

**ΑΡΙΘΜΗΤΙΚΕΣ ΜΕΘΟΔΟΙ ΓΙΑ ΤΟΝ
ΥΠΟΛΟΓΙΣΜΟ ΙΔΙΟΤΙΜΩΝ ΚΑΙ
ΙΔΙΟΔΙΑΝΥΣΜΑΤΩΝ**

Διπλωματική Εργασία

ΤΑΡΟΥΔΑΚΗ ΒΙΚΤΩΡΙΑ

Επιβλέπων Καθηγητής: Παναγιώτης Χατζηπαντελίδης

ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ, ΠΑΝΕΠΙΣΤΗΜΙΟ ΚΡΗΤΗΣ, ΗΡΑΚΛΕΙΟ

Εαρινό εξάμηνο, 2008

Την επιτροπή αξιολόγησης της διπλωματικής αυτής εργασίας αποτέλεσαν οι:
Χατζηπαντελίδης Παναγιώτης (Επιβλέπων)
Μακρινδάκης Χαράλαμπος
Ζουράρης Γεώργιος

0.1. Εισαγωγή

Για να αναγνωρίσουμε αν δυο οργανισμοί έχουν όμοιο DNA, αρκεί να ελέγξουμε αν τα ιδιοδιανύσματα που αντιστοιχούν στις μεγαλύτερες ιδιοτιμές των πινάκων που κατασκευάζονται σύμφωνα με τα υπόλοιπα φυσικά χαρακτηριστικά του προβλήματος έχουν την ίδια κατεύθυνση. Για τη μελέτη του ρυθμού αύξησης ή μείωσης των ερυθρών αιμοσφαιρίων σε έναν οργανισμό είναι επίσης απαραίτητη η γνώση του αν η μεγαλύτερη ιδιοτιμή ενός πίνακα 2×2 είναι μεγαλύτερη, ίση ή μικρότερη της μονάδας. Στη μελέτη των στοιβάδων ατόμων και κατά συνέπεια και των χημικών ενώσεων είναι απαραίτητη η γνώση κάποιων ιδιοτιμών. Γραμμικά συστήματα προκύπτουν και από τη μελέτη ταλαντώσεων δυναμικά συνδεδεμένων σωμάτων για την οποία απαραίτητη είναι η εύρεση ιδιοτιμών πινάκων.

Είναι φανερό λοιπόν ότι η γνώση των ιδιοτιμών και των ιδιοδιανυσμάτων πινάκων είναι πολύ σημαντική και χρησιμοποιείται ευρύτατα σε πολλές επιστημονικές περιοχές όπως η βιολογία, η χημεία, η φυσική και άλλες. Επομένως είναι αναγκαία η χρήση μεθόδων εύρεσης ιδιοτιμών και ιδιοδιανυσμάτων των πινάκων που προκύπτουν από τα εκάστοτε προβλήματα.

Η εργασία αυτή έχει ως σκοπό τη μελέτη ορισμένων μεθόδων αριθμητικής γραμμικής άλγεβρας για τον υπολογισμό των ιδιοτιμών και των ιδιοδιανυσμάτων τυχαίων πινάκων.

Για το σκοπό αυτό, το **πρώτο κεφάλαιο**, με το οποίο ξεκινά η εργασία, περιέχει κάποιους βασικούς ορισμούς και απλά θεωρήματα τα οποία είναι απαραίτητα για την κατανόηση της υπόλοιπης εργασίας. Αρχικά ορίζονται οι ιδιοτιμές και τα ιδιοδιανύσματα που είναι και το θέμα της εργασίας. Στη συνέχεια προσπαθούμε να δούμε πως η διαγωνιοποίηση ενός πίνακα μπορεί να βοηθήσει στη μελέτη που κάνουμε και τότε αυτή είναι δυνατή ορίζοντας το χαρακτηριστικό πολυώνυμο και η ορίζουσα ενός πίνακα που μπορεί κανείς να θεωρήσει ως τον ευκολότερο τρόπο εύρεσης ιδιοτιμών καθώς και την αλγεβρική και τη γεωμετρική πολλαπλότητα των ιδιοτιμών. Η έννοια του μετασχηματισμού ομοιότητας που αναφέρεται εδώ για πρώτη φορά χρησιμοποιείται ευρύτατα στη συνέχεια χωρίς αυτό να αναγράφεται ρητά. Τέλος, ορίζεται και η κανονική μορφή του Schur.

