριθμο, ισχύουν οι εξισώσεις (5.7)-(5.15) στις οποίες θέτοντας το k ίσο με το μηδέν συνεπάγεται $A^0 = \underline{Q}^{(0)} = \underline{R}^{(0)} = I$ και $A^{(0)} = A$, απ’ όπου οι εξισώσεις (5.16) και (5.17) είναι προφανείς. Έστω τώρα ότι έχουμε την περίπτωση $k \geq 1$ για τη Μέθοδο των Ταυτόχρονων Επαναλήψεων. Η εξίσωση (5.17) είναι πανομοιότυπη με την (5.10), έτσι το μόνο που χρειαζόμαστε είναι να επαληθεύσουμε την σχέση (5.16) η οποία μπορεί να γίνει ως εξής:

$$\mathrm{A^k = A\underline{Q}^{(k-1)}\underline{R}^{(k-1)} = \underline{Q}^{(k)}R^{(k)}\underline{R}^{(k-1)} = \underline{Q}^{(k)}\underline{R}^{(k)}}.$$

Η πρώτη ισότητα προκύπτει από την επαγωγική υπόθεση (5.16), η δεύτερη από (5.8) και (5.9) και η τρίτη από την (5.15).

Από την άλλη μεριά, θεωρούμε για τον Αλγόριθμο QR την περίπτωση όπου $k \geq 1$. Μπορούμε να επαληθεύσουμε την (5.16) από την ακολουθία

$$\mathrm{A^k} = \mathrm{A}\underline{\mathrm{Q}}^{(\mathrm{k}-1)}\underline{\mathrm{R}}^{(\mathrm{k}-1)} = \underline{\mathrm{Q}}^{(\mathrm{k}-1)}A^{(\mathrm{k}-1)}\underline{\mathrm{R}}^{(\mathrm{k}-1)} = \underline{\mathrm{Q}}^{(\mathrm{k})}\underline{\mathrm{R}}^{(\mathrm{k})}.$$

Η πρώτη ισότητα προκύπτει από την επαγωγική υπόθεση (5.16), η δεύτερη από την επαγωγική υπόθεση (5.17) και η τρίτη προκύπτει από την (5.12) μαζί με την (5.14) και την (5.15). Τελικά μπορούμε να επαληθεύσουμε την (5.17) από την ακολουθία

$$\mathrm{A}^{(k)} = \left(\mathrm{Q}^{(\mathrm{k})}\right)^{\mathrm{T}} \mathrm{A}^{(\mathrm{k}-1)}\mathrm{Q}^{(\mathrm{k})} = \left(\underline{\mathrm{Q}}^{(\mathrm{k})}\right)^{\mathrm{T}} \mathrm{A}\underline{\mathrm{Q}}^{(\mathrm{k})}.$$

Η πρώτη ισότητα προκύπτει από την (5.12) και την (5.13) και η δεύτερη από την επαγωγική υπόθεση της (5.17). □

5.5. Σύγκλιση του Αλγορίθμου QR

Μπορούμε τώρα να μιλήσουμε για την σύγκλιση του αλγορίθμου QR.

Οι σχέσεις στις οποίες θα βασιστούμε προκειμένου να κατανοήσουμε την σύγκλιση του Αλγορίθμου QR είναι οι (5.16) και (5.17). Η πρώτη από αυτές μας δείχνει ότι ο Αλγόριθμος QR μπορεί να βρει τα ιδιοδιανύσματα του A κατασκευάζοντας ορθοκανονικές βάσεις διαδοχικών δυνάμεων $\mathrm{A^k}$ του πίνακα A. Η δεύτερη εξίσωση μας εξηγεί πως ο αλγόριθμος βρίσκει τις ιδιοτιμές. Συγκεκριμένα από αυτή την εξίσωση προκύπτει ότι τα διαγώνια στοιχεία του $\mathrm{A}^{(k)}$ είναι τα Πηλίκα του Rayleigh του A που αντιστοιχούν στις στήλες του $\underline{Q}^{(\mathrm{k})}$. Έτσι καθώς αυτές οι στήλες συγκλίνουν στα ιδιοδιανύσματα, τα Πηλίκα του Rayleigh συγκλίνουν στις αντίστοιχες ιδιοτιμές. Εν τω μεταξύ η σχέση (5.17) δείχνει ότι τα μη-διαγώνια στοιχεία του $\mathrm{A}^{(k)}$ αντιστοιχούν σε γενικευμένα Πηλίκα του Rayleigh που περιλαμβάνουν προσεγγίσεις διακριτών ιδιοδιανυσμάτων του πίνακα A. Αυτές οι προσεγγίσεις πρέπει να είναι ορθογώνιες, καθώς συγκλίνουν στα διακριτά ιδιοδιανύσματα, και άρα τα μη-διαγώνια στοιχεία του $\mathrm{A}^{(k)}$ πρέπει να συγκλίνουν στο 0.

Για μια πιο ποσοτική κατανόηση έχουμε το ακόλουθο θεώρημα.

Θεώρημα 5.4. Έστω ότι ο απλός Αλγόριθμος QR (αλγόριθμος 5.1) εφαρμόζεται σε ένα πραγματικό, συμμετρικό πίνακα A για του οποίου τις ιδιοτιμές ισχύει ότι $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_m|$ και του οποίου ο πίνακας των αντίστοιχων ιδιοδιανυσμάτων Q έχει όλους τους κύριους υποπίνακες αντιστρέψιμους. Τότε καθώς το $k \to \infty$, ο $\mathrm{A}^{(k)}$ συγκλίνει γραμμικά στον $diag(\lambda_1, \dots, \lambda_m)$ με σταθερά $max_j \left|\lambda_{j+1}\right| / \left|\lambda_j\right|$ και ο πίνακας $\underline{Q}^{(k)}$ (με πρόσημο των στηλών του όπως χρειάζεται κάθε φορά) συγκλίνει στον Q με τον ίδιο ρυθμό.

ΚΕΦΑΛΑΙΟ 6

Αλγόριθμος QR με Εναλλαγές

Αυτό που κάνει την Επαναληπτική Μέθοδο QR να είναι πολύ καλή είναι η εισαγωγή εναλλαγών $A \rightarrow A - \mu I$ σε κάθε βήμα της επανάληψης. Στο κεφάλαιο αυτό θα εξηγήσουμε πώς αυτή η ιδέα οδηγεί σε κυβική σύγκλιση της μεθόδου, χάρις σε μια έμμεση σύνδεση με την Επαναληπτική Μέθοδο του Πηλίκου Rayleigh.

6.1. Σύνδεση με Αντίστροφη Επανάληψη

Συνεχίζουμε να υποθέτουμε ότι ο $A \in \mathbb{R}^{m\times m}$ είναι πραγματικός και συμμετρικός πίνακας, με πραγματικές ιδιοτιμές, $\{\lambda_j\}$ και ορθοκανονικά ιδιοδιανύσματα $\{q_j\}$.

Όπως έχουμε δει, ο Απλός Αλγόριθμος QR (Αλγόριθμος 5.1) είναι ισοδύναμος με τη Μέθοδο των Ταυτόχρονων Επαναλήψεων εφαρμοζόμενη στον ταυτοτικό πίνακα. Συγκεκριμένα η πρώτη στήλη του αποτελέσματος προκύπτει σύμφωνα με την Επαναληπτική Μέθοδο των Δυνάμεων εφαρμοζόμενη στο διάνυσμα e_1. Υπάρχει και ένας δεύτερος τρόπος για να το παρατηρήσουμε αυτό. Ο Αλγόριθμος 5.1 είναι επίσης ισοδύναμος με τις Ταυτόχρονες Αντίστροφες Επαναλήψεις εφαρμοζόμενες σε έναν αλλαγμένο πίνακα P και συγκεκριμένα η $m - o\sigma\tau$ή στήλη του αποτελέσματος προκύπτει σύμφωνα με την Αντίστροφη Επανάληψη εφαρμοζόμενη στο διάνυσμα e_m.

Μπορούμε να αποδείξουμε αυτόν τον ισχυρισμό ως εξής:

Έστω $\mathrm{Q}^{(k)}$ να είναι όπως και στο προηγούμενο κεφάλαιο ο ορθογώνιος παράγοντας στο k-οστό βήμα του Αλγορίθμου QR. Στο προηγούμενο κεφάλαιο, δείξαμε ότι το γινόμενο (5.14) αυτών των πινάκων

$$\underline{Q}^{(k)} = \prod_{\mathrm{j}=1}^{\mathrm{k}} \mathrm{Q}^{(\mathrm{j})} = \left[\mathrm{q}_1^{(\mathrm{k})} \,\middle|\, \mathrm{q}_2^{(\mathrm{k})} \,\middle|\, \dots \,\middle|\, \quad \mathrm{q}_\mathrm{m}^{(\mathrm{k})}\right],$$

είναι ο ίδιος ορθογώνιος πίνακας ο οποίος εμφανίζεται στο k βήμα (5.9) της Μεθόδου των Ταυτόχρονων Επαναλήψεων. Ένας άλλος τρόπος να το δούμε αυτό είναι να πούμε ότι ο $\underline{Q}^{(k)}$ είναι ο ορθογώνιος παράγοντας σε μια παραγοντοποίηση QR (5.16)

$$A^{(k)} = \underline{Q}^{(k)}\underline{R}^{(k)}.$$

Τώρα εξετάζουμε τι συμβαίνει αν αντιστρέψουμε αυτόν τον τύπο. Υπολογίζουμε

$$A^{-k} = \left(\underline{R}^{(k)}\right)^{-1}\underline{Q}^{(k)T} = \underline{Q}^{(k)}\left(\underline{R}^{(k)}\right)^{-T}, \tag{6.1}$$

για τη δεύτερη ισότητα έχουμε χρησιμοποιήσει το γεγονός ότι ο A^{-1} είναι συμμετρικός. Έστω ότι ο P ορίζει ένα $m \times m$ πίνακα μετάθεσης ο οποίος αντιστρέφει τη σειρά των γραμμών και των στηλών.

$$P = \begin{bmatrix} & & & 1 \\ & & 1 & \\ & \cdots & & \\ 1 & & & \end{bmatrix}.$$

Αφού $P^2 = I$, η (6.1) μπορεί να γραφεί ως εξής:

$$A^{-k}P = \left[\underline{Q}^{(k)}P\right]\left[P\left(\underline{R}^{(k)}\right)^{-T}P\right]. \tag{6.2}$$

Ο πρώτος παράγοντας σε αυτό το γινόμενο, ο $\underline{Q}^{(k)}P$ είναι ορθογώνιος. Ο δεύτερος παράγοντας $P\left(\underline{R}^{(k)}\right)^{-T}P$ είναι άνω τριγωνικός (ξεκινά με τον άνω τριγωνικό πίνακα $\left(\underline{R}^{(k)}\right)^{-T}$ και εφαρμόζοντας τον P από δεξιά γίνεται κάτω τριγωνικός, ενώ αν εφαρμοστεί από αριστερά, γίνεται πάλι άνω τριγωνικός). Έτσι η (6.2) μπορεί να ερμηνευτεί ως μία παραγοντοποίηση QR του $A^{-k}P$. Με λίγα λόγια εκτελούμε αποτελεσματικά την Μέθοδο των Ταυτόχρονων Επαναλήψεων στον A^{-1} εφαρμοζόμενο στον αρχικό πίνακα P, την οποία μέθοδο καλούμε Μέθοδο των Ταυτόχρονων Αντίστροφων Επαναλήψεων του A. Συγκεκριμένα η πρώτη στήλη του $\underline{Q}^{(k)}P$, που είναι η τελευταία στήλη του $\underline{Q}^{(k)}$, είναι το αποτέλεσμα της εφαρμογής k-βημάτων από τη Μέθοδο της Αντίστροφης Επανάληψης στο διάνυσμα e_m.

6.2. Σύνδεση με Εναλλασσόμενη Αντίστροφη Επανάληψη

Επομένως ο Αλγόριθμος QR είναι παράλληλα και Ταυτόχρονες Επαναλήψεις και Ταυτόχρονες Αντίστροφες Επαναλήψεις: η συμμετρία είναι τέλεια. Ωστόσο όπως είδαμε στο Κεφάλαιο 4 υπάρχει μια μεγάλη διαφορά μεταξύ της Μεθόδου των Δυνάμεων και της Αντίστροφης Επανάληψης. Συγκεκριμένα η τελευταία μπορεί να προκύψει με τη χρήση αυθαίρετων εναλλαγών. Όσο καλύτερα μπορούμε να εκτιμήσουμε μια ιδιοτιμή $\mu \approx \lambda_J$, τόσο γρηγορότερα θα επιτύχουμε καλύτερες εκτιμήσεις με ένα βήμα της Αντίστροφης Επανάληψης χρησιμοποιώντας τον πίνακα εναλλαγών $A - \mu I$. Ο " Πρακτικός " Αλγόριθμος δείχνει το πως οι εναλλαγές εισάγονται στα βήματα του Αλγορίθμου QR. Αυτό είναι ακριβώς η εισαγωγή εναλλαγών στις αντίστοιχες μεθόδους των Ταυτόχρονων Επαναλήψεων και των Αντίστροφων Επαναλήψεων, και άρα η επίδραση της εισαγωγής αυτών των εναλλαγών είναι ακριβώς η ίδια.

Έστω ότι $\mu^{(k)}$ συμβολίζει την εκτίμηση της ιδιοτιμής που επιλέγουμε στο k βήμα του Αλγορίθμου QR. Στον " Πρακτικό " Αλγόριθμο (με τις εναλλαγές) η σχέση μεταξύ των βημάτων $k-1$ και k είναι

$$A^{(k-1)} - \mu^{(k)} I = Q^{(k)} R^{(k)},$$

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

Αυτό συνεπάγεται

$$A^{(k)} = \left(Q^{(k)}\right)^T A^{(k-1)} Q^{(k)}, \tag{6.3}$$

και επαγωγικά προκύπτει ότι

$$A^{(k)} = \left(\underline{Q}^{(k)}\right)^T A \underline{Q}^{(k)},$$

η οποία είναι η αμετάβλητη σχέση (5.17). Όμως η (5.16) δεν ισχύει πλέον. Αντί για αυτή, έχουμε την παραγοντοποίηση

$$\left(A - \mu^{(k)} I\right)\left(A - \mu^{(k-1)} I\right) \dots \left(A - \mu^{(1)} I\right) = \underline{Q}^{(k)} \underline{R}^{(k)}, \tag{6.4}$$

η οποία είναι μια παραλλαγή με εναλλαγές της Μεθόδου των Ταυτόχρονων Επαναλήψεων. Δηλαδή το γινόμενο $\underline{Q}^{(k)} = \prod_{\mathrm{j}=1}^{\mathrm{k}} \mathrm{Q}^{(\mathrm{j})}$ είναι μια ορθογωνοποίηση του πίνακα $\prod_{j=k}^{1}\left(A - \mu^{(j)} I\right)$. Η πρώτη στήλη του $\underline{Q}^{(k)}$ είναι το αποτέλεσμα της εφαρμογής της Μεθόδου της Εναλλασσόμενης Επαναληπτικής Μεθόδου των Δυνάμεων στο διάνυσμα e_1 χρησιμοποιώντας τις εναλλαγές $\mu^{(j)}$, και η τελευταία στήλη είναι το αποτέλεσμα της εφαρμογής k βημάτων της Εναλλασσόμενης Αντίστροφης Επανάληψης στο διάνυσμα e_m με τις ίδιες εναλλαγές. Αν οι εναλλαγές είναι καλές εκτιμήσεις των ιδιοτιμών, τότε η τελευταία στήλη του πίνακα $\underline{Q}^{(k)}$ συγκλίνει σε ένα ιδιοδιάνυσμα.

6.3. Σύνδεση με την Επαναληπτική Μέθοδο του Πηλίκου Rayleigh

Έχουμε ανακαλύψει ένα ισχυρό εργαλείο που κρύβεται στον Εναλλασσόμενο Αλγόριθμο QR, αυτό είναι η Μέθοδος της Εναλλασσόμενης Αντίστροφης Επανάληψης. Για να ολοκληρώσουμε την ιδέα αυτή, χρειαζόμαστε ένα τρόπο με τον οποίο θα επιλέγουμε τις εναλλαγές $\mu^{(k)}$ με σκοπό να επιτύχουμε γρήγορη σύγκλιση στην τελευταία στήλη του $\underline{Q}^{(k)}$.

Στην αρχή προκειμένου να εκτιμήσουμε την ιδιοτιμή που αντιστοιχεί στο ιδιοδιάνυσμα που προσεγγίζεται από την τελευταία στήλη του πίνακα $\underline{Q}^{(k)}$, θα χρησιμοποιήσουμε στη στήλη αυτή το Πηλίκο Rayleigh. Αυτό μας δίνει

$$\mu^{(k)} = \frac{\left(q_m^{(k)}\right)^T A q_m^{(k)}}{\left(q_m^{(k)}\right)^T q_m^{(k)}} = \left(q_m^{(k)}\right)^T A q_m^{(k)}. \tag{6.5}$$

Αν αυτός ο αριθμός που επιλέχθηκε παραμένει σταθερός καθ' όλη τη διάρκεια των επαναλήψεων, τότε οι εκτιμήσεις των ιδιοτιμών $\mu^{(k)}$ και των ιδιοδιανυσμάτων $q_m^{(k)}$ είναι πανομοιότυπες με εκείνες που υπολογίστηκαν με τη χρήση της Επαναληπτικής Μεθόδου του Πηλίκου Rayleigh που ξεκινά με το διάνυσμα e_m. Άρα ο Αλγόριθμος QR συγκλίνει κυβικά με την έννοια ότι το διάνυσμα $q_m^{(k)}$ συγκλίνει κυβικά σε ένα ιδιοδιάνυσμα.

Σημειώνουμε ότι στον Αλγόριθμο QR, το Πηλίκο Rayleigh $r\left(q_m^{(k)}\right)$ εμφανίζεται ως το στοιχείο της m γραμμής και m στήλης του πίνακα $A^{(k)}$. Το αναφέραμε αυτό στο τέλος του προηγούμενου κεφαλαίου. Ξεκινώντας με την σχέση (6.3) έχουμε

$$\mathrm{A}_{\mathrm{mm}}^{(\mathrm{k})} = \mathrm{e}_{\mathrm{m}}^{\mathrm{T}} \mathrm{A}^{(\mathrm{k})} \mathrm{e}_{\mathrm{m}} = \mathrm{e}_{\mathrm{m}}^{\mathrm{T}} \underline{\mathrm{Q}}^{(\mathrm{k})\mathrm{T}} \mathrm{A} \underline{\mathrm{Q}}^{(\mathrm{k})} \mathrm{e}_{\mathrm{m}} = \mathrm{q}_{\mathrm{m}}^{(\mathrm{k})\mathrm{T}} \mathrm{A} \mathrm{q}_{\mathrm{m}}^{(\mathrm{k})}. \tag{6.6}$$

Άρα η (6.5) είναι ίδια καθώς θέτουμε $\mu^{(k)} = \mathrm{A}_{\mathrm{mm}}^{(\mathrm{k})}$. Αυτό είναι γνωστό ως η εναλλαγή του Πηλίκου Rayleigh.

6.4. Εναλλαγές Wilkinson

Παρόλο που οι εναλλαγές του Πηλίκου Rayleigh συγκλίνουν κυβικά στη γενική περίπτωση, η σύγκλιση δεν είναι εγγυημένη για όλες τις αρχικές συνθήκες. Αυτό μπορεί να γίνει αντιληπτό με ένα απλό παράδειγμα. Θεωρούμε τον πίνακα

$$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}. \tag{6.7}$$

Ο Αλγόριθμος QR χωρίς εναλλαγές δεν συγκλίνει

$$A = Q^{(1)} R^{(1)} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix},$$

$$A^{(1)} = R^{(1)} Q^{(1)} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = A.$$

Η εναλλαγή του Πηλίκου Rayleigh $\mu = A_{mm}$, δεν έχει ούτε αυτή αποτέλεσμα, καθώς $A_{mm} = 0$. Έτσι είναι ξεκάθαρο ότι στη χειρότερη περίπτωση, ο Αλγόριθμος QR με την εναλλαγή του Πηλίκου Rayleigh μπορεί να αποτύχει.

Το πρόβλημα προκύπτει λόγω της συμμετρίας των ιδιοτιμών. Η μια ιδιοτιμή είναι το +1 και η άλλη το -1, έτσι όταν προσπαθούμε να βελτιώσουμε την εκτίμηση της ιδιοτιμής $\mu = 0$, αυτή τείνει το ίδιο και στις δύο ιδιοτιμές και έτσι η εκτίμηση δεν βελτιώνεται. Αυτό που χρειάζεται είναι μια εκτίμηση ιδιοτιμής που μπορεί να σπάσει τη συμμετρία. Για να βρούμε μια τέτοια ιδιοτιμή, κάνουμε το εξής: Ορίζουμε ο B να είναι ο 2×2 κάτω δεξιά υποπίνακας του $A^{(k)}$.

$$B = \begin{bmatrix} a_{m-1} & b_{m-1} \\ b_{m-1} & a_m \end{bmatrix}.$$

Η εναλλαγή Wilkinson ορίζεται ως η ιδιοτιμή του πίνακα B η οποία είναι πιο κοντά στο a_m. Στη περίπτωση που κάποιες ιδιοτιμές είναι εξίσου κοντά τότε επιλέγουμε τυχαία μια από τις αυτές.

Ένας αριθμητικά σταθερός τύπος για την εναλλαγή Wilkinson είναι

$$\mu = \mathrm{a_m} - \mathrm{sing}(\delta)\mathrm{b}_{\mathrm{m-1}}^2 / \left(|\delta\,| + \sqrt{\delta^2 + \mathrm{b}_{\mathrm{m-1}}^2}\right), \tag{6.8}$$

όπου $\delta = (a_{m-1} - a_m)/2$. Αν $\delta = 0$, τότε το πρόσημο του δ μπορεί να θεωρηθεί τυχαία ίσο με 1 ή με -1.

Όπως η εναλλαγή του Πηλίκου Rayleigh, έτσι και η εναλλαγή Wilkinson επιτυγχάνει κυβική σύγκλιση στη γενική περίπτωση. Επιπλέον μπορεί να δειχθεί ότι στη χειρότερη περίπτωση μπορεί να επιτύχει τουλάχιστον τετραγωνική σύγκλιση. Στην πράξη ο Αλγόριθμος QR με την εναλλαγή Wilkinson πάντα συγκλίνει.

Στο παράδειγμα (6.7) η εναλλαγή Wilkinson είναι είτε +1 είτε -1. Έτσι η συμμετρία χαλάει και η σύγκλιση πραγματοποιείται σε ένα βήμα.

6.5. Ευστάθεια και Ακρίβεια

Εδώ ολοκληρώνουμε τη συζήτησή μας για τους μηχανισμούς του Αλγορίθμου QR, αν και πολλές πρακτικές λεπτομέρειες έχουν παραλειφθεί. Μένει να πούμε λίγα λόγια για τη ευστάθεια και την ακρίβεια.

Όπως μπορεί κανείς να περιμένει από τη χρήση ορθογώνιων πινάκων στον Αλγόριθμο QR, η μέθοδος είναι backward ευσταθής. Όπως και σε προηγούμενο κεφάλαιο, ο ευκολότερος τρόπος να διαμορφώσουμε ένα τύπο για το αποτέλεσμα είναι να ορίσουμε ως $\tilde{\Lambda}$ τον πίνακα που είναι διαγωνοποίηση του A όπως υπολογίζεται με απλής ακρίβειας αριθμητική και ως $\tilde{Q}$ τον ορθογώνιο πίνακα που συνδέεται με το γινόμενο όλων των αριθμητικά υπολογισμένων ανακλαστών Householder (ή περιστροφών Givens) που χρησιμοποιήθηκαν κατά την πορεία. Το ακόλουθο θεώρημα μπορεί να αποδειχθεί.

Θεώρημα 6.1. Έστω ένας πραγματικός, συμμετρικός και τριδιαγώνιος πίνακας $A \in \mathbb{R}^{m \times m}$ που έχει διαγωνιοποιηθεί από τον " Πρακτικό " Αλγόριθμο QR (Αλγόριθμος 5.2) σε υπολογιστή που ικανοποιεί τα (3.1) και (3.2) αξιώματα και έστω $\tilde{\Lambda}$ και $\tilde{Q}$ να είναι όπως ορίστηκαν παραπάνω. Τότε έχουμε

$$\tilde{Q}\tilde{\Lambda}\tilde{\mathrm{Q}}^{*} = \mathrm{A} + \delta A, \qquad \frac{\|\delta A\|}{\|A\|} = O(\epsilon_{\mu\eta\chi\alpha\nu\acute{\eta}\varsigma}) \tag{6.9}$$

για κάποιο $\delta A \in \mathbb{C}^{m\times m}$.

Ο Αλγόριθμος QR δίνει μια ακριβή λύση σε ένα ελάχιστα διαταραγμένο πρόβλημα. Συνδυάζοντας τα Θεωρήματα 3.1 και 6.1 παρατηρούμε ότι η απλοποίηση σε τριδιαγώνια μορφή που προκύπτει από τον Αλγόριθμο QR είναι ένας backward ευσταθής αλγόριθμος υπολογισμού ιδιοτιμών των πινάκων. Για να δούμε τι συνεπάγεται αυτό για την ακρίβεια των υπολογισμένων ιδιοτιμών, συνδυάζουμε αυτό το συμπέρασμα με το αποτέλεσμα (3.4) όσον αφορά την διαταραχή των ιδιοτιμών των πραγματικών και συμμετρικών πινάκων (μια ειδική περίπτωση κανονικών πινάκων). Το συμπέρασμα είναι ότι οι υπολογισμένες ιδιοτιμές $\tilde{\lambda}_j$ ικανοποιούν

$$\frac{|\tilde{\lambda}_j - \lambda_j|}{\|A\|} = O\big(\epsilon_{\mu\eta\chi\alpha\nu\acute{\eta}\varsigma}\big). \tag{6.10}$$

Αυτό δεν είναι καθόλου κακό αποτέλεσμα για έναν αλγόριθμο ο οποίος απαιτεί μόλις $\sim\frac{4}{3}m^3$ πράξεις, δηλαδή τα $\frac{2}{3}$ των πράξεων που απαιτούνται για τον υπολογισμό του γινομένου ενός ζεύγους $m \times m$ πινάκων.

ΚΕΦΑΛΑΙΟ 7

Άλλοι Αλγόριθμοι εύρεσης ιδιοτιμών

Εκτός από τον Αλγόριθμο QR υπάρχουν και άλλοι αλγόριθμοι εύρεσης ιδιοτιμών. Σε αυτό το κεφάλαιο θα αναφερθούν τρείς δημοφιλείς αλγόριθμοι για πραγματικά, συμμετρικά προβλήματα ιδιοτιμών. Αυτοί οι αλγόριθμοι είναι, ο Αλγόριθμος Jacobi για τους κανονικούς πίνακες, η Διχοτόμηση και Αλγόριθμοι Διαίρει και Βασίλευε για τους τριδιαγώνιους πίνακες.

7.1. Jacobi

Μία από τις παλαιότερες ιδέες για των υπολογισμό ιδιοτιμών πινάκων είναι ο Αλγόριθμος Jacobi που θεσπίστηκε από τον Jacobi το 1845. Είναι μία μέθοδος η οποία κατάφερε να προσελκύσει την προσοχή από την εποχή που εμφανίστηκαν οι υπολογιστές, και ιδιαίτερα μετά την έλευση του παράλληλου προγραμματισμού.

Η ιδέα είναι η εξής: Για πίνακες διαστάσεως 5 ή μεγαλύτερους, γνωρίζουμε ότι οι ιδιοτιμές μπορούν να αποκτηθούν μόνο με αναδρομή (Κεφάλαιο 2), ωστόσο για τους πίνακες που είναι μικρότερης διάστασης μπορούμε να βρούμε τις ιδιοτιμές τους σε ένα μόνο βήμα. Θα μπορούσαμε λοιπόν να διαγωνοποιήσουμε ένα μικρό υποπίνακα του A, μετά έναν άλλον και ούτω καθεξής με σκοπό να συγκλίνουν τελικά σε μια διαγωνοποίηση του κανονικού πίνακα.

Η ιδέα αυτή έχει δοκιμαστεί για υποπίνακες 4×4, αλλά συνήθως βασίζεται σε υποπίνακες 2×2 . Ένας 2×2 πραγματικός συμμετρικός πίνακας μπορεί να διαγωνοποιηθεί στη μορφή

$$J^T \begin{bmatrix} a & d \\ d & b \end{bmatrix} J = \begin{bmatrix} \neq 0 & 0 \\ 0 & \neq 0 \end{bmatrix}, \tag{7.1}$$

όπου ο J είναι ένας ορθογώνιος πίνακας. Υπάρχουν αρκετοί τρόποι για να επιλέξουμε τον J. Ένας από αυτούς είναι να θεωρήσουμε ότι είναι ένας 2×2 ανακλαστής Householder της μορφής

$$F = \begin{bmatrix} -c & s \\ s & c \end{bmatrix}, \tag{7.2}$$

όπου $s = \sin\theta$ και $c = \cos\theta$ για κάποιο θ. Σημειώνουμε ότι $detF = -1$ το οποίο αποτελεί το χαρακτηριστικό γνώρισμα ενός ανακλαστή. Εναλλακτικά θα μπορούσε κάποιος να πάρει τον πίνακα J όχι ως ένα ανακλαστή αλλά ως ένα πίνακα περιστροφής,

$$J = \begin{bmatrix} c & s \\ -s & c \end{bmatrix}, \tag{7.3}$$

με $detJ = 1$. Αυτή είναι η τυπική προσέγγιση του Αλγορίθμου Jacobi. Μπορεί να αποδειχθεί ότι η διαγωνιοποίηση (7.1) επιτυγχάνεται αν το θ ικανοποιεί τη σχέση

$$\tan 2\theta = \frac{2d}{b-a}, \quad (7.4)$$

και ο πίνακας J ο οποίος βασίζεται σε αυτή την επιλογή καλείται περιστροφή Jacobi.

Έστω ότι ο $A \in \mathbb{R}^{m\times m}$ είναι ένας συμμετρικός πίνακας. Ο Αλγόριθμος Jacobi συνίσταται από την αναδρομική εφαρμογή των μετασχηματισμών (7.1) βασιζόμενη στους πίνακες που ορίστηκαν από τις σχέσεις (7.3) και (7.4). Ο πίνακας J είναι τώρα επαυξημένος σε έναν $m \times m$ πίνακα που όλα τα στοιχεία του εκτός από τέσσερα, είναι ίσα με τη μονάδα, και έχει τη μορφή του πίνακα της σχέσης (7.3). Εφαρμόζοντας τον J^T από τα αριστερά του A μεταβάλλονται οι δύο σειρές του πίνακα A, κατόπιν εφαρμόζοντας τον J από τα δεξιά μεταβάλλονται οι δύο στήλες του A. Σε κάθε βήμα της επανάληψης ένα συμμετρικό ζεύγος μηδενικών εισάγεται στο πίνακα αλλά τα προηγούμενα μηδενικά καταστρέφονται. Όπως και με τον Αλγόριθμο QR, το σύνηθες αποτέλεσμα είναι τα πλήθος αυτών των μη μηδενικών στοιχείων να μειώνεται σταθερά.

Όσο αφορά το ποιά μη διαγώνια στοιχεία a_{ij} πρέπει να μηδενίζονται σε κάθε βήμα της επανάληψης, η απάντηση είναι ότι πρέπει να επιλέγουμε το μεγαλύτερο μη διαγώνιο στοιχείο σε κάθε βήμα. Σε αυτή την περίπτωση η ανάλυση της σύγκλισης γίνεται προφανής για εκείνον που μπορεί να αποδείξει ότι το άθροισμα των τετραγώνων των μη διαγώνιων στοιχείων μειώνεται τουλάχιστον κατά ένα παράγοντα $1 - 2/(m^2 - m)$ σε κάθε βήμα. Μετά από $O(m^2)$ βήματα, καθένα από τα οποία απαιτεί $O(m)$ πράξεις, το άθροισμα των τετραγώνων πρέπει να μειωθεί κατά ένα σταθερό παράγοντα, και η μέθοδος συγκλίνει με ακρίβεια μηχανής $\epsilon_{machine}$ μετά από $O\left(m^3 \log(\epsilon_{μηχανής})\right)$ πράξεις. Στην πράξη, έχει παρατηρηθεί ότι η σύγκλιση είναι καλύτερη από την παραπάνω (τελικά τετραγωνική παρά γραμμική), ώστε το σύνολο των πράξεων που απαιτείται να είναι $O\left(m^3 \log(\left|\log(\epsilon_{μηχανής}\right|)\right)$.

Σε έναν υπολογιστή, τα μη διαγώνια στοιχεία απαλείφονται με ένα κυκλικό τρόπο ο οποίος αποφεύγει τις $O(m^2)$ πράξεις που εκτελούνται με σκοπό την αναζήτηση του μεγαλυτέρου μη διαγώνιου στοιχείου. Για παράδειγμα αν τα $m(m-1)/2$ υπερδιαγώνια στοιχεία απαλείφονται κατά γραμμή, ξεκινώντας με τα στοιχεία $a_{12}, a_{13}, \ldots$, η ταχεία ασυμπτωτική σύγκλιση είναι και πάλι εγγυημένη. Μετά από μια σάρωση του 2×2 πίνακα που περιέχει όλα τα $m(m-1)/2$ ζευγάρια των μη διαγώνιων στοιχείων, η ακρίβεια της σύγκλισης έχει βελτιωθεί σε γενικές γραμμές και η σύγκλιση είναι τελικά τετραγωνική.

Η Μέθοδος Jacobi έχει ενδιαφέρον διότι ασχολείται μόνο με ζεύγη σειρών και στηλών κάθε φορά, κάνοντάς την εύκολα παραλληλοποιήσιμη . Ο πίνακας σε αυτή την περίπτωση δεν είναι από πριν τριδιαγωνοποιημένος, καθώς οι επαναλήψεις της μεθόδου Jacobi θα κατέστρεφαν αυτή τη δομή. Τυπικά η σύγκλιση πινάκων διαστάσεως $m \leq 1000$ μπορεί να επιτευχθεί σε λιγότερες από 10 σαρώσεις, και η τελική ακρίβεια ως προς τις συνιστώσες είναι σε γενικές γραμμές ακόμη καλύτερη απ' ότι μπορεί να επιτευχθεί από τον Αλγόριθμο QR. Ωστόσο, ο Αλγόριθμος Jacobi δεν είναι συνήθως τόσο γρήγορος όσο η τριδιαγωνιοποίηση που προκύπτει από τον Αλγόριθμο QR ή τον Αλγόριθμο του Διαίρει και Βασίλευε.