Στο **δεύτερο κεφάλαιο**, αναφέρονται οι δυσκολίες που αντιμετωπίζει κάποιος όταν προσπαθεί να υπολογίσει με τους πιο προφανείς τρόπους τις ιδιοτιμές πινάκων είτε ως ρίζα του χαρακτηριστικού πολυωνύμου είτε ως όριο μιας ακολουθίας. Η παραγοντοποίηση Schur αναφέρεται στο σημείο αυτό καθώς χρησιμοποιείται στους περισσότερους αλγορίθμους εύρεσης ιδιοτιμών. Εδώ διαχωρίζουμε και τους αλγορίθμους εύρεσης ιδιοτιμών σε δύο φάσεις.

Το **τρίτο κεφάλαιο** αναφέρεται στη μετατροπή ενός τυχαίο πίνακα σε μορφή Hessenberg (η οποία είναι διαγώνια μορφή στην περίπτωση που ο αρχικός πίνακας είναι ερμητιανός). Η μορφή Hessenberg είναι μια άνω κλιμακωτή μορφή με μηδενικά στοιχεία κάτω από την πρώτη υποδιαγώνιο που προκύπτει χρησιμοποιώντας ανακλαστές Householder. Η δημιουργία άνω τριγωνικού πίνακα θεωρείται στο κεφάλαιο αυτό ως μια κακή ιδέα. Το κεφάλαιο ολοκληρώνεται με την αναφορά στο κόστος σε πράξεις και της σταθερότητας της μεθόδου αυτής.

Το **τέταρτο κεφάλαιο** αφιερώνεται σε μεθόδους εύρεσης μιας ιδιοτιμής. Η Μέθοδος των Δυνάμεων που βρίσκει μόνο τη μεγαλύτερη ιδιοτιμή του πίνακα είναι η απλούστερη. Παραλλαγές αυτής της μεθόδου μπορούν να υπολογίσουν και τις υπόλοιπες ιδιοτιμές.

Μια κάθε φορά. Από αυτές τις παραλλαγές παρουσιάζεται η Μέθοδος της Αντίστροφης Επανάληψης και μια ακόμη ταχύτερη μέθοδος η Μέθοδος του πηλίκου Rayleigh η οποία όμως απαιτεί περισσότερες πράξεις όπως φαίνεται στην τελευταία ενότητα του κεφαλαίου αυτού.

Η Μέθοδος QR με την πιο "πρακτική" παραλλαγή της και η Μέθοδος των Ταυτόχρονων Επαναλήψεων που μπορεί να παραλληλιστεί με τη Μέθοδο του πηλίκου Rayleigh, υπολογίζουν περισσότερες ιδιοτιμές και παρουσιάζονται στο **πέμπτο κεφάλαιο**.

Η θεωρητική μελέτη που έγινε κατά τη διάρκεια της εργασίας, ολοκληρώνεται στο **έκτο κεφάλαιο** όπου γίνεται μια παραλληλία της Μεθόδου QR με άλλες μεθόδους και αναφέρονται οι εναλλαγές Wilkinson που είναι χρήσιμες για την προσέγγιση ιδιοτιμών όταν οι πίνακες που προκύπτουν σε κάθε βήμα της επανάληψης της Μεθόδου των Ταυτόχρονων Επαναλήψεων είναι μιας ειδικής μορφής.

Το **έβδομο κεφάλαιο**, το οποίο είναι και το τελευταίο, είναι μια καταγραφή κάποιων αριθμητικών αποτελεσμάτων που έχουν προκύψει από προγράμματα που κατασκευάστηκαν κατά τη διάρκεια εκπόνησης της εργασίας για το σκοπό αυτό. Σε αυτό το κεφάλαιο επιβεβαιώνονται κάποια από τα αποτελέσματα που είχαν αποδειχθεί στα προηγούμενα κεφάλαια.