7.2. Διχοτόμηση

Ο επόμενος αλγόριθμος υπολογισμού ιδιοτιμών, είναι η Μέθοδος της Διχοτόμησης, η οποία έχει μεγάλη πρακτική σημασία. Αφού λοιπόν ένας συμμετρικός πίνακας τριδιαγωνιοποιηθεί, η Μέθοδος της Διχοτόμησης αποτελεί το επόμενο βήμα σε περίπτωση που κάποιος δεν θέλει όλες τις ιδιοτιμές του πίνακα, αλλά ένα υποσύνολο αυτών. Για παράδειγμα η Μέθοδος της Διχοτόμησης μπορεί να βρει το 10% των μεγαλύτερων ιδιοτιμών, ή τις 30 μικρότερες ιδιοτιμές ενός πίνακα, ή όλες τις ιδιοτιμές που υπάρχουν στο διάστημα [1,2]. Μόλις βρεθούν οι επιθυμητές ιδιοτιμές, τα αντίστοιχα ιδιοδιανύσματα μπορούν να βρεθούν από ένα βήμα της Αντίστροφης Επανάληψης (Αλγόριθμος 4.2).

Το σημείο έναρξης είναι βασικό. Αφού οι ιδιοτιμές ενός πραγματικού και συμμετρικού πίνακα είναι πραγματικές, μπορούμε να τις βρούμε ψάχνοντας τις πραγματικές ρίζες του πολυωνύμου $p(x) = \det(A - xI)$. Αυτό ακούγεται σαν μια κακή ιδέα, καθώς στο Κεφάλαιο 2 αναφέραμε ότι η εύρεση ριζών πολυωνύμων είναι μια ιδιαίτερα ασταθής διαδικασία για την εύρεση ιδιοτιμών. Η διαφορά εδώ είναι ότι αυτές οι παρατηρήσεις αφορούσαν την εύρεση ριζών μέσω του υπολογισμού των πολυωνυμικών συντελεστών του χαρακτηριστικού πολυωνύμου. Τώρα όμως η ιδέα είναι να βρούμε τις ρίζες εκτιμώντας το πολυώνυμο $p(x)$ σε διάφορα σημεία x, χωρίς να ασχολούμαστε με τους συντελεστές του και να εφαρμόσουμε τη συνήθη διαδικασία διχοτόμησης για μη γραμμικές συναρτήσεις. Αυτό μπορεί να γίνει, για παράδειγμα, με Γκαουσιανή απαλοιφή με οδήγηση, και ο αλγόριθμος που θα προκύψει θα είναι πολύ σταθερός.

Αυτό που δίνει στη Μέθοδο της Διχοτόμησης τη δύναμή της, είναι μερικές επιπρόσθετες ιδιότητες των ιδιοτιμών και των οριζουσών οι οποίες δεν είναι άμεσα προφανής.

Δίνεται ένας συμμετρικός πίνακας $A \in \mathbb{R}^{m\times m}$, έστω ότι οι $A^{(1)}, \dots, A^{(m)}$ υποδηλώνουν τους κύριους (άνω- αριστερά) τετραγωνικούς υποπίνακες διατάσεως $1, \dots, m$. Μπορεί να διχτεί ότι οι ιδιοτιμές αυτών των πινάκων είναι αυστηρά διατεταγμένες. Πριν ορίσουμε αυτή την ιδιότητα, ας υποθέσουμε ότι ο A είναι τριδιαγώνιος και μη αναγώγιμος υπό την έννοια ότι όλα τα μη διαγώνια στοιχεία του είναι μη μηδενικά:

$$A=\begin{bmatrix} a_1 & b_1 & & & \\ b_1 & a_2 & b_2 & & \\ & b_2 & a_3 & \ddots & \\ & & \ddots & \ddots & b_{m-1} \\ & & & b_{m-1} & a_m \end{bmatrix}, \quad b_j \neq 0\ . \tag{7.5}$$

(Αν υπάρχουν μηδενικά μη διαγώνια στοιχεία, τότε το πρόβλημα ιδιοτιμών μπορεί να διαχωριστεί σε μικρότερα προβλήματα, όπως στον " Πρακτικό " Αλγόριθμο QR.)

Έστω ότι για τις ιδιοτιμές του $A^{(k)}$ ισχύει

$$\lambda_1^{(k)} < \lambda_2^{(k)} < \cdots < \lambda_k^{(k)}.$$

Η ιδιότητα στην οποία οφείλει η Μέθοδος της Διχοτόμησης τη δύναμή της είναι ότι αυτές οι ιδιοτιμές είναι αυστηρά διατεταγμένες, ικανοποιώντας τις ανισότητες

$$\lambda_j^{(k+1)} < \lambda_j^{(k)} < \lambda_{j+1}^{(k+1)}. \tag{7.6}$$

Για $k=1,2,\dots,m$ και για $j=1,2,\dots,k-1$.

Ολοκληρώνοντας την περιγραφή του Αλγορίθμου της Διχοτόμησης παρατηρούμε ότι για ένα τριδιαγώνιο πίνακα, οι ορίζουσες των πινάκων $\{A^{(k)}\}$ σχετίζονται με μία αναδρομική σχέση η οποία αποτελείται από τρείς όρους. Επεκτείνοντας το $\det\left(\mathrm{A}^{(k)}\right)$ κατά τα μικρότερα σε σχέση με τα στοιχεία του b_{k-1} και a_k στη γραμμή k, δίνει από την (7.5),

$$\det\left(\mathrm{A}^{(k)}\right) = a_k \det\left(\mathrm{A}^{(k-1)}\right) - \mathrm{b}_{\mathrm{k}-1}^2 \det\left(\mathrm{A}^{(k-2)}\right). \tag{7.7}$$

Εισάγοντας τη μετατόπιση κατά xI και γράφοντας $p^{(k)}(x)=\det\left(A^{(k)}-xI\right)$, παίρνουμε

$$p^{(k)}(x) = (a_k - x)p^{(k-1)}(x) - \mathrm{b}_{\mathrm{k}-1}^2 \mathrm{p}^{(\mathrm{k}-2)}(\mathrm{x}). \tag{7.8}$$

Αν ορίσουμε $p^{(-1)}(x)=0$ και $p^{(0)}(x)=1$, τότε αυτή η αναδρομή ισχύει για όλα τα $k=1,2,\dots,m$.

Εφαρμόζοντας την σχέση (7.8) για μια διαδοχική σειρά τιμών του x και μετρώντας τις αλλαγές των προσήμων κατά την πορεία, ο αλγόριθμος της Μεθόδου της Διχοτόμησης τοποθετεί τις ιδιοτιμές σε μικρά τυχαία διαστήματα. Το κόστος είναι $O(m)$ πράξεις για κάθε εκτίμηση της ακολουθίας, ως εκ τούτου $O\left(mlog\left(\epsilon_{\mu\eta\chi\alpha\nu\acute{\eta}\varsigma}\right)\right)$ πράξεις συνολικά για να βρούμε μια ιδιοτιμή με σχετική ακρίβεια $\epsilon_{\mu\eta\chi\alpha\nu\acute{\eta}\varsigma}$. Αν χρειάζεται ένας μικρός αριθμός ιδιοτιμών, αυτή είναι μια διακριτή βελτίωση έναντι των $O(m^2)$ πράξεων που απαιτούνται για τον Αλγόριθμο QR.

7.3. Μέθοδος Διαίρει και Βασίλευε

Ο αλγόριθμος της Μεθόδου Διαίρει και Βασίλευε, ο οποίος εισήχθηκε για πρώτη φορά από τον Cuppen το 1981, παρουσίασε την πιο σημαντική πρόοδο στους αλγορίθμους εύρεσης ιδιοτιμών πίνακα από το 1960. Ο λόγος είναι ότι βασίζεται σε μια αναδρομική υποδιαίρεση ενός συμμετρικού τριδιαγώνιου προβλήματος ιδιοτιμών σε προβλήματα μικρότερων διαστάσεων. Σε περίπτωση που ζητούνται τα ιδιοδιανύσματα καθώς και οι ιδιοτιμές ενός πίνακα, αυτή η μέθοδος είναι περισσότερο από δύο φορές πιο γρήγορη απ' ότι ο Αλγόριθμος QR. Έστω $T \in \mathbb{R}^{m \times m}$ με $m \geq 2$ ένας συμμετρικός, τριδιαγώνιος και μη αναγώγιμος πίνακας, δηλαδή ένας πίνακας που έχει μόνο μη μηδενικά μη διαγώνια στοιχεία. Τότε για κάθε n στην περιοχή $1 \leq n < m$, ο πίνακας T μπορεί να διασπαστεί σε δύο υποπίνακες ως εξής:

$$T = \begin{bmatrix} T_1 & \beta \\ \beta & T_2 \end{bmatrix} = \begin{bmatrix} \widehat{T_1} & \\ & \widehat{T_2} \end{bmatrix} + \begin{bmatrix} \beta & \beta \\ \beta & \beta \end{bmatrix}. \tag{7.9}$$

Εδώ ο T_1 είναι ο άνω-αριστερά $n \times n$ κύριος υποπίνακας του T, ο T_2 είναι ο κάτω-δεξιά $(m-n) \times (m-n)$ κύριος υποπίνακας, και $\beta = t_{n+1,n} = t_{n,n+1} \neq 0.$ Η μόνη διαφορά μεταξύ του T_1 και του $\widehat{T_1}$ είναι ότι το κάτω-δεξιά στοιχείο t_{nn} έχει αντικατασταθεί από το $t_{nn} - \beta$, και η μόνη διαφορά μεταξύ του T_2 και του $\widehat{T_2}$ είναι ότι το άνω-αριστερά στοιχείο $t_{n+1,n+1}$ έχει αντικατασταθεί από το $t_{n+1,n+1} - \beta$. Αυτές οι τροποποιήσεις των δύο στοιχείων έγιναν προκειμένου ο δεξιότερος πίνακας της σχέσης (7.9) να είναι τάξης ένα.

Στο σημείο αυτό θα εκφράσουμε τη σχέση (7.9) με λέξεις. *Ένας τριδιαγώνιος πίνακας μπορεί να γραφτεί ως το άθροισμα ενός* 2×2 *block διαγώνιου πίνακα με τριδιαγώνια blocks και μια τάξης ένα διόρθωση.*

Ο Αλγόριθμος του Διαίρει και Βασίλευε ενεργεί ως εξής. Διαιρεί τον πίνακα T όπως στη σχέση (7.9) με $n \approx m/2$. Υποθέτουμε ότι οι ιδιοτιμές των $\widehat{T_1}$ και $\widehat{T_2}$ είναι γνωστές. Δεδομένου ότι η διόρθωση του πίνακα είναι τάξης ένα, ένας μη γραμμικός αλλά γρήγορος υπολογισμός μπορεί να χρησιμοποιηθεί για να πάρει από τις ιδιοτιμές των $\widehat{T_1}$ και $\widehat{T_2}$ εκείνες του T. Τώρα θα επαναλάβουμε αυτή την ιδέα, βρίσκοντας τις ιδιοτιμές των $\widehat{T_1}$ και $\widehat{T_2}$ με περαιτέρω υποδιαιρέσεις, με τάξης-ένα διορθώσεις, και ούτω καθεξής. Με αυτόν τον τρόπο ένα $m \times m$ πρόβλημα ιδιοτιμών ανάγεται σε ένα σύνολο 1×1 προβλημάτων ιδιοτιμών μαζί με μια συλλογή από τάξης-ένα διορθώσεις. (Στην πράξη, για μέγιστη αποτελεσματικότητα, είναι σύνηθες να στρεφόμαστε στον Αλγόριθμο QR όταν οι υποπίνακες είναι επαρκώς μικρής διάστασης.)

Σε αυτή τη διαδικασία υπάρχει ένα βασικό σημείο. Αν οι ιδιοτιμές των $\widehat{T_1}$ και $\widehat{T_2}$ είναι γνωστές, τότε προκύπτει το ερώτημα για το πως μπορούν να βρεθούν οι ιδιοτιμές του T. Για να απαντήσουμε σε αυτή την ερώτηση, υποθέτουμε ότι οι διαγωνοποιήσεις

$$\widehat{T}_1 = Q_1 D_1 Q_1^T, \qquad \widehat{T}_2 = Q_2 D_2 Q_2^T$$

έχουν υπολογιστεί. Τότε από τη σχέση (7.9) προκύπτει ότι

$$T = \begin{bmatrix} Q_1 & \\ & Q_2 \end{bmatrix} \left(\begin{bmatrix} D_1 & \\ & D_2 \end{bmatrix} + \beta z z^T \right) \begin{bmatrix} Q_1^T & \\ & Q_2^T \end{bmatrix}. \tag{7.10}$$

Με $z^T = (q_1^T, q_2^T)$, όπου το q_1^T είναι η τελευταία γραμμή του Q_1 και το q_2^T είναι η πρώτη γραμμή του Q_2. Καθώς αυτή η εξίσωση είναι μετασχηματισμός ομοιότητας, έχουμε περιορίσει το μαθηματικό πρόβλημα σε ένα πρόβλημα εύρεσης των ιδιοτιμών ενός διαγώνιου πίνακα σύν μια τάξης –ένα διόρθωση.

Για να δείξουμε πως γίνεται αυτό, απλοποιούμε τον συμβολισμό ως εξής. Υποθέτουμε ότι θέλουμε να βρούμε τις ιδιοτιμές του $D + ww^T$, όπου ο $D \in \mathbb{R}^{m \times m}$ είναι ένας διαγώνιος πίνακας με διακριτά διαγώνια στοιχεία $\{d_j\}$ και το $w \in \mathbb{R}^m$είναι ένα διάνυσμα. (Η επιλογή του πρόσημου συν αντιστοιχεί στο $\beta > 0$, για $\beta < 0$ θα πρέπει να θεωρήσουμε $D - ww^T$.) Μπορούμε να υποθέσουμε ότι ισχύει $w_j \neq 0$ για κάθε j, διαφορετικά το πρόβλημα είναι αναγώγιμο. Τότε οι ιδιοτιμές του $D + ww^T$είναι οι ρίζες της ρητής συνάρτησης

$$\mathrm{f}(\lambda) = 1 + \sum_{j=1}^{m} \frac{w_j^2}{d_j - \lambda}. \tag{7.11}$$

Ο ισχυρισμός αυτός μπορεί να δικαιολογηθεί σημειώνοντας ότι αν ισχύει ($D + ww^T)q = \lambda q$ για κάποια $q \neq 0$, τότε $(D - \lambda I)q + w(w^T q) = 0$, το οποίο συνεπάγεται ότι $q + (D - \lambda I)^{-1} w(w^T q) = 0$, και στη συνέχεια πολλαπλασιάζοντας με w^T προκύπτει $w^T q + w^T (D - \lambda I)^{-1} w(w^T q) = 0$. Αυτό ισοδυναμεί με την εξίσωση $\mathrm{f}(\lambda)(w^T q) = 0$, στην οποία το $w^T q$ πρέπει να είναι μη μηδενικό, διαφορετικά το q θα είναι ένα ιδιοδιάνυσμα του D , και ως εκ τούτου είναι μη μηδενική μόνο μία θέση, πράγμα που σημαίνει τελικά ότι $w^T q \neq 0$. Καταλήγουμε στο συμπέρασμα ότι αν το q είναι ένα ιδιοδιάνυσμα του $D + ww^T$ με ιδιοτιμή λ, τότε το $\mathrm{f}(\lambda)$ πρέπει να είναι μηδέν και το αντίστροφο προκύπτει από τη μορφή του $\mathrm{f}(\lambda)$ η οποία εγγυάται ότι αυτό έχει ακριβώς m μηδενικά. Η εξίσωση $\mathrm{f}(\lambda) = 0$ είναι γνωστή ως χαρακτηριστική εξίσωση.

Σε κάθε αναδρομικό βήμα του Αλγορίθμου Διαίρει και Βασίλευε, οι ρίζες της εξίσωσης (7.11) βρίσκονται με μια τυχαία επαναληπτική διαδικασία σχετική με τη μέθοδο του Newton. Μόνο $O(1)$ επαναλήψεις απαιτούνται για κάθε ρίζα (ή $O\left(log\left(\left|\log\left(\epsilon_{\mu\eta\chi\alpha\nu\eta\varsigma}\right)\right|\right)\right)$ επαναλήψεις αν το $\epsilon_{machine}$ θεωρηθεί ως μια μεταβλητή), καθιστώντας το σύνολο των πράξεων σε $O(m)$ πράξεις για κάθε ρίζα ενός $m \times m$ πίνακα ή $O(m^2)$ πράξεις συνολικά. Αν φανταστούμε μια αναδρομή κατά την οποία ένας πίνακας διάστασης m διασπάται στο μισό σε κάθε βήμα, το σύνολο των πράξεων που χρειάζεται για την εύρεση των ιδιοτιμών ενός τριδιαγώνιου πίνακα με τον Αλγόριθμο Διαίρει και Βασίλευε είναι

$$\mathrm{O}(\mathrm{m}^2 + 2\left(\frac{\mathrm{m}}{2}\right)^2 + 4\left(\frac{\mathrm{m}}{4}\right)^2 + 8\left(\frac{\mathrm{m}}{8}\right)^2 + \cdots + \mathrm{m}\left(\frac{\mathrm{m}}{\mathrm{m}}\right)^2), \qquad (7.12)$$

η σειρά αυτή συγκλίνει στο $\mathrm{O}(\mathrm{m}^2)$ (όχι σε $\mathrm{O}(\mathrm{m}^2\mathrm{logm})$) εξαιτίας των τετραγώνων στους παρονομαστές. Έτσι το σύνολο των πράξεων που απαιτούνται φαίνεται να είναι της ίδιας τάξης $O(m^2)$ όπως και στον Αλγόριθμο QR.