Αρωγός μου στη μελέτη αριθμητικών μεθόδων εύρεσης ιδιοτιμών και ιδιοδιανυσμάτων ήταν ο Επίκουρος Καθηγητής του Τμήματος Μαθηματικών του Πανεπιστημίου Κρήτης, Χατζηπαντελίδης Παναγιώτης. Για την πολύτιμη βοήθειά του, τη συνεργασία και τη συνεχή στήριξή του, χωρίς τις οποίες η εργασία δε θα είχε πραγματοποιηθεί, τον ευχαριστώ θερμά.

Περιεχόμενα

0.1. Εισαγωγή	iii
Κεφάλαιο 1. Βασική Θεωρία	1
1.1. Ιδιοτιμές και Ιδιοδιανύσματα	1
1.2. Διαγωνιοποίηση πινάκων	2
1.3. Γεωμετρική Πολλαπλότητα	3
1.4. Χαρακτηριστικό Πολυώνυμο	3
1.5. Αλγεβρική Πολλαπλότητα	4
1.6. Μετασχηματισμοί ομοιότητας	5
1.7. Ιδιόμορφες ιδιοτιμές και ιδιόμορφοι πίνακες	6
1.8. Διαγωνιοποίηση	6
1.9. Ορίζουσα και Ίχνος	7
1.10. Ορθοκανονική διαγωνιοποίηση	7
1.11. Κανονική μορφή του Schur	8
1.12. Παραγοντοποιήσεις πινάκων για την εύρεση ιδιοτιμών	8
Κεφάλαιο 2. Γενική περιγραφή αλγορίθμων εύρεσης ιδιοτιμών	11
2.1. Μειονεκτήματα προφανών αλγορίθμων	11
2.2. Μια βασική δυσκολία	12
2.3. Παραγοντοποίηση Schur και διαγωνιοποίηση	15
2.4. Δυο φάσεις στον υπολογισμό ιδιοτιμών	15
Κεφάλαιο 3. Απλοποίηση σε Hessenberg ή τριδιαγώνια μορφή	17
3.1. Μια κακή ιδέα	17
3.2. Μια καλή ιδέα	18
3.3. Μέτρηση πράξεων	19
3.4. Η περίπτωση Ερμητιανών πινάκων: Μετατροπή σε τριδιαγώνια μορφή	19
3.5. Σταθερότητα	20
Κεφάλαιο 4. Πηλίκo Rayleigh, Αντίστροφη Αναδρομή	21
4.1. Περιορισμός σε πραγματικούς συμμετρικούς πίνακες	21
4.2. Πηλίκo Rayleigh	21
4.3. Μέθοδος Δυνάμεων	23
4.4. Μέθοδος Αντίστροφης Αναδρομής	25
4.5. Επαναληπτική Μέθοδος πηλίκου Rayleigh	26
4.6. Μέτρηση πράξεων	28
Κεφάλαιο 5. Αλγόριθμος QR χωρίς εναλλαγές	29
5.1. Ο απλός αλγόριθμος QR	29

5.2. Μη κανονικοποιημένες παράλληλες Επαναλήψεις	31
5.3. Ταυτόχρονες επαναλήψεις	32
5.4. Ισοδυναμία Ταυτόχρονων Επαναλήψεων και Αλγορίθμου QR	33
5.5. Σύγκλιση του αλγορίθμου QR	35
Κεφάλαιο 6. Αλγόριθμος QR με εναλλαγές	37
6.1. Σύνδεση με αντίστροφη αναδρομή	37
6.2. Σύνδεση με την εναλλασσόμενη αντίστροφη αναδρομή	38
6.3. Σύνδεση με την επαναληπτική μέθοδο του πηλίκου Rayleigh	39
6.4. Εναλλαγές Wilkinson	39
6.5. Σταθερότητα και ακρίβεια	40
Κεφάλαιο 7. Αριθμητικά Παραδείγματα-Παρατηρήσεις	43
7.1. Παράδειγμα 1	44
7.2. Παράδειγμα 2	46
7.3. Παράδειγμα 3	48
7.4. Παράδειγμα 4	50
7.5. Παράδειγμα 5	52
7.6. Παράδειγμα 6	54
7.7. Παράδειγμα 7	56
7.8. Παρατηρήσεις	58
Βιβλιογραφία	59