Μέχρι στιγμής, δεν είναι ξεκάθαρο γιατί ο Αλγόριθμος του Διαίρει και Βασίλευε είναι πλεονεκτικός. Δεδομένου ότι η αναγωγή ενός πλήρους πίνακα σε τριδιαγώνια μορφή (φάση 1 στην εισαγωγή του Κεφαλαίου 2) απαιτεί $4m^3/3$ πράξεις (3.2), φαίνεται ότι οποιαδήποτε μείωση $O(m^2)$ στις πράξεις που απαιτεί για τη διαγωνοποίηση αυτού του τριδιαγώνιου πίνακα (φάση 2) είναι πολύ σημαντική. Όμως η οικονομία στις πράξεις αλλάζει αν κάποιος υπολογίσει τα ιδιοδιανύσματα καθώς και τις ιδιοτιμές. Σε αυτή την περίπτωση η φάση 1 απαιτεί $8m^3/3$ πράξεις και η φάση 2 $O(m^3)$ πράξεις.

Ο πολλαπλασιασμός του Q_j με και Q_j^T στη σχέση (7.10) απαιτεί $O(m^3)$ πράξεις. Το σύνολο των πράξεων που απαιτείται για όλα τα βήματα της αναδρομής, είναι $4m^3/3$ πράξεις. Προσθέτοντας στις $8m^3/3$ πράξεις της φάσης 1 ο αριθμός των πράξεων μειώνεται από $\approx 9m^3$ σε $4m^3$.

Στην πραγματικότητα, ο Αλγόριθμος Διαίρει και Βασίλευε, μπορεί να τα καταφέρει ακόμα καλύτερα με το θέμα των πράξεων, για κάποιο λόγο που δεν είναι προφανής. Για τους περισσότερους πίνακες A, πολλά από τα διανύσματα z και τους πίνακες Q_j που βρίσκονται στη σχέση (7.10) αποδεικνύεται ότι είναι αριθμητικά αραιοί, υπό την έννοια ότι πολλά από τα στοιχεία τους έχουν σχετικά μεγέθη μικρότερα από την ακρίβεια της μηχανής. Αυτή η αραιότητα επιτρέπει τη διαδικασία μιας αριθμητικής μείωσης, μέσω της οποίας τριδιαγώνια προβλήματα ιδιοτιμών μετατρέπονται σε μη περίπλοκα προβλήματα μικρότερων διαστάσεων. Έτσι, αυτό έχει σαν αποτέλεσμα σε τυπικές περιπτώσεις το σύνολο των πράξεων που απαιτούνται για τη φάση 2 να είναι λιγότερο από m^3 πράξεις, ενώ το σύνολο των πράξεων που απαιτείται για τη φάση 1 και 2 μειώνεται σε $\frac{8}{3}m^3$ πράξεις συνολικά. Για τον υπολογισμό των ιδιοτιμών και μόνο, σύμφωνα με τη σχέση (7.12) το σύνολο των πράξεων στη φάση 2 είναι λιγότερο από m^2 πράξεις.

ΚΕΦΑΛΑΙΟ 8

Παραδείγματα με MATLAB.

8.1 Μέθοδος των Δυνάμεων.

Η Μέθοδος των Δυνάμεων είναι μια επαναληπτική μέθοδος με την οποία μπορούμε να βρούμε την κατά μέτρο μεγαλύτερη ιδιοτιμή ενός πίνακα A και το ιδιοδιάνυσμα που αντιστοιχεί σε αυτή. Όπως είδαμε στο Κεφάλαιο 4 ο αλγόριθμος που περιγράφει τον τρόπο με τον οποίο δουλεύει η Μέθοδος των Δυνάμεων είναι ο εξής:

$u^{(0)} = \text{κάποιο διάνυσμα με } \|u^{(0)}\| = 1$
$\text{για } k = 1,2,\dots$
$w = Au^{(k-1)}$ $\qquad$ *εφαρμόζει τον πίνακα* A *.*
$u^{(k)} = \frac{w}{\|w\|}$ $\qquad$ *κανονικοποιεί το νέο διάνυσμα.*
$\lambda^{(k)} = \left(u^{(k)}\right)^T Au^{(k)}$ $\qquad$ *βρίσκει το Πηλίκο του Rayleigh.*

Παίρνουμε λοιπόν ως δεδομένο ένα πίνακα $A \in \mathbb{R}^{n\times n}$ και μια αρχική εκτίμηση w για το ιδιοδιάνυσμα και στη συνέχεια πραγματοποιούμε $k = 1,2,\dots$ επαναλήψεις, όπου σε κάθε επανάληψη εκτελούνται τα ακόλουθα βήματα: α) εφαρμόζουμε τον πίνακα A στο κανονικοποιημένο διάνυσμα που έχουμε από την προηγούμενη επανάληψη, β) κανονικοποιούμε το νέο διάνυσμα, γ) βρίσκουμε το Πηλίκο του Rayleigh που αντιστοιχεί στο νέο διάνυσμα.

Ο αλγόριθμος λοιπόν της Μεθόδου των Δυνάμεων παράγει επαναληπτικά την ακολουθία $u^{(k)}$ και θεωρεί στην k επανάληψη, το διάνυσμα $u^{(k)}$ ως την προσέγγιση του u_1 και το Πηλίκο Rayleigh ως την προσέγγιση της ιδιοτιμής λ_1.

Με βάση λοιπόν τα βήματα που ακολουθεί η Μέθοδος των Δυνάμεων και θεωρώντας ότι οι επαναλήψεις συνεχίζονται μέχρις ότου η διαφορά μεταξύ δύο διαδοχικών ιδιοτιμών να είναι μικρότερη από την απαιτούμενη ακρίβεια, θα κατασκευάσουμε ένα πρόγραμμα στο περιβάλλον της Matlab.

Το ακόλουθο m-file της Matlab υλοποιεί τη Μέθοδο των Δυνάμεων.

```
function [lambda,x,relerr,nit] = powermethod(A,w,toler)
 % Find the eigenvector corresponding to the dominant
 % eigenvalue of matrix A by
```

```
% the Power Method.
% INPUT:
% A - coefficient matrix
% w - initial estimate of the eigenvector of the matrix A.
% toler - tolerance for estimated relative error
% OUTPUT:
% lambda - final estimate of dominant eigenvalue
% x - final estimate of corresponding eigenvector
% relerr - estimated relative error in infinity norm
% nit - number of iterations
[n,m]=size(A); % check size of matrix
if n~=m
error('The input matrix is not square')
disp('')
return;
end
relerr = inf;
nit = 1;
while relerr >= toler
u = w/norm(w);% normalize
w = A*u; % Apply A
lamda1 = u'*w; % Rayleigh quotient
if nit>1,
relerr = abs( lambda1-lambda0 )/abs( lambda0 );
end
lambda0 = lambda1;
nit = nit+1;
end
lambda = lambda1;
x =u;
end
```

Παράδειγμα 1

Ορίζουμε τον πίνακα:

$$A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 8 & -1 \\ 1 & -1 & 2 \end{bmatrix}.$$

Αρχικά με την συνάρτηση eig(A) της Matlab θα βρούμε τις ιδιοτιμές και τα αντίστοιχα ιδιοδιανύσματα του πίνακα, προκειμένου να τα συγκρίνουμε με τα αποτελέσματα που θα πάρουμε από την εκτέλεση της Μεθόδου των Δυνάμεων.

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

Ιδιοτιμή	Ιδιοδιάνυσμα
0.344297211778027	-0.834120939686035 0.071441569898443 0.546935425866465
2.489765906101409	0.551108594642854 0.148999621589125 0.821022794858386
8.165936882120560	0.022838014100700 -0.986253119060590 0.163655767558595

Παρατηρούμε ότι η μεγαλύτερη κατά μέτρο ιδιοτιμή είναι η 8.165936882120560.

Στη συνέχεια καλούμε τη συνάρτηση powermethod με αρχική εκτίμηση $[1 \quad 1 \quad 1]^T$ και απαιτούμενη ακρίβεια toler= 10^{-6}.

>> [lambda,x,relerr,nit] = powermethod(A,[1;1;1],1e-6)

Στον πίνακα που ακολουθεί περιέχονται τα αποτελέσματα που προκύπτουν από την εκτέλεση της παραπάνω εντολής :

	Μέθοδος των Δυνάμεων
lambda	8.165936767593001
x	-0.022759731409771 0.986274273826213 -0.163539143371641
relerr	1.368436028247686e-07
nit	10

Διαπιστώνουμε λοιπόν ότι η μέθοδος, έπειτα από 10 επαναλήψεις, συγκλίνει στην ιδιοτιμή 8.165936767593001, δηλαδή προσεγγίζει με πολύ μικρό σφάλμα, τη μεγαλύτερη κατά μέτρο ιδιοτιμή του πίνακα, που είναι σύμφωνα με την εντολή της MATLAB η 8.165936882120560. Οπότε τα αποτελέσματα που προέκυψαν από τη Μέθοδο των Δυνάμεων, συγκλίνουν με αυτά που μας έδωσε η εντολή βιβλιοθήκης `eig` της MATLAB.

Παράδειγμα 2

Στο παράδειγμα αυτό θα μελετήσουμε την περίπτωση στην οποία οι δύο, κατά μέτρο, μεγαλύτερες ιδιοτιμές είναι πολύ κοντά. Όπως αναφέραμε στο Κεφάλαιο 4 η σύγκλιση της Μεθόδου των Δυνάμεων είναι γραμμική, μειώνοντας το σφάλμα κατά ένα σταθερό όρο $\approx |\lambda_2/\lambda_1|$ σε κάθε επανάληψη. Αν οι δύο μεγαλύτερες ιδιοτιμές είναι πολύ κοντά σε μέγεθος, τότε η σύγκλιση θα είναι πολύ αργή.

Ορίζουμε τον πίνακα :

$$A = \begin{bmatrix} 15 & 0 & -1 \\ 0 & 15.1 & 2 \\ -1 & 2 & 0 \end{bmatrix}.$$

Με την εντολή [X,D] = eig(A) βρίσκουμε τις ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα A.

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

Ιδιοτιμή	**Ιδιοδιάνυσμα**
-0.324580970687017	0.064577835415014 -0.128318334171901 0.989628267729080
15.015969864263340	-0.934648870645293 -0.355258633518452 0.014926215598088
15.408611106423672	0.349658678963628 -0.925918845485904 -0.142874419681970

Η μεγαλύτερη κατά μέτρο ιδιοτιμή είναι η 15.408611106423672.

Στη συνέχεια καλούμε τη συνάρτηση powermethod με αρχική εκτίμηση $[1 \quad 1 \quad 1]^T$ και απαιτούμενη ακρίβεια toler= 10^{-6}.

>> [lambda,x,relerr,nit] = powermethod(A,[1;1;1],1e-6)

Στο παρακάτω πίνακα καταγράφονται τα αποτελέσματα που προκύπτουν για τη Μέθοδο των Δυνάμεων:

	Μέθοδος των Δυνάμεων
lambda	15.408323372950827
x	-0.324229069768437 0.935196569676223 0.142417998802538
relerr	9.885848298351703e-07
nit	164

Η μέθοδος λοιπόν συγκλίνει στην ιδιοτιμή 15.408323372950827 η οποία προσεγγίζει με πολύ μικρό σφάλμα τη μεγαλύτερη κατά μέτρο ιδιοτιμή, δηλαδή την 15.408611106423672. Σύμφωνα λοιπόν με τα δύο προηγούμενα παραδείγματα βλέπουμε ότι στην δεύτερη περίπτωση όπου οι δυο μεγαλύτερες ιδιοτιμές είναι πολύ κοντά, η σύγκλιση είναι αρκετά αργή (συγκλίνει μετά από 164 επαναλήψεις) καθώς ο λόγος $|\lambda_2/\lambda_1| \approx 0{,}97$ είναι πολύ μεγάλος. Αντίθετα στο πρώτο παράδειγμα όπου η μεγαλύτερη ιδιοτιμή είναι σημαντικά μεγαλύτερη από τις άλλες, και $|\lambda_2/\lambda_1| \approx 0.30$, η σύγκλιση πραγματοποιείται μετά από 10 επαναλήψεις.

8.2 Μέθοδος της Αντίστροφης Επανάληψης.

Έστω A αντιστρέψιμος πίνακας. Η Μέθοδος της Αντίστροφης Επανάληψης βασίζεται στο γεγονός ότι οι ιδιοτιμές του A^{-1} είναι οι αντίστροφες του A. Άρα με την εφαρμογή της Μεθόδου των Δυνάμεων στον πίνακα A^{-1} υπολογίζουμε την μεγαλύτερη κατά μέτρο ιδιοτιμή του A^{-1}, το αντίστροφο της οποίας είναι η κατά μέτρο μικρότερη ιδιοτιμή του A. Επίσης βασιζόμενοι στο γεγονός ότι μια μετατόπιση ενός πίνακα κατά μI δίνει μετατόπιση μ σε όλες τις ιδιοτιμές, μπορούμε να εφαρμόσουμε τη Μέθοδο των Δυνάμεων στον πίνακα $(A-\mu I)^{-1}$ για να υπολογίσουμε την πλησιέστερη ιδιοτιμή προς τον πραγματικό αριθμό μ.

Ο αλγόριθμος που περιγράφει τον τρόπο με τον οποίο δουλεύει η Μέθοδος της Αντίστροφης Επανάληψης είναι ο εξής:

$u^{(0)} = \textit{κάποιο διάνυσμα με } \|u^{(0)}\| = 1$

$\textit{για } k = 1,2,\ldots$

$\textit{λύνει το σύστημα } (A-\mu I)w = u^{(k-1)} \textit{ως προς } w \qquad \textit{εφαρμόζει τον } (A-\mu I)^{-1}$

$u^{(k)} = w/\|w\| \qquad$ *κανονικοποιεί το νέο διάνυσμα.*

$\lambda^{(k)} = \left(u^{(k)}\right)^T A u^{(k)} \qquad$ *υπολογίζει το Πηλίκο Rayleigh.*

Σύμφωνα λοιπόν με τον παραπάνω αλγόριθμο, έχοντας ως δεδομένο ένα πίνακα $A \in \mathbb{R}^{n\times n}$ και μια αρχική εκτίμηση w για το ιδιοδιάνυσμα, για $k = 1,2,\ldots$ εκτελούμε τα εξής βήματα: α)

Λύνουμε το σύστημα $(A-\mu I)w=u^{(k-1)}$ για να υπολογίσουμε το w, β) Κανονικοποιούμε το νέο διάνυσμα και γ) Υπολογίζουμε το Πηλίκο Rayleigh που αντιστοιχεί στο νέο αυτό διάνυσμα.

Με βάση λοιπόν τα βήματα αυτά και υποθέτοντας ότι οι επαναλήψεις σταματούν όταν η διαφορά μεταξύ δύο διαδοχικών ιδιοτιμών είναι μικρότερη της απαιτούμενης ακρίβειας, θα κατασκευάσουμε ένα πρόγραμμα στο περιβάλλον της Matlab το οποίο θα υλοποιεί τη Μέθοδος της Αντίστροφης Επανάληψης.

```
function [lambda,x0,relerr,nit] = inverseIter(A,w,s,toler)
% inverseIter.m - Inverse Iteration for iterative approximation
% of dominant eigenvalue of A
% INPUT:
% A - coefficient matrix
% w – initial estimate of the eigenvector of the matrix A.
% toler - tolerance for estimated relative error
% s-real number who is not an eingenvalue of A
% OUTPUT:
% lambda - final estimate of dominant eigenvalue
% x0 - final estimate of corresponding eigenvector
% relerr - estimated relative error in infinity norm
% nit - number of iterations
[n,m]=size(A); % check size of matrix
if n~=m
error('The input matrix is not square')
disp('')
return;
end
relerr = inf;
nit = 1;
while relerr >= toler
u = w/norm(w); % normalize
w = (A-s*eye(n))\u;
lambda1 = u'*A*u; % Rayleigh quotient
if nit>1,
relerr = abs( lambda1-lambda0 )/abs( lambda0 );
end
lambda0 = lambda1;
nit = nit+1;
end
lambda = lambda1;
x0 =u;
end
```

Παράδειγμα 3

Έστω ο πίνακας $A = \begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}$.

Σύμφωνα με τη συνάρτηση eig, ο πίνακας A έχει ιδιοτιμές και ιδιοδιανύσματα:

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

Ιδιοτιμή	**Ιδιοδιάνυσμα**
0.585786437626905	0.500000000000000 0.707106781186547 0.500000000000000
2.000000000000000	-0.707106781186547 0.000000000000000 0.707106781186547
3.414213562373095	-0.500000000000000 0.707106781186547 -0.500000000000000

Η μικρότερη κατά μέτρο ιδιοτιμή του A και άρα η μεγαλύτερη του A^{-1} είναι η 0.585786437626905.

Στη συνέχεια καλούμε τη συνάρτηση inverseIter με αρχική εκτίμηση $[0 \quad 1 \quad 0]^T$, απαιτούμενη ακρίβεια toler= $1e-6$ και μια προσέγγιση της ιδιοτιμής $s = 0.5$.

>> [lambda,x0,relerr,nit] = inverseIter(A,[0;1;0],0.5,1e-6)

Στον επόμενο πίνακα καταγράφονται τα αποτελέσματα που προέκυψαν για τη Μέθοδο της Αντίστροφης Επανάληψης.

	Αντίστροφη Επανάληψη
lambda	0.585786437628500
x0	0.499999624543868 0.707107312161303 0.499999624543868
relerr	3.139154924390718e-09
nit	6

Βλέπουμε ότι η μέθοδος συγκλίνει στην ιδιοτιμή 0.585786437628500 η οποία είναι η μικρότερη ιδιοτιμή του πίνακα A. Προκύπτει επίσης ότι η ταχύτητα σύγκλισης είναι πολύ μεγάλη, αφού έπειτα από 6 μόλις επαναλήψεις, το πρόγραμμα μας επιστρέφει την ιδιοτιμή που είναι πιο κοντά στη προσέγγιση 0.5 που δώσαμε, καθώς και το ιδιοδιάνυσμα που αντιστοιχεί σε αυτή.

Στη συνέχεια για το ίδιο παράδειγμα θα καλέσουμε τη Μέθοδο των Δυνάμεων:

>> [lambda,x,relerr,nit] = powermethod(A,[0;1;0],1e-6)

Προκύπτει ο εξής πίνακας:

	Μέθοδος των Δυνάμεων
lambda	3.414213499851323
x	-0.499925656084639 0.707211903729489 -0.499925656084639
relerr	6.037635603779616e-07
nit	7

Από τα παραπάνω αποτελέσματα, φαίνεται ότι η Μέθοδος των Δυνάμεων συγκλίνει έπειτα από 7 επαναλήψεις στην ιδιοτιμή 3.414213499851323, η οποία προσεγγίζει με μικρό σφάλμα την κατά μέτρο μεγαλύτερη ιδιοτιμή του πίνακα A.

Παράδειγμα 4.

Επανερχόμαστε στη συνέχεια στο Παράδειγμα 2 και εφαρμόζουμε τη Μέθοδο της Αντίστροφης Επανάληψης. Επιλέγουμε μια προσέγγιση της ιδιοτιμής το 15.3 που είναι πιο κοντά στο 15.408611106423672 απ' ότι στις άλλες ιδιοτιμές του πίνακα A . Γι'αυτό η μέθοδος πλησιάζει όπως θα δούμε παρακάτω την ιδιοτιμή 15.408611106423672.

>> [lambda,x0,relerr,nit] = inverseIter(A,[1;1;1],15.3,1e-6)

Τα αποτελέσματα που προκύπτουν από την εκτέλεση της εντολής αυτής περιέχονται στον πίνακα που έπεται:

	Αντίστροφη Επανάληψη
lambda	15.408609342342096
x0	-0.351639009274388 0.925163746330797 0.142905736874665
relerr	6.684459704712823e-07
nit	9

Παρατηρούμε λοιπόν ότι και οι δύο μέθοδοι συγκλίνουν με πολύ μικρό σφάλμα στην κατά μέτρο μεγαλύτερη ιδιοτιμή, δηλαδή στην 15.408611106423672. Ακόμη η Μέθοδος της Αντίστροφης Επανάληψης είναι αρκετά ταχύτερη από τη Μέθοδο των Δυνάμεων, απαιτεί μόλις 9 επαναλήψεις για να επιτευχθεί σύγκλιση, σε αντίθεση με τη Μέθοδο των Δυνάμεων η οποία απαιτεί 164. Αυτό οφείλεται στο γεγονός ότι για τη Μέθοδο της Αντίστροφης Επανάληψης ο λόγος $\left|\frac{\mu-\lambda_J}{\mu-\lambda_K}\right| \approx 0.39$ είναι αρκετά μικρότερος της μονάδας.

Αν τώρα επιλέξουμε μια προσέγγιση της ιδιοτιμής να είναι το 15 τότε η Μέθοδος της Αντίστροφης Επανάληψης θα προσεγγίσει την ιδιοτιμή 15.015969864263340. Επίσης, καθώς ο λόγος $\left|\frac{\mu-\lambda_J}{\mu-\lambda_K}\right| \approx 0.05$, είναι δηλαδή πολύ μικρός, η Μέθοδος θα είναι γρήγορη. Πράγματι:

>> [lambda,x0,relerr,nit] = inverseIter(A,[1;1;1],15,1e-6)

Τα αποτελέσματα που προέκυψαν φαίνονται στον ακόλουθο πίνακα:

	Αντίστροφη Επανάληψη
lambda	15.015969864708541
x0	0.934637096067235 0.355289811724689 -0.014921405417785
relerr	1.937943082404070e-08
nit	5

8.3. Επαναληπτική Μέθοδος Πηλίκου Rayleigh.

Στο Κεφάλαιο 4 αναφερθήκαμε στη επαναληπτική μέθοδο του Πηλίκου Rayleigh. Η επαναληπτική αυτή μέθοδος η οποία βρίσκει τη μεγαλύτερη κατά μέτρο ιδιοτιμή ενός πίνακα A και το ιδιοδιάνυσμα που αντιστοιχεί σε αυτή, απαιτεί μια αρχική εκτίμηση για το ιδιοδιάνυσμα έτσι ώστε να προκύψει μια ακριβής προσέγγιση της αντίστοιχης ιδιοτιμής του.

Ο αλγόριθμος που περιγράφει τα βήματα που ακολουθεί η Επαναληπτική Μέθοδος του Πηλίκου Rayleigh είναι ο ακόλουθος:

$$u^{(0)} = \textit{κάποιο διάνυσμα με } \left\|u^{(0)}\right\| = 1$$
$$\lambda^{(0)} = \left(u^{(0)}\right)^T A u^{(0)} = \textit{το αντίστοιχο στο } u^{(0)} \textit{ Πηλίκο Rayleigh.}$$
$$\textit{για } k = 1,2,..$$
$$\qquad \textit{λύνουμε το } \left(A - \lambda^{(k-1)} I\right) w = u^{(k-1)} \textit{ ως πρός } w \qquad \textit{εφαρμόζει τον } \left(A - \lambda^{(k-1)} I\right)^{-1}.$$
$$\qquad u^{(k)} = w/\|w\| \qquad \textit{κανονικοποιεί το νέο διάνυσμα.}$$
$$\qquad \lambda^{(k)} = \left(u^{(k)}\right)^T A u^{(k)} \quad \textit{υπολογίζει το Πηλίκο Rayleigh}.$$

Για έναν πίνακα λοιπόν $A \in \mathbb{R}^{n\times n}$ και μια αρχική εκτίμηση $u^{(0)}$ για το ιδιοδιάνυσμα, υπολογίζουμε το Πηλίκο Rayleigh $\lambda^{(0)} = \left(u^{(0)}\right)^T A u^{(0)}$ που αντιστοιχεί στο $u^{(0)}$ και κατόπιν για $k = 1,2,..$ πραγματοποιούμε τα ακόλουθα βήματα: α) λύνουμε το σύστημα $\left(A - \lambda^{(k-1)} I\right) w = u^{(k-1)}$ για να βρούμε το w, β) κανονικοποιούμε το ιδιοδιάνυσμα που προκύπτει και γ) υπολογίζουμε το Πηλίκο Rayleigh που αντιστοιχεί στο νέο αυτό ιδιοδιάνυσμα.

Βασιζόμενοι λοιπόν στον προηγούμενο αλγόριθμο θα κατασκευάσουμε ένα πρόγραμμα στο περιβάλλον της MATLAB το οποίο θα υλοποιεί την Επαναληπτική Μέθοδο του Πηλίκου Rayleigh.

```
function [lambda,x,relerr,nit] = rayleigh_quotient(A,u0,toler)
% This function finds the dominant eigenvalue and corresponding
% eigenvector of a matrix A, using the Rayleigh-
% Quotient iteration method.
% INPUT:
% A - coefficient matrix
% u0-initial guess
% toler-tolerance for estimated relative error
% OUTPUT:
% lambda - final estimate of dominant eigenvalue
% x - final estimate of corresponding eigenvector
% relerr - estimated relative error in infinity norm
% nit - number of iterations.
```

```
u1 = u0/norm(u0,2);
if norm(A*u1)<=eps;
error('Bad choice for initial guess')
disp('')
return;
end
[n,m] = size(A); % check size of matrix
if n~=m
error('The input matrix is not square!')
disp('')
return;
end
relerr=inf;
nit = 1;
while relerr >=toler
lambda1 = u1'*A*u1; % Rayleight quotient
w = (A-lambda1*eye(n))\u1;
u = w/norm(w,2);% normalize
u1 = u;
if nit>1
relerr = abs( lambda1-lambda0 )/abs(lambda0);
end
lambda0 = lambda1;
nit = nit + 1;
end
lambda=lambda1;
x=u/norm(u,2);
end
```

Παράδειγμα 5

Έστω ο πίνακας $A = \begin{bmatrix} 1 & 4 & 0 \\ 4 & 3 & 1 \\ 0 & 1 & 8 \end{bmatrix}$, με ιδιοτιμές και ιδιοδιανύσματα:

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

Ιδιοτιμή	Ιδιοδιάνυσμα
-2.160657181791619	0.783162602273291 -0.618827125846423 0.060904242193644
5.841487472620544	-0.599821280095929 -0.726006803348915 0.336345883630485
8.319169709171074	0.163923062271933 0.299945178003824 0.939767056162135

Στη συνέχεια καλούμε τη συνάρτηση rayleigh_quotient με αρχική εκτίμηση $[1 \quad 1 \quad 1]^T$ και απαιτούμενη ακρίβεια tol= $1e-6$.

>> [lambda,x,relerr,nit] = rayleigh_quotient(A,[1;1;1],1e-6)

Οπότε προκύπτουν τα αποτελέσματα που βρίσκονται καταχωρημένα στο παρακάτω πίνακα:

	Rayleigh Quotient
lambda	8.319169709171073
x	0.163923062271933 0.299945178003824 0.939767056162135
relerr	3.302090676526759e-07
nit	6

Στο σημείο αυτό θα συγκρίνουμε τα αποτελέσματα της Μεθόδου του Πηλίκου Rayleigh με αυτά που προκύπτουν από τη Μέθοδο των Δυνάμεων αλλά και με αυτά που δίνει η Μέθοδος της Αντίστροφης Επανάληψης.

Έχουμε λοιπόν:

>> [lambda,x,relerr,nit]=powermethod(A,[1;1;1],1e-6)

Και τα αποτελέσματα που προκύπτουν από την εκτέλεση της εντολής αυτής είναι:

	Μέθοδος των Δυνάμεων
lambda	8.319162305799344
x	0.164959662369353 0.301199697930223 0.939184248035353
relerr	9.150184288504288e-07
nit	19

Τέλος θα καλέσουμε την συνάρτηση inverseIter

>> [lambda,x0,relerr,nit] = inverseIter(A,[1;1;1],8,1e-6)

Όπου επιλέξαμε μια προσέγγιση της ιδιοτιμής το 8 που είναι πιο κοντά στο 8.319169709171074 (δηλαδή στη μέγιστη κατά μέτρο ιδιοτιμή) απ' ότι στις άλλες ιδιοτιμές του πίνακα A. Για το λόγο αυτό η μέθοδος πλησιάζει, όπως θα δούμε παρακάτω, στην ιδιοτιμή 8.319169709171074

Στον πίνακα που ακολουθεί φαίνονται τα αποτελέσματα που προέκυψαν για την Μέθοδο της Αντίστροφης Επανάληψης:

	Αντίστροφη Επανάληψη.
lambda	8.319169703018966
x0	0.163893169248131 0.299909003892028 0.939783814745867
relerr	3.308333645389806e-08
nit	7

Όπως παρατηρούμε και οι τρείς μέθοδοι προσεγγίζουν με πολύ μικρό σφάλμα την ιδιοτιμή 8.319169709171074, δηλαδή την κατά μέτρο μεγαλύτερη ιδιοτιμή του πίνακα A. Επίσης σύμφωνα με τα αποτελέσματα που προέκυψαν, η ιδιοτιμή στην οποία συγκλίνει η Μέθοδος του Πηλίκου Rayleigh είναι πιο κοντά στην τιμή 8.319169709171074 απ' ότι οι ιδιοτιμές στις οποίες συγκλίνουν οι δύο άλλες μέθοδοι. Ακόμη η Μέθοδος του Πηλίκου Rayleigh είναι αρκετά πιο γρήγορη σε σχέση με τις υπόλοιπες, καθώς απαιτεί μόλις 6 επαναλήψεις προκειμένου να επιτευχθεί σύγκλιση, ενώ η Μέθοδος των Δυνάμεων συγκλίνει μετά από 19 επαναλήψεις, και η Μέθοδος της Αντίστροφης επανάληψης μετά από 7.

Παράδειγμα 6

Έστω ο πίνακας $A = \begin{bmatrix} 3 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & -1 & -4 \end{bmatrix}$.

Με την ακόλουθη εντολή βρίσκουμε τις ιδιοτιμές και τα ιδιοδιανύσματα του πίνακα αυτού.

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

Ιδιοτιμή	**Ιδιοδιάνυσμα**
3.140054944640259	0.990334273778511 0.000000000000000 0.138701211889401
-4.140054944640259	-0.138701211889401 0.000000000000000 0.990334273778511
1.000000000000000	0.089087080637475 0.979957887012223 -0.178174161274950

Παρατηρούμε ότι η κατά μέτρο μεγαλύτερη ιδιοτιμή είναι αρνητικός αριθμός (-4.140054944640259).

Κατόπιν με αρχική εκτίμηση $[1 \quad 1 \quad 1]^T$ και απαιτούμενη ακρίβεια tol= $1e-6$ καλούμε τη συνάρτηση rayleigh_quotient με σκοπό να βρούμε την ιδιοτιμή του πίνακα A στην οποία συγκλίνει η Μέθοδος του Πηλίκου Rayleigh καθώς και το ιδιοδιάνυσμα που αντιστοιχεί στην ιδιοτιμή αυτή.

>> [lambda,x,relerr,nit] = rayleigh_quotient(A,[1;1;1],1e-6)

	Rayleigh Quotient
lambda	0.999999999999238
x	-0.089087080637475 -0.979957887012223 0.178174161274950
relerr	4.880647762142419e-07
nit	6

Σύμφωνα με τα αποτελέσματα που προέκυψαν, βλέπουμε ότι για τον συγκεκριμένο πίνακα η Επαναληπτική Μέθοδος του Πηλίκου Rayleigh δεν συγκλίνει στη μεγαλύτερη κατά μέτρο ιδιοτιμή του A, αλλά προσεγγίζει την ιδιοτιμή 1.000000000000000.

Παρατηρούμε ότι η τιμή που παίρνει το lambda μετά από 6 επαναλήψεις της μεθόδου, είναι πιο κοντά στο Πηλίκο Rayleigh που αντιστοιχεί στο αρχικό διάνυσμα. Πράγματι πραγματοποιώντας μια επανάληψη της παραπάνω μεθόδου παίρνουμε τα ακόλουθα αποτελέσματα :

	Rayleigh Quotient
lambda1	0.333333333333333
x	0.328199606059184 0.904550133772873 -0.272165526975909
relerr	Inf
nit	1

Κατόπιν για τον ίδιο πίνακα θα χρησιμοποιήσουμε τη Μέθοδο των Δυνάμεων και της Αντίστροφης Επανάληψης:

>> [lambda,x,relerr,nit] = powermethod(A,[1;1;1],1e-6)

	Μέθοδος των Δυνάμεων
lambda	-4.140050629750609
X	-0.137938744801885 0.000000000000000 0.990440761823987
relerr	7.695332115658415e-07
nit	28

>> [lambda,x0,relerr,nit] = inverseIter(A,[1;1;1],-3,1e-6)

Όπου επιλέξαμε μια προσέγγιση της ιδιοτιμής το -3 που είναι πιο κοντά στο -4.140054944640259 (δηλαδή στη μέγιστη κατά μέτρο ιδιοτιμή) απ’ ότι στις άλλες ιδιοτιμές του πίνακα A .

	Αντίστροφη Επανάληψης
lambda	-4.140055217126239
x0	-0.138701192534556 0.000000275145855 0.990334276489215
relerr	2.967442026255750e-07
nit	14

Με βάση τα στοιχεία των παραπάνω πινάκων, παρατηρούμε ότι και οι δύο επαναληπτικές μέθοδοι προσεγγίζουν με πολύ μικρό σχετικό σφάλμα, τη μεγαλύτερη κατά μέτρο ιδιοτιμή. Ωστόσο η Μέθοδος της Αντίστροφης Επανάληψης είναι ταχύτερη από τη Μέθοδο των Δυνάμεων, καθώς η πρώτη συγκλίνει μετά από 14 επαναλήψεις, ενώ η δεύτερη μετά από 28.

8.4. Απλός Αλγόριθμος QR.