ΚΕΦΑΛΑΙΟ 1

Βασική Θεωρία

Ένα ιδιαίτερα ενδιαφέρον πρόβλημα στην περιοχή των επιστημονικών υπολογισμών είναι ο υπολογισμός ιδιοτιμών και ιδιοδιανυσμάτων ενός πίνακα. Ένας λόγος είναι ότι οι ιδιοτιμές είναι απαραίτητες για την επίλυση διαφόρων προβλημάτων της φυσικής και άλλων επιστημών όπως αναφέρθηκε νωρίτερα, ενώ ένας άλλος είναι ότι οι καλύτεροι αλγόριθμοι για την εύρεση ιδιοτιμών είναι ισχυροί κάτι το οποίο θα προσπαθήσουμε να το εξηγήσουμε αργότερα αναλύοντας τους.

Σε αυτό το κεφάλαιο θα διατυπώσουμε απαραίτητους ορισμούς και θεωρήματα για την μελέτη και κατανόηση των αλγορίθμων που θα παρουσιαστούν στα επόμενα κεφάλαια.

1.1. Ιδιοτιμές και Ιδιοδιανύσματα

ΟΡΙΣΜΟΣ 1.1. Έστω $m \in \mathbb{N}$ και $A \in \mathbb{C}^{m \times m}$ 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

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

Αυτό που φαίνεται από την (1.1) είναι ότι το γινόμενο Ax λειτουργεί ως βαθμωτό γινόμενο λx που εξαρτάται από τον πίνακα A και το διάνυσμα x . Ο υπόχωρος \mathcal{S} του \mathbb{C}^m που αποτελείται από τα διανύσματα $x \in \mathbb{C}^m$ τέτοια ώστε $Ax = \lambda x$ ονομάζεται **ιδιόχωρος** του πίνακα A που αντιστοιχεί στην ιδιοτιμή λ και τα μη μηδενικά $x \in \mathcal{S}$ είναι ιδιοδιανύσματα του A που αντιστοιχούν στην ιδιοτιμή λ .

ΟΡΙΣΜΟΣ 1.2. Το σύνολο όλων των ιδιοτιμών ενός πίνακα A καλείται **φάσμα** του A και είναι ένα υποσύνολο του \mathbb{C} το οποίο συμβολίζεται ως $\Lambda(A)$.

Τα προβλήματα ιδιοτιμών έχουν ένα πολύ διαφορετικό χαρακτήρα από τα προβλήματα τετραγωνικών ή ορθογώνιων γραμμικών συστημάτων εξισώσεων. Για ένα σύστημα εξισώσεων, το πεδίο ορισμού μπορεί να είναι ένας χώρος και το σύνολο τιμών ένας άλλος. Στο παρακάτω παράδειγμα ο πίνακας A αντιστοιχεί διανύσματα μήκους n με πολυωνυμικούς συντελεστές σε διανύσματα μήκους m απλών πολυωνύμων.

ΠΑΡΑΔΕΙΓΜΑ 1.3. Πίνακας Vandermonde

Σταθεροποιούμε μια ακολουθία αριθμών x_1, x_2, \dots, x_m . Αν p και q είναι πολυώνυμα βαθμού $< n$ και α είναι ένας αριθμός, τότε τα $p+q$ και αq είναι επίσης πολυώνυμα βαθμού $< n$. Επιπλέον οι τιμές αυτών των πολυωνύμων στα σημεία x_i ικανοποιούν τις παρακάτω ιδιότητες γραμμικότητας:

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

$$(\alpha p)(x_i) = \alpha(p(x_i))$$

Έτσι, η αντιστοίχιση από διανύσματα με συντελεστές πολυώνυμα p βαθμού $< n$ σε διανύσματα $(p(x_1), p(x_2), \dots, p(x_m))$ απλών πολυωνυμικών τιμών είναι γραμμική. Οποιαδήποτε γραμμική αντιστοίχιση μπορεί να εκφραστεί ως πολυπλασιασμός με πίνακα. Εδώ έχουμε ένα παράδειγμα. Στην πραγματικότητα, εκφράζεται μέσω ενός $m \times n$ πίνακα Vandermonde