Ο Αλγόριθμος QR είναι μια αριθμητική μέθοδος προσέγγισης των ιδιοτιμών ενός πίνακα $A \in \mathbb{R}^{n\times n}$, ο οποίος βασίζεται στην παραγοντοποίηση QR κατά την οποία ο A γράφεται ως $A = QR$, όπου ο Q είναι ένας πίνακας του οποίου οι στήλες είναι ορθοκανονικές, και ο R είναι ένας άνω τριγωνικός πίνακας. Στο Κεφάλαιο 5 περιγράψαμε τον αλγόριθμο αυτό ως εξής:

$A^{(0)} = A$

$για\ k = 1,2, \ldots$

$$Q^{(k)}R^{(k)} = A^{(k-1)} \qquad παραγοντοποίηση\ QR\ του\ A^{(k-1)}$$

$$A^{(k)} = R^{(k)}Q^{(k)} \quad αναδιατάσσει\ τους\ παράγοντες\ με\ αντίστροφη\ σειρά.$$

Παίρνοντας λοιπόν ως δεδομένο ένα πίνακα $A \in \mathbb{R}^{n\times n}$ πραγματοποιούμε k επαναλήψεις (όπου $k = 1,2, \ldots$) μέχρις ότου επέλθει σύγκλιση. Σε κάθε επανάληψη υπολογίζουμε την QR παραγοντοποίηση του πίνακα $A^{(k-1)}$ και στη συνέχεια αναδιατάσσουμε τους παράγοντες $Q^{(k)}$ και $R^{(k)}$ για να πάρουμε τον επόμενο πίνακα $A^{(k)}$.

Κατόπιν στο περιβάλλον του MATLAB κατασκευάσαμε ένα m-file το οποίο υλοποιεί αυτή τη μέθοδο.

```
function [B,Q] = qr_iteration(A,tol)
% Given a square matrix A, this function use
% the QR algorithm to produce a diagonal matrix B,
% with A's eigenvalues on the main diagonal,and a matrix Q with
% A's eigenvectors on his columns.
[m,n]=size(A); % check size of matrix
if m~=n
error('The input matrix is not square!');
else
B = A;
n=size(A,1);
Q=eye(n);
Qb=Q;
lowtri=2*tol;
nit=0;
while lowtri > tol
[Q,R]=qr(B); % built-in Matlab command for QR factorization.
Qb=Qb*Q;
B = R*Q;
```

```
lowtri = max(max(abs(tril(B,-1))));
nit=nit+1;
end
disp('Number of Iterations :')
disp(nit)
end
```

Παράδειγμα 7

Έστω ο πίνακας : $A = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & -1 \\ 1 & -1 & 0 \end{bmatrix}$.

Με τη συνάρτηση eig της Matlab παίρνουμε τις ιδιοτιμές του πίνακα A καθώς και τα ιδιοδιανύσματα που αντιστοιχούν σε αυτές:

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

Ιδιοτιμή	Ιδιοδιάνυσμα
-2.000000000000000	-0.577350269189626 0.577350269189626 0.577350269189626
1.000000000000000	0.393825375182167 0.816328687059106 -0.422503311876939
1.000000000000000	0.715239988066447 0.016557214470529 0.698682773595918

Στη συνέχεια για τον παραπάνω πίνακα καλούμε τη μέθοδο qr_iteration με απαιτούμενη ακρίβεια toler= 10^{-10} .

>> [B,Q]=qr_iteration(A,1e-10)

Τα αποτελέσματα που προέκυψαν περιέχονται στον ακόλουθο πίνακα.

Απλός Αλγόριθμος QR			
Ιδιοτιμές (B)	-1.999999999999998	0.000000000053467	-0.000000000030869
	0.000000000053467	1.000000000000000	0.000000000000000
	-0.000000000030869	0.000000000000000	1.000000000000000
Ιδιοδιανύσματα (Q)	-0.577350269206429	-0.707106781176255	0.408248290457928
	0.577350269181224	-0.707106781196841	-0.408248290457916
	0.577350269181224	-0.000000000010283	0.816496580933666
Αριθμός επαναλήψεων	36		

Παρατηρούμε λοιπόν ότι μετά από 36 επαναλήψεις, προέκυψε ένας πίνακας (B) του οποίου τα διαγώνια στοιχεία είναι οι ιδιοτιμές του A.

Επίσης όπως έχουμε ήδη αναφέρει οι στήλες του $Q^{(k)}$, στις οποίες αντιστοιχούν τα διαγώνια στοιχεία του $A^{(k)}$, συγκλίνουν στα ιδιοδιανύσματα. Έτσι ο πίνακας Q περιέχει τα ιδιοδιανύσματα ως στήλες του, όπως αυτά προέκυψαν μετά από 36 επαναλήψεις.

Παράδειγμα 8

Έστω ο πίνακας $A = \begin{bmatrix} 6 & -2 & 2 \\ -2 & 5 & 0 \\ 2 & 0 & 7 \end{bmatrix}$.

Με ιδιοτιμές και ιδιοδιανύσματα :

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

Ιδιοτιμή	Ιδιοδιάνυσμα
3.000000000000000	0.666666666666667 0.666666666666667 -0.333333333333333
6.000000000000000	-0.333333333333333 0.666666666666667 0.666666666666667
9.000000000000000	0.666666666666667 -0.333333333333333 0.666666666666667

Κατόπιν καλούμε τη μέθοδο qr_iteration με απαιτούμενη ακρίβεια 10^{-10}. Πληκτρολογούμε την εντολή:

[B,Q]=qr_iteration(A,1e-10)

Τα αποτελέσματα που προκύπτουν βρίσκονται καταχωρημένα στον παρακάτω πίνακα.

Απλός Αλγόριθμος QR			
Ιδιοτιμές(B)	9.000000000000007	0.000000000091790	-0.000000000000001
	0.000000000091791	6.000000000000004	-0.000000000000001
	0.000000000000000	0.000000000000000	3.000000000000003
Ιδιοδιανύσματα (Q)	0.666666666676867	-0.333333333312936	-0.666666666666666
	-0.333333333353732	0.666666666656468	-0.666666666666666
	0.666666666646271	0.666666666687065	0.333333333333333
Αριθμός επαναλήψεων	58		

Παρατηρούμε λοιπόν ότι στις 58 επαναλήψεις τα διαγώνια στοιχεία του πίνακα B που προκύπτει, προσεγγίζουν με μεγάλη ακρίβεια τις ιδιοτιμές του πίνακα A.

Παράδειγμα 9

Έστω ο πίνακας $A = \begin{bmatrix} 1 & -1 & 3 \\ -1 & -4 & -1 \\ 0 & 1 & -3 \end{bmatrix}$.

Ο πίνακας A έχει ιδιοτιμές και ιδιοδιανύσματα:

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

Ιδιοτιμή	Ιδιοδιάνυσμα
1.048917339522306	0.981605307496518 -0.185352041026693 -0.045778173640996
-3.356895867892206	-0.587311778886737 -0.272050508145210 0.762268584816946
-3.692021471630098	-0.543244208071300 -0.477760727286989 0.690384253774081

Στη συνέχεια δίνουμε την εντολή να εκτελεστεί η Μέθοδος QR με απαιτούμενη ακρίβεια toler= 10^{-10}.

Απλός Αλγόριθμος QR	
Ιδιοτιμές (B)	-3.692021472045054 -1.479320234828664 2.112431720316421 0.000000000094003 -3.356895867477256 -2.519770598596110 0.000000000000000 -0.000000000000000 1.048917339522307
Ιδιοδιανύσματα (Q)	0.543244207998310 -0.260216812812335 0.798231132446035 0.477760727533165 0.877621781980363 -0.039047343323137 -0.690384253661159 0.402575729675055 0.601084323679569
Αριθμός επαναλήψεων	215

Σύμφωνα με τα παραπάνω αποτελέσματα ο αλγόριθμος είναι αρκετά αργός καθώς, απαιτεί πολλές επαναλήψεις μέχρις ότου να συγκλίνει σε ένα πίνακα του οποίου τα διαγώνια στοιχεία ταυτίζονται με τις ιδιοτιμές του πίνακα A.

Παρατηρώντας λοιπόν τα αποτελέσματα που προέκυψαν από τα τρία προηγούμενα παραδείγματα βλέπουμε ότι όσο πιο κοντά είναι οι ιδιοτιμές, τόσο πιο αργός είναι ο αλγόριθμος. Πράγματι, στο τελευταίο παράδειγμα όπου οι δύο από τις τρείς ιδιοτιμές του πίνακα ήταν αρκετά κοντά, χρειάστηκαν 215 επαναλήψεις της Μεθόδου QR προκειμένου να συγκλίνει στον πίνακα που περιέχει τις ιδιοτιμές του πίνακα A και στον πίνακα που περιέχει τα αντίστοιχα ιδιοδιανύσματα ως στήλες του. Ωστόσο παρατηρούμε ότι οι προσεγγίσεις των ιδιοδιανυσμάτων της δεύτερης και τρίτης στήλης δεν είναι πολύ καλές.

8.5. "Πρακτικός" Αλγόριθμος QR.

Όπως στη Μέθοδο της Αντίστροφης Επανάληψης εισάγαμε εναλλαγές προκειμένου να επιταχύνουμε τη σύγκλιση, έτσι θα τροποποιήσουμε τον Αλγόριθμο QR που παρουσιάστηκε προηγουμένως, εισάγοντας εναλλαγές σε κάθε βήμα.

Προκύπτει λοιπόν ο Πρακτικός Αλγόριθμος ο οποίος δουλεύει ως εξής (Κεφάλαιο 5):

$$\left(Q^{(0)}\right)^T A^{(0)} Q^{(0)} = A \quad Το\ A^{(0)}\ είναι\ μια\ τριδιαγωνιοποίηση\ του\ A$$

$$για\ k = 1,2,\dots$$

$$επέλεξε\ ένα\ \mu^{(k)}$$

$$Q^{(k)}R^{(k)} = A^{(k-1)} - \mu^{(k)}I \qquad παραγοντοποίηση\ QR\ του\ A^{(k-1)} - \mu^{(k)}I$$

$$A^{(k)} = R^{(k)}Q^{(k)} + \mu^{(k)}I \qquad αναδιατάσσει\ παράγοντες\ με\ αντίστροφη\ σειρά$$

Αν κάποιο μη διαγώνιο στοιχείο $A_{j,j+1}^{(k)}$ είναι αρκετά κοντά στο μηδέν,

θέτουμε $A_{j,j+1} = A_{j+1,j} = 0$ ώστε να πάρουμε

$$\begin{bmatrix} A_1 & 0 \\ 0 & A_2 \end{bmatrix} = A^{(k)}$$

Και τώρα εφήρμοσε τον αλγόριθμο QR για το A_1 και A_2

Με βάση τον παραπάνω αλγόριθμο, θα κατασκευάσουμε ένα m-file στο MATLAB το οποίο θα υλοποιεί την Μέθοδο QR.

```
function [B] = shifted_qr_iteration(A,tol)
% Given a square matrix A, this function use
% the shifted QR algorithm to produce a diagonal matrix B,
% with A's eigenvalues on the main diagonal, and a matrix Q with
% A's eigenvectors on his columns.
[m,n]=size(A); % check size of matrix
if m~=n
error('The input matrix is not square!');
else
B = A;
lowtri=2*tol;
nit=0;
while lowtri > tol
n=size(B,1);
I=eye(n,n);
m = B(n,n);
[Q,R]=qr(B-m*I); % built-in Matlab command for QR factorization.
B = R*Q+m*I;
lowtri = max(max(abs(tril(B,-1))));
nit=nit+1;
end
disp('Number of Iterations :')
disp(nit)
end
```

Παράδειγμα 10

Για τον πίνακα A που χρησιμοποιήσαμε στο Παράδειγμα 9, θα καλέσουμε την μέθοδο shifted_qr_iteration, με απαιτούμενη ακρίβεια toler= 10^{-10}.

Πρακτικός Αλγόριθμος QR	
Ιδιοτιμές (Β)	1.048917339527337 -2.568140005517884 2.103466935302694 0.000000000009286 -3.692021471635126 1.407154491908836 0.000000000000000 0.000000000000000 -3.356895867892210
Αριθ. επαναλήψεων	11

Παρατηρώντας τα αποτελέσματα που προέκυψαν για τις ιδιοτιμές του πίνακα A από τη χρήση του Απλού Αλγορίθμου QR στο Παράδειγμα 9, και από τον Πρακτικό Αλγόριθμο QR, διαπιστώνουμε ότι ο δεύτερος συγκλίνει πολύ πιο γρήγορα από τον πρώτο.

Παράδειγμα 11

Στο παράδειγμα αυτό θα χρησιμοποιήσουμε τον πίνακα του Παραδείγματος 7.

Τα αποτελέσματα που προκύπτουν από την εκτέλεση της εντολής : **>> [B]=shifted_qr_iteration(A,1e-10)**

είναι τα ακόλουθα:

Πρακτικός Αλγόριθμος QR	
Ιδιοτιμές (B)	-2.000000000000001 0.000000000003342 0.000000000001929 0.000000000003342 1.000000000000000 0.000000000000000 -0.000000000001929 0.000000000000000 1.000000000000000
Αριθμός επαναλήψεων	4

Προκύπτει λοιπόν ότι ο " Πρακτικός " Αλγόριθμος QR συγκλίνει γρηγορότερα σε ένα πίνακα του οποίου τα διαγώνια στοιχεία προσεγγίζουν με αρκετά μεγάλη ακρίβεια τις ιδιοτιμές του πίνακα A . Αντίθετα ο Απλός Αλγόριθμος QR είναι αργός καθώς απαιτεί 36 επαναλήψεις προκειμένου να συγκλίνει.

8.6. Μέθοδος Ταυτοχρόνων Επαναλήψεων.

Όπως αναφέραμε στο Κεφάλαιο 5, η ιδέα στην οποία βασίζεται η Μέθοδος των Ταυτόχρονων Επαναλήψεων, είναι η εφαρμογή της Μεθόδου των Δυνάμεων ταυτοχρόνως σε περισσότερα από ένα διανύσματα σε κάθε βήμα.

Ο Αλγόριθμος της Μεθόδου των Ταυτοχρόνων Επαναλήψεων είναι ο εξής:

$$
\begin{aligned}
&\textit{Πάρε } \hat{Q}^{(0)} \in \mathbb{R}^{m\times n} \textit{ με ορθοκανονικές στήλες.}\\
&\textit{για } k = 1,2,\dots\\
&\quad Z = A\hat{Q}^{(k-1)}\\
&\quad \hat{Q}^{(k)}\hat{R}^{(k)} = Z \qquad \textit{γίνεται παραγοντοποίηση } QR \textit{ του } Z
\end{aligned}
$$

Σε κάθε βήμα αυτής της επαναληπτικής μεθόδου παίρνουμε ένα νέο πίνακα Z , πολλαπλασιάζοντας τον πίνακα A με τον πίνακα $\hat{Q}^{(k-1)}$. Στη συνέχεια μπορούμε να πάρουμε τα ορθοκανονικά διανύσματα στήλες του πίνακα Q, από την QR παραγοντοποίηση του πίνακα Z. Αυτή η διαδικασία κατόπιν επαναλαμβάνεται.

Το παρακάτω m-file περιγράφει τη Μέθοδο των Ταυτοχρόνων Επαναλήψεων.

```
function [S,Q]= Simit(A,tol)
% Performs simultaneous iterations on square matrix A to calculate
% its real Schur form.
[m,n]=size(A);
if m~=n
error('Matrix must be square!');
disp('')
return
end
n=size(A,1);
Q=eye(n);
lowtri=2*tol;
nit=0;
while lowtri > tol
Z=A*Q;
[Q,R]=qr(Z);
S=Q'*A*Q;
lowtri = max(max(abs(tril(S,-1))));
nit=nit+1;
end
disp('Number of Iterations :')
disp(nit)
end
```

Παράδειγμα 12.

Για τον πίνακα A που χρησιμοποιήσαμε στο Παράδειγμα 9, θα καλέσουμε την μέθοδο Simit με σκοπό να τη συγκρίνουμε με τις δύο άλλες μεθόδους. Τα αποτελέσματα που προέκυψαν από τη μέθοδο αυτή είναι τα εξής:

>> [S,Q]=Simit(A,1e-10)

Μέθοδος των Ταυτοχρόνων Επαναλήψεων			
Ιδιοτιμές (S)	-3.692021472045046	-1.479320234828657	-2.112431720316429
	0.000000000094003	-3.356895867477260	2.519770598596097
	0.000000000000001	0.000000000000000	1.048917339522306
Ιδιοδιάνυσμα (Q)	-0.543244207998310	0.260216812812333	0.798231132446036
	-0.477760727533159	-0.877621781980365	-0.039047343323137
	0.690384253661159	-0.402575729675051	0.601084323679569
Αριθμός επαναλήψεων	215		

Με βάση τα παραπάνω αποτελέσματα παρατηρούμε ότι η Μέθοδος των Ταυτοχρόνων Επαναλήψεων είναι πολύ αργή όταν οι ιδιοτιμές του πίνακα είναι κοντά, όπως συνέβη και με τη Μέθοδο QR. Μάλιστα παρατηρούμε ότι οι δύο αυτές μέθοδοι συγκλίνουν μετά από τον ίδιο αριθμό επαναλήψεων και τα αποτελέσματα τους, τα οποία διαφέρουν ελάχιστα, προσεγγίζουν αρκετά τα αποτελέσματα που προέκυψαν από την συνάρτηση `eig` της Matlab ως προς τις ιδιοτιμές.

Παράδειγμα 13

Έστω ο πίνακας $A = \begin{bmatrix} 3 & -1 & -3 \\ -1 & 3 & -1 \\ -3 & -1 & 3 \end{bmatrix}$, οι ιδιοτιμές του οποίου καθώς και τα ιδιοδιανύσματα που αντιστοιχούν σε αυτές περιλαμβάνονται στον πίνακα που ακολουθεί:

[X,D] = eig(A)

Ιδιοτιμή	Ιδιοδιάνυσμα
-0.561552812808831	0.657192299694123 0.369048184449539 0.657192299694123
3.561552812808830	0.260956473808853 -0.929410263314592 0.260956473808852
6.000000000000000	0.707106781186547 0.000000000000000 -0.707106781186548

Στο σημείο αυτό, θα καλέσουμε την μέθοδο Simit, με απαιτούμενη ακρίβεια toler= 10^{-10}.

Τα αποτελέσματα που προέκυψαν είναι τα εξής:

Μέθοδος των Ταυτοχρόνων Επαναλήψεων			
Ιδιοτιμές (S)	5.999999999999998 0.000000000097202 -0.000000000000002	0.000000000097202 3.561552812808829 -0.000000000000000	-0.000000000000002 -0.000000000000000 -0.561552812808830
Ιδιοδιάνυσμα (Q)	-0.707106781196950 0.000000000037048 0.707106781176145	0.260956473780666 -0.929410263314592 0.260956473837039	0.657192299694123 0.369048184449538 0.657192299694123
Αριθμός επαναλήψεων	44		

Με βάσει τα στοιχεία του προηγούμενου πίνακα διαπιστώνουμε ότι μετά από 44 επαναλήψεις η Μέθοδος των Ταυτοχρόνων Επαναλήψεων συγκλίνει σε ένα πίνακα (S), του οποίου τα διαγώνια στοιχεία είναι οι ιδιοτιμές του πίνακα A, και σε ένα πίνακα (Q) ο οποίος έχει ως στήλες τα ιδιοδιανύσματα που αντιστοιχούν στις ιδιοτιμές του (S).

Στη συνέχεια για τον ίδιο πίνακα A θα χρησιμοποιήσουμε τον Απλό Αλγόριθμο QR καθώς και τον " Πρακτικό " Αλγόριθμο QR με σκοπό να συγκρίνουμε τα αποτελέσματα που προκύπτουν για κάθε αριθμητική μέθοδο, και να δούμε πια από αυτές συγκλίνει με τον μικρότερο αριθμό επαναλήψεων.

Απλός Αλγόριθμος QR	
Ιδιοτιμές (B)	5.999999999999994 0.000000000097202 -0.000000000000000 0.000000000097202 3.561552812808828 -0.000000000000000 -0.000000000000000 -0.000000000000000 -0.561552812808830
Ιδιοδιανύσματα (Q)	0.707106781196949 -0.260956473780666 0.657192299694123 -0.000000000037048 0.929410263314592 0.369048184449539 -0.707106781176145 -0.260956473837039 0.657192299694123
Αριθμός επαναλήψεων	44

Πρακτικός Αλγόριθμος QR	
Ιδιοτιμές (B)	-0.561552812808828 -0.000000000067531 -0.000000000000001 -0.000000000067530 5.999999999999998 0.000000000000000 0.000000000000000 0.000000000000000 3.561552812808830
Αριθμός επαναλήψεων	49

Σύμφωνα λοιπόν με τα αποτελέσματα που φαίνονται στα παραπάνω 3 πινάκια παρατηρούμε ότι οι δύο πρώτες μέθοδοι συγκλίνουν μετά από 44 επαναλήψεις κάτι που περιμέναμε, καθώς οι δύο αυτές μέθοδοι είναι ισοδύναμες, ενώ η τελευταία μέθοδος συγκλίνει μετά από 49 επαναλήψεις. Συμπεραίνουμε λοιπόν ότι στο συγκεκριμένο παράδειγμα η " Πρακτική " Μέθοδος QR είναι πιο αργή από τις δύο άλλες. Επιπλέον τα αποτελέσματα που προκύπτουν από τις τρείς μεθόδους συμφωνούν με αυτά της συνάρτησης `eig` της Matlab.

Παράδειγμα 14

Έστω ο πίνακας $A = \begin{bmatrix} 2 & 1 & 0 \\ 2 & 4 & 2 \\ 2 & 0 & 6 \end{bmatrix}$, με ιδιοτιμές και αντίστοιχα ιδιοδιανύσματα:

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

Ιδιοτιμή	**Ιδιοδιάνυσμα**
1.550510257216821	0.843924811050375 -0.379335546247375 -0.379335546247375
6.449489742783175	-0.156949095431439 -0.698343390261285 -0.698343390261286
4.000000000000000	-0.408248290463863 -0.816496580927726 0.408248290463863

Για το ίδιο παράδειγμα θα καλέσουμε τη Μέθοδο των Δυνάμεων, τη Μέθοδο της Αντίστροφης Επανάληψης καθώς και τη Μέθοδο του Πηλίκου Rayleigh.

Στους τρείς παρακάτω πίνακες φαίνονται τα αποτελέσματα που προκύπτουν για τις τρείς αυτές μεθόδους.

>> [lambda,x,relerr,nit] = powermethod(A,[1;1;1],1e-10)

	Μέθοδος των Δυνάμεων
lambda	6.449489742890446
x	0.156949095481379 0.698343390255674 0.698343390255674
relerr	5.255058269086633e-11
nit	18

>> [lambda,x0,relerr,nit] = inverseIter(A,[1;1;1],6,1e-10)

	Αντίστροφη Επανάληψη.
lambda	6.449489742773545
x0	0.156949095426954 0.698343390261790 0.698343390261790
relerr	1.628015247811624e-11
nit	13

>> [lambda,x,relerr,nit] = rayleigh_quotient(A,[1;1;1],1e-10)

	Rayleigh Quotient
lambda	6.449489742783175
x	0.156949095431439 0.698343390261286 0.698343390261286
relerr	2.754259499967344e-16
nit	7

Με βάση τα αποτελέσματα που προέκυψαν, παρατηρούμε ότι και οι τρείς επαναληπτικές μέθοδοι προσεγγίζουν με πολύ μικρό σχετικό σφάλμα, την κατά μέτρο μεγαλύτερη ιδιοτιμή, δηλαδή την 6.449489742783175. Ωστόσο η Μέθοδος του Πηλίκου Rayleigh είναι ταχύτερη από τις υπόλοιπες, με τη Μέθοδο της Αντίστροφης Επανάληψης να ακολουθεί.

Στους τρείς πίνακες που ακολουθούν περιέχονται ο πίνακας Β (του οποίου η διαγώνιος αποτελείται από τις ιδιοτιμές του πίνακα A) και ο πίνακας Q (ο οποίος περιέχει ως στήλες τα διανύσματα που αντιστοιχούν στις ιδιοτιμές αυτές) όπως αυτοί προέκυψαν από τον Απλό Αλγόριθμο QR, τον " Πρακτικό " Αλγόριθμο QR και τη Μέθοδο των Ταυτοχρόνων Επαναλήψεων.

Απλός Αλγόριθμος QR	
Ιδιοτιμές (Β)	6.449489742783170 -0.912741568124067 -2.672502537185260 -0.000000000000000 4.000000000019617 1.012241581076691 0.000000000000001 -0.000000000047472 1.550510257197205
Ιδιοδιανύσματα (Q)	0.156949095431439 0.377186706211767 -0.912741568079581 0.698343390261286 0.611119448994518 0.372625184788966 0.698343390261286 -0.695890226772699 -0.167491198457164
Αριθμός επαναλήψεων	26

Πρακτικός Αλγόριθμος QR	
Ιδιοτιμές (B)	1.550510257216828 0.473178241855949 -2.657798574897605 -0.000000000000004 3.999999999999965 1.308514075853308 -0.000000000000007 -0.000000000000064 6.449489742783209
Αριθμός επαναλήψεων	4

Μέθοδος των Ταυτοχρόνων Επαναλήψεων	
Ιδιοτιμές (S)	6.449489742783181 -0.912741568124064 2.672502537185265 0.000000000000000 4.000000000019617 -1.012241581076691 -0.000000000000000 0.000000000047472 1.550510257197204
Ιδιοδιάνυσμα (Q)	-0.156949095431439 -0.377186706211767 -0.912741568079581 -0.698343390261286 -0.611119448994517 0.372625184788966 -0.698343390261286 0.695890226772698 -0.167491198457164
Αριθμός επαναλήψεων	26

Βάσει των παραπάνω αποτελεσμάτων, παρατηρούμε ο " Πρακτικός " Αλγόριθμος QR είναι ταχύτερος από τις υπόλοιπες τρεις μεθόδους.

Κατόπιν με την εντολή tic, toc θα βρούμε το χρόνο που απαιτεί η παραπάνω εργασία που κάναμε στην Matlab, δηλαδή θα χρονομετρήσουμε το χρόνο που απαιτούν οι παραπάνω εντολές: *tic,[lambda,x,relerr,nit] = powermethod(A,[1;1;1],1e-10),toc* κ.ο.κ .

Χρόνος που απαιτεί η κάθε μέθοδος	
Αριθμητικές Μέθοδοι	**Χρόνος**
Μέθοδος των Δυνάμεων.	Elapsed time is 0.000876 seconds.
Αντίστροφη Επανάληψη.	Elapsed time is 0.001100 seconds.
Πηλίκο Rayleigh.	Elapsed time is 0.001599 seconds.
Απλός Αλγόριθμος QR.	26 επαναλήψεις: Elapsed time is 0.000839 seconds.
Πρακτικός Αλγόριθμος QR.	4 επαναλήψεις: Elapsed time is 0.000390 seconds.
Ταυτόχρονες Επαναλήψεις.	26 επαναλήψεις: Elapsed time is 0.000951 seconds.

Διαπιστώνουμε λοιπόν ότι από τις τρείς πρώτες μεθόδους, η επαναληπτική μέθοδος του Πηλίκου Rayleigh είναι αυτή που απαιτεί τον περισσότερο χρόνο προκειμένου να πραγματοποιηθεί. Αυτό είναι κάτι το οποίο περιμέναμε, καθώς η μέθοδος αυτή απαιτεί τις περισσότερες πράξεις για την εκτέλεση της, σε σχέση με τις υπόλοιπες μεθόδους. Επίσης σύμφωνα με τα αποτελέσματα που προέκυψαν από την εντολή tic, toc η μέθοδος των Ταυτοχρόνων Επαναλήψεων απαιτεί περισσότερο χρόνο ώστε να πραγματοποιηθεί απ’ ότι η Απλή αλλά και η " Πρακτική " μέθοδος QR.

8.7. Δυο συγκριτικά παραδείγματα μεγαλύτερης διάστασης.

1) Πίνακας Vandermonde.

Ο πίνακας Vandermonde, τον οποίο συναντήσαμε στο Κεφάλαιο 1, είναι ένας πίνακας ο οποίος σε κάθε σειρά του έχει τους όρους μιας γεωμετρικής προόδου. Ένας πίνακας m$\times n$ Vandermonde είναι της μορφής:

$$A=\begin{bmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \dots & x_2^{n-1} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^{n-1} \end{bmatrix}.$$

Η εντολή λοιπόν της Matlab `V = vander(X)` επιστρέφει τον πίνακα Vandermonde του οποίου οι στήλες είναι δυνάμεις του διανύσματος `X`. Ισχύει `A(i,j) = v(i)^(n-j)`, όπου `n = length(X)`.

Με το ακόλουθο m-file δημιουργούμε ένα πίνακα Vandermonde διαστάσεως `nmax`.

```
function[V]=vandermonde(nmax)
for n=1:nmax,
% Interpolation points
X = linspace(0,1,n)';
% Vandermonde matrix
V = vander(X);
end
end
```

Στη συνέχεια πραγματοποιούμε τις ακόλουθες αλλαγές στο m-file powermethod καθώς και σε αυτά των μεθόδων InverseIter.m και Rayleigh_quotient :

Επειδή στη συγκεκριμένη περίπτωση το αρχικό διάνυσμα που πρέπει να εισάγουμε είναι μεγάλης διάστασης, θα τοποθετήσουμε την εντολή w=ones(n,1) μέσα στα προγράμματα που προαναφέραμε, με σκοπό την κατασκευή ενός μοναδιαίου διανύσματος διαστάσεως n.

Προχωρούμε στην κατασκευή ενός πίνακα Vandermonde διαστάσεως 50:

>> [V]=vandermonde(50)

Αρχικά θα χρησιμοποιήσουμε την εντολή της Matlab , [X,D] = eig(V), για να δούμε ποια είναι η μεγαλύτερη κατά μέτρο ιδιοτιμή του πίνακα Vandermonde διαστάσεως 50.

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

10.861680925734795	0.000567144931947	0.000000000384078	0.00000000000000+ 0.000000000000000i	-0.000000000000000 - 0.000000000000000i
-6.510387505595932	-0.000160342063697	-0.000000000076734	0.000000000000001 - 0.000000000000000i	0.000000000000000
2.985918150496272	0.000043719716749	0.000000000014791	0.000000000000000	0.00000000000000+ 0.000000000000000i
-1.233009899463014	-0.000011506072847	-0.000000000002747	-0.000000000000000	0.000000000000000 - 0.000000000000000i
0.475883618850483	0.000002924490454	0.000000000000491	-0.00000000000000+ 0.000000000000000i	-0.00000000000000+ 0.000000000000000i
-0.173576057557874	-0.000000718162126	-0.000000000000088	-0.000000000000000 - 0.000000000000000i	-0.000000000000000 - 0.000000000000000i
0.060216266961219	0.000000170431949	0.000000000000021	0.00000000000000+ 0.000000000000000i	0.000000000000000
-0.019955308326081	-0.000000039091265	-0.000000000000007	0.000000000000000 - 0.000000000000000i	-0.000000000000000
0.006336865333383	0.000000008665491	0.000000000000001	0.000000000000000	-0.00000000000000+ 0.000000000000000i
-0.001932704091976	-0.000000001856171	-0.000000000000001	-0.00000000000000+ 0.000000000000000i	-0.000000000000000 - 0.000000000000000i

Σύμφωνα με τα στοιχεία του πίνακα, η κατά μέτρο μεγαλύτερη ιδιοτιμή του V είναι η 10.861680925734795. Στη συνέχεια θα χρησιμοποιήσουμε τις επαναληπτικές Μεθόδους των Δυνάμεων, της Αντίστροφης Επανάληψης και του Πηλίκου Rayleigh και θα συγκρίνουμε τα αποτελέσματα που δίνει η κάθε μια με το αποτέλεσμά της εντολής της Matlab, [X,D] = eig(V), καθώς και την ταχύτητα με την οποία συγκλίνει η κάθε μέθοδος.

Επίσης το ιδιοδιάνυσμα στο οποίο συγκλίνει η ιδιοτιμή 10.861680925734795 σύμφωνα με τη συνάρτηση `eig` της Matlab είναι:

-0.046560059002302	-0.055148113042865	-0.068110725757251	-0.090582860012516	-0.143679948505144
-0.047288841030603	-0.056200064741442	-0.069791252345570	-0.093822889219020	-0.153816983408767
-0.048042372743894	-0.057296493065208	-0.071566741178359	-0.097343833376313	-0.165939598617370
-0.048822002742927	-0.058440459527378	-0.073446115836935	-0.101187689654506	-0.180779881844638
-0.049629183169002	-0.059635325632052	-0.075439511980656	-0.105405798054156	-0.199492086386799
-0.050465480110650	-0.060884791673032	-0.077558497544141	-0.110061702093633	-0.223984718938094
-0.051332585316780	-0.062192941847663	-0.079816343939871	-0.115235161060217	-0.257570423094409
-0.052232329414693	-0.063564296938247	-0.082228362861796	-0.121027905387000	-0.306273149398363
-0.053166696866959	-0.065003876109488	-0.084812328333285	-0.127572105489442	-0.381514407090922
-0.054137842944176	-0.066517269748192	-0.087589010767271	-0.135043197763656	-0.505720504766389

Για τη Μέθοδο των Δυνάμεων παίρνουμε τα εξής αποτελέσματα:

>> tic,[lambda,x,relerr,nit] = powermethod(V,1e-10),toc

	Μέθοδος των Δυνάμεων
lambda	10.861680925329344
relerr	9.960009092175687e-11
nit	43
Elapsed time	0.001554 seconds

Το ιδιοδιάνυσμα που αντιστοιχεί στην ιδιοτιμή 10.861680925329344 είναι:

x =

0.046560058951160	0.055148112985988	0.068110725693624	0.090582859942145	0.143679948439717
0.047288840978933	0.056200064683933	0.069791252281228	0.093822889148154	0.153816983347095
0.048042372691687	0.057296493007057	0.071566741113300	0.097343833305038	0.165939598561205
0.048822002690172	0.058440459468574	0.073446115771159	0.101187689582934	0.180779881796614
0.049629183115690	0.059635325572585	0.075439511914168	0.105405797982434	0.199492086351023
0.050465480056771	0.060884791612893	0.077558497476948	0.110061702021953	0.223984718921247
0.051332585262322	0.062192941786842	0.079816343871987	0.115235160988831	0.257570423107871
0.052232329359646	0.063564296876736	0.082228362793239	0.121027905316245	0.306273149462432
0.053166696811313	0.065003876047278	0.084812328264082	0.127572105419770	0.381514407243408
0.054137842887919	0.066517269685276	0.087589010697459	0.135043197695681	0.505720505080468

Από τη Μέθοδο της Αντίστροφής Επανάληψης προκύπτουν τα ακόλουθα:

tic,[lambda,x0,relerr,nit] = inverseIter(V,9,1e-10),toc

	Αντίστροφη Επανάληψη.
lambda	10.861680925873323
relerr	5.396296237360224e-11
nit	20
Elapsed time	0.008356 seconds

x0 =

0.046560059027402	0.055148113066404	0.068110725776311	0.090582860019864	0.143679948479844
0.047288841055622	0.056200064764704	0.069791252363899	0.093822889224427	0.153816983377988
0.048042372768819	0.057296493088164	0.071566741195885	0.097343833379570	0.165939598580717
0.048822002767743	0.058440459549996	0.073446115853577	0.101187689655375	0.180779881801952
0.049629183193695	0.059635325654296	0.075439511996326	0.105405798052372	0.199492086338484
0.050465480135201	0.060884791694864	0.077558497558740	0.110061702088897	0.223984718885825
0.051332585341171	0.062192941869042	0.079816343953291	0.115235161052194	0.257570423042654
0.052232329438905	0.063564296959126	0.082228362873916	0.121027905375315	0.306273149357662
0.053166696890971	0.065003876129817	0.084812328343970	0.127572105473680	0.381514407085015
0.054137842967964	0.066517269767917	0.087589010776371	0.135043197743364	0.505720504847759

Για τη Μέθοδο του Πηλίκου Rayleigh παίρνουμε τα παρακάτω αποτελέσματα:

>> tic,[lambda,x,relerr,nit] = rayleigh_quotient(V,1e-10),toc

	Πηλίκο Rayleigh.
lambda	2.985918150496273
relerr	1.487278577198283e-16
nit	12
Elapsed time	0.005479 seconds

x =

0.129368123243159	0.120280236994133	0.095228998949484	0.030704190679428	-0.148348432052306
0.128873627038568	0.118712084514786	0.091179385372793	0.020057121151319	-0.178513638283248
0.128305228472615	0.116981400875335	0.086732623234367	0.008260582154639	-0.210961471340840
0.127657124039469	0.115074006402233	0.081848003460627	-0.004827624026800	-0.244517228381122
0.126923003659549	0.112974273976634	0.076479777639105	-0.019368775003995	-0.276277781195323
0.126095998718079	0.110664952628492	0.070576419530825	-0.035544586991556	-0.299636165993386
0.125168623856577	0.108126965884263	0.064079766239910	-0.053557263310092	-0.300020318990018
0.124132711649183	0.105339180737009	0.056924020317761	-0.073626628094244	-0.245631261539832
0.122979339158636	0.102278142362764	0.049034593172684	-0.095981153757528	-0.067209299478199
0.121698745205492	0.098917768823206	0.040326771162301	-0.120836482115531	0.386282627287387

Σύμφωνα με τα αποτελέσματα που προέκυψαν από τις τρείς προηγούμενες μεθόδους παρατηρούμε ότι η Μέθοδος των Δυνάμεων και η Μέθοδος της Αντίστροφης Επανάληψης συγκλίνουν στην κατά μέτρο μεγαλύτερη ιδιοτιμή του 50×50 πίνακα Vandermonde, η πρώτη μετά από 43 επαναλήψεις και η δεύτερη μετά από 20. Ακόμη βλέπουμε ότι η Μέθοδος του Πηλίκου Rayleigh συγκλίνει στην ιδιοτιμή 2.985918150496273, η οποία είναι πιο κοντά στο Πηλίκο Rayleigh που αντιστοιχεί στο αρχικό διάνυσμα. Πράγματι πραγματοποιώντας μια επανάληψη της παραπάνω μεθόδου παίρνουμε τα ακόλουθα αποτελέσματα :

	Πηλίκο Rayleigh.
lambda	4.960532032764734
relerr	Inf
nit	1
Elapsed time	0.097509 seconds.

Εφαρμόζοντας τη **Μέθοδο QR** με την εντολή **>> tic,[B,Q]=qr_iteration(V,1e-10);,toc, diag(B),** προκύπτουν τα ακόλουθα αριθμητικά αποτελέσματα:

Number of Iterations :

52

Elapsed time is 0.084453 seconds.

eig(V)	**Απλός Αλγόριθμος QR.** (52 επαναλήψεις)
10.861680925734795	10.861680925738215
-6.510387505595932	-6.510387505599375
2.985918150496272	2.985918150496272
-1.233009899463014	-1.233009899463009
0.475883618850483	0.475883618850482
-0.173576057557874	-0.173576057557874
0.060216266961219	0.060216266961219
-0.019955308326081	-0.019955308326081
0.006336865333383	0.006336865333383
-0.001932704091976	-0.001932704091977
0.000567144931947	0.000567144931947
-0.000160342063697	-0.000160342063697
0.000043719716749	0.000043719716749
-0.000011506072847	-0.000011506072847
0.000002924490454	0.000002924490454
-0.000000718162126	-0.000000718162126
0.000000170431949	0.000000170431949
-0.000000039091265	-0.000000039091265
0.000000008665491	0.000000008665494
-0.000000001856171	-0.000000001856176
0.000000000384078	0.000000000384086
-0.000000000076734	-0.000000000076742
0.000000000014791	0.000000000014797
-0.000000000002747	-0.000000000002751
0.000000000000491	0.000000000000491
-0.000000000000088	-0.000000000000083
0.000000000000021	0.000000000000014
-0.000000000000007	-0.000000000000003
0.000000000000001	0.000000000000005
-0.000000000000001	-0.000000000000004
0.000000000000001+ 0.000000000000000i	-0.000000000000000
0.000000000000001- 0.000000000000000i	0.000000000000000
0.000000000000000	-0.000000000000000
-0.000000000000000	-0.000000000000000
- 0.000000000000000+	-0.000000000000000

0.000000000000000i	
- 0.000000000000000- 0.000000000000000i	-0.000000000000000
0.000000000000000- 0.000000000000000i	0.000000000000000
0.000000000000000	0.000000000000000
- 0.000000000000000+ 0.000000000000000i	0.000000000000000
- 0.000000000000000- 0.000000000000000i	-0.000000000000000
0.000000000000000	-0.000000000000000
0.000000000000000+ 0.000000000000000i	0.000000000000000
0.000000000000000	0.000000000000000
0.000000000000000- 0.000000000000000i	0.000000000000000
- 0.000000000000000+ 0.000000000000000i	-0.000000000000000
- 0.000000000000000- 0.000000000000000i	-0.000000000000000
0.000000000000000	0.000000000000000
-0.000000000000000	0.000000000000000
- 0.000000000000000+ 0.000000000000000i	-0.000000000000000
- 0.000000000000000- 0.000000000000000i	0.000000000000000

Στη συνέχεια εφαρμόζουμε τον **" Πρακτικό " αλγόριθμο QR** με την ακόλουθη εντολή:

>> tic,[B]=shifted_qr_iteration(V,1e-10);,toc, diag(B),

Στη συγκεκριμένη περίπτωση ο αλγόριθμος απέτυχε.

Από τη **Μέθοδος των Ταυτοχρόνων Επαναλήψεων** παίρνουμε τα εξής αποτελέσματα:

>> tic,[S,Q]=Simit(V,1e-10);,toc, diag(S),

Number of Iterations :

52

Elapsed time is 0.138577 seconds.

X

eig(V)	**Μέθοδος Ταυτόχρονων Επαναλήψεων.** (52 επαναλήψεις)
10.861680925734795	10.861680925738209
-6.510387505595932	-6.510387505599370
2.985918150496272	2.985918150496274
-1.233009899463014	-1.233009899463011
0.475883618850483	0.475883618850482
-0.173576057557874	-0.173576057557874
0.060216266961219	0.060216266961218
-0.019955308326081	-0.019955308326081
0.006336865333383	0.006336865333383
-0.001932704091976	-0.001932704091976
0.000567144931947	0.000567144931946
-0.000160342063697	-0.000160342063697
0.000043719716749	0.000043719716749
-0.000011506072847	-0.000011506072847
0.000002924490454	0.000002924490454
-0.000000718162126	-0.000000718162125
0.000000170431949	0.000000170431947
-0.000000039091265	-0.000000039091264
0.000000008665491	0.000000008665493
-0.000000001856171	-0.000000001856176
0.000000000384078	0.000000000384087
-0.000000000076734	-0.000000000076744
0.000000000014791	0.000000000014799
-0.000000000002747	-0.000000000002753
0.000000000000491	0.000000000000495
-0.000000000000088	-0.000000000000088
0.000000000000021	0.000000000000016
-0.000000000000007	-0.000000000000002
0.000000000000001	-0.000000000000001
-0.000000000000001	0.000000000000001
0.000000000000001+ 0.000000000000000i	0.000000000000000
0.000000000000001- 0.000000000000000i	-0.000000000000000
0.000000000000000	-0.000000000000000
-0.000000000000000	-0.000000000000000
-	0.000000000000000

0.000000000000000+ 0.000000000000000i	
- 0.000000000000000- 0.000000000000000i	0.000000000000000
0.000000000000000- 0.000000000000000i	0.000000000000000
0.000000000000000	-0.000000000000000
- 0.000000000000000+ 0.000000000000000i	0.000000000000000
- 0.000000000000000- 0.000000000000000i	-0.000000000000000
0.000000000000000	0.000000000000000
0.000000000000000+ 0.000000000000000i	0.000000000000000
0.000000000000000	-0.000000000000000
0.000000000000000- 0.000000000000000i	0.000000000000000
- 0.000000000000000+ 0.000000000000000i	0.000000000000000
- 0.000000000000000- 0.000000000000000i	-0.000000000000000
0.000000000000000	-0.000000000000000
-0.000000000000000	-0.000000000000000
- 0.000000000000000+ 0.000000000000000i	-0.000000000000000
- 0.000000000000000- 0.000000000000000i	0.000000000000000

Από τα αποτελέσματα που προέκυψαν παρατηρούμε ότι ο Απλός Αλγόριθμος QR και η Μέθοδος των Ταυτόχρονων Επαναλήψεων συγκλίνουν μετά από τον ίδιο αριθμό επαναλήψεων σε ένα πίνακα του οποίου τα διαγώνια στοιχεία είναι οι ιδιοτιμές του 50×50 πίνακα Vandermonde. Διαπιστώνουμε επίσης ότι τα αποτελέσματα που προκύπτουν από τις δύο μεθόδους συγκλίνουν αρκετά με αυτά που μας έδωσε η εντολή βιβλιοθήκης `eig` της MATLAB.

2) Πίνακας Redheffer με στοιχεία 1 και 0.

Με την εντολή A=gallery('redheff',n) δημιουργούμε ένα $n \times n$ πίνακα του οποίου τα στοιχεία είναι είτε 0 είτε 1. Συγκεκριμένα αν το $j = 1$ ή το i διαιρεί το j, τότε $A(i,j) = 1$ διαφορετικά, $A(i,j) = 0$.

Στο σημείο αυτό θα κατασκευάσουμε ένα 30×30 πίνακα Redheffer :

>> A=gallery('redheff',30)

Κατόπιν με τη βοήθεια της συνάρτησης eig της Matlab θα βρούμε τις ιδιοτιμές του πίνακα που μόλις δημιουργήσαμε, καθώς και το ιδιοδιάνυσμα που αντιστοιχεί στην κατά μέτρο μεγαλύτερη ιδιοτιμή.

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

Ιδιοτιμές		
7.185774680516740	0.999997858188983 - 0.000001008525641i	1.000000012909947
-3.312906641399024	1.000000197562346+ 0.000002359203117i	1.000000000000000+ 0.000000008729863i
0.180101411588960+ 0.363153901916259i	1.000000197562346 - 0.000002359203117i	1.000000000000000 - 0.000000008729863i
0.180101411588960 - 0.363153901916259i	1.000001944248665+ 0.000001350677474i	1.000000000000000+ 0.000000000000002i
0.766929137704361	1.000001944248665 - 0.000001350677474i	1.000000000000000 - 0.000000000000002i
0.999976133673270+ 0.000023867210433i	0.999999987090049	1.000000000000000
0.999976133673270 - 0.000023867210433i	0.999999991664190+ 0.000000007603532i	1.000000000000002
1.000023866326737+ 0.000023865427352i	0.999999991664190 - 0.000000007603532i	1.000000000000000
1.000023866326737 - 0.000023865427352i	1.000000008335813+ 0.000000007603534i	1.000000000000000
0.999997858188983+ 0.000001008525641i	1.000000008335813 - 0.000000007603534i	0.999999999999999

Ιδιοδιάνυσμα που αντιστοιχεί στη μεγαλύτερη ιδιοτιμή		
-0.622011628721844	-0.116811044163492	-0.100555170669405
-0.371883861184753	-0.116811044163492	-0.100555170669405
-0.268306426126270	-0.116811044163492	-0.100555170669405
-0.205974245591839	-0.116811044163492	-0.100555170669405
-0.189718372097752	-0.116811044163492	-0.100555170669405
-0.168206609298389	-0.100555170669405	-0.100555170669405
-0.151950735804303	-0.100555170669405	-0.100555170669405
-0.133066917657578	-0.100555170669405	-0.100555170669405
-0.133066917657578	-0.100555170669405	-0.100555170669405
-0.133066917657578	-0.100555170669405	-0.100555170669405

Στη συνέχεια θα καλέσουμε της αριθμητικές μεθόδους των Δυνάμεων, της Αντίστροφης Επανάληψης και του Πηλίκου Rayleigh.

>> tic,[lambda,x,relerr,nit] = powermethod(A,1e-10),toc

	Μέθοδος των Δυνάμεων
lambda	7.185774680297249
relerr	9.680261756938280e-11
iterations	31
Elapsed time	0.007412 seconds

Και το αντίστοιχο ιδιοδιάνυσμα της ιδιοτιμής είναι 7.185774680297249:

x =

0.622011628784207	0.116811044147626	0.100555170649828
0.371883861202068	0.116811044147626	0.100555170649828
0.268306426132354	0.116811044147626	0.100555170649828
0.205974245591549	0.116811044147626	0.100555170649828
0.189718372093751	0.116811044147626	0.100555170649828
0.168206609292665	0.100555170649828	0.100555170649828
0.151950735794867	0.100555170649828	0.100555170649828
0.133066917645425	0.100555170649828	0.100555170649828
0.133066917645425	0.100555170649828	0.100555170649828
0.133066917645425	0.100555170649828	0.100555170649828

Από τη Μέθοδο της Αντίστροφής Επανάληψης προκύπτουν τα ακόλουθα:

>> tic,[lambda,x0,relerr,nit] = inverseIter(A,9,1e-10),toc

	Αντίστροφη Επανάληψη.
lambda	7.185774680573083
relerr	8.154786646981744e-11
nit	14
Elapsed time	0.029058 seconds

x0 =

0.622011628695511	0.116811044180970	0.100555170641703
0.371883860964918	0.116811044180970	0.100555170641703
0.268306426188170	0.116811044180970	0.100555170641703
0.205974245752076	0.116811044180970	0.100555170641703
0.189718372212809	0.116811044180970	0.100555170641703
0.168206609423695	0.100555170641703	0.100555170641703
0.151950735884428	0.100555170641703	0.100555170641703
0.133066917720237	0.100555170641703	0.100555170641703
0.133066917720237	0.100555170641703	0.100555170641703
0.133066917720237	0.100555170641703	0.100555170641703

>> tic,[lambda,x,relerr,nit] = rayleigh_quotient(A,1e-10),toc

	Πηλίκο Rayleigh.
lambda	7.185774680516752
relerr	0
nit	8
Elapsed time	0.029129 seconds

x =

-0.622011628721844	-0.116811044163491	-0.100555170669405
-0.371883861184753	-0.116811044163492	-0.100555170669405
-0.268306426126270	-0.116811044163491	-0.100555170669405
-0.205974245591838	-0.116811044163491	-0.100555170669405
-0.189718372097752	-0.116811044163491	-0.100555170669405
-0.168206609298389	-0.100555170669405	-0.100555170669405
-0.151950735804303	-0.100555170669405	-0.100555170669405
-0.133066917657578	-0.100555170669405	-0.100555170669405
-0.133066917657578	-0.100555170669405	-0.100555170669405
-0.133066917657578	-0.100555170669405	-0.100555170669405

Σύμφωνα λοιπόν με τα παραπάνω αποτελέσματα, συμπεραίνουμε ότι για ένα 30×30 πίνακα Redheffer, η Μέθοδος του Πηλίκου Rayleigh προσεγγίζει με μεγαλύτερη ακρίβεια την κατά μέτρο μεγαλύτερη ιδιοτιμή του και σε μικρότερο αριθμό επαναλήψεων απ' ότι οι δύο υπόλοιπες αριθμητικές μέθοδοι.

Σημειώνουμε επίσης ότι για τον συγκεκριμένο πίνακα ο απλός όπως και ο " Πρακτικός " Αλγόριθμος QR, καθώς και η Μέθοδος των Ταυτόχρονων Επαναλήψεων, απέτυχαν.

ΒΙΒΛΙΟΓΡΑΦΙΑ

1. Numerical Linear Algebra, Lloyd N. Trefethen – David Bau, III, Siam, Philadelphia, 1997.

2. Γραμμική Άλγεβρα και Αναλυτική Γεωμετρία, Ανάργυρος Φελλούρης, Αθήνα, 2006.

3. Εισαγωγή στην Αριθμητική Ανάλυση, Α.Μπακόπουλος – Ι. Χρυσοβέργης, Εκδόσεις Συμεών, 2009.

4. Σημειώσεις Αριθμητικής Γραμμικής Άλγεβρας, Β. Δούγαλης – Δ. Νούτσος – Α. Χατζηδήμος, Ιωάννινα, 2012.

5. Αριθμητική Ανάλυση με εφαρμογές σε Matlab και Mathematica, Γ. Σ. Παπαγεωργίου – Χ. Γ. Τσίτουρας, Εκδόσεις Συμεών, Αθήνα, 2006.

6. Θέματα Ανάλυσης Πινάκων, Παναγιώτης Ι. Ψαρράκος, Αθήνα 2014.