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

Αν c είναι το διάνυσμα στήλη των συντελεστών του p ,

$$p(x) = c_0 + c_1x + \dots + c_{n-1}x^{n-1},$$

τότε το γινόμενο Ac δίνει τις απλές πολυωνυμικές τιμές. Δηλαδή για κάθε i από το 1 έως το m , έχουμε

$$(Ac)_i = c_0 + c_1x_i + \dots + c_{n-1}(x_i)^{n-1} = p(x_i)$$

Σ'αυτό το παράδειγμα, είναι ξεκάθαρο ότι δε χρειάζεται να σκεφτόμαστε το γινόμενο ενός πίνακα με ένα διάνυσμα, Ac , ως m διακριτά βαθμωτά αδροίσματα το καθένα εκ των οποίων δίνει ένα διαφορετικό γραμμικό συνδυασμό των συνιστωσών του c . Αντιθέτως, μπορούμε να δούμε τον πίνακα A ως πίνακα στηλών όπου κάθε στήλη περιέχει απλές τιμές ενός μονωνύμου $A = (1 \mid x \mid x^2 \mid \dots \mid x^{n-1})$ και το γινόμενο Ac ως ένα μόνο διάνυσμα άδροισης που δίνει απ' ευθείας ένα γραμμικό συνδυασμό αυτών των μονωνύμων,

$$Ac = c_0 + c_1x + \dots + c_{n-1}x^{n-1} = p(x) \quad \square$$

Το να αναζητήσουμε τις ιδιοτιμές ενός τέτοιου πίνακα A θα ήταν ανούσιο. Τα προβλήματα ιδιοτιμών έχουν νόημα μόνο όταν οι χώροι του συνόλου τιμών και το πεδίο ορισμού είναι οι ίδιοι. Αυτό αντανακλά το γεγονός ότι στις εφαρμογές, οι ιδιοτιμές χρησιμοποιούνται κυρίως όταν θέλουμε να συνθέσουμε ένα πίνακα αναδρομικά, είτε κατηγορηματικά ως δύναμη A^k είτε υπονοούμενα σε συναρτησιακή μορφή όπως: e^{tA} , όπως, π.χ., στην λύση διαφορικών εξισώσεων.

Γενικά μιλώντας, οι ιδιοτιμές και τα ιδιοδιανύσματα είναι χρήσιμα για δύο λόγους, έναν αλγοριθμικό και έναν που αφορά τη φυσική. Αλγοριθμικά, η ανάλυση ιδιοτιμών μπορεί να απλοποιήσει λύσεις σε ορισμένα προβλήματα, μετατρέποντας ένα περίπλοκο σύστημα σε συλλογή από βαθμωτά προβλήματα. Όσον αφορά τη φυσική, η ανάλυση ιδιοτιμών μπορεί να δώσει μια άποψη ως προς τη συμπεριφορά συστημάτων που διέπονται από γραμμικές εξισώσεις, όπως, για παράδειγμα συστήματα δύο ή περισσότερων σωμάτων συνδεδεμένα με ελατήρια διαφορετικής τάσης το καθένα που αφορούν την κινηματική.

1.2. Διαγωνιοποίηση πινάκων

Διαγωνιοποίηση ενός τετραγωνικού πίνακα A καλείται η παραγοντοποίησή του, όταν φυσικά αυτή είναι δυνατή, σε

$$(1.2) \quad A = X\Lambda X^{-1}$$

όπου X είναι αντιστρέψιμος και Λ διαγώνιος πίνακας. Η παραπάνω σχέση, (1.2), μπορεί να γραφεί ισοδύναμα ως

$$(1.3) \quad AX = Q\Lambda$$

Δηλαδή

$$[A] \cdot [x_1|x_2|\cdots|x_m] = [x_1|x_2|\cdots|x_m] \cdot \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_m \end{bmatrix}.$$

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

Η διαγωνιοποίηση εκφράζει μια αλλαγή βάσης στις «συντεταγμένες του ιδιοδιανύσματος». Αν $Ax = b$ και $A = X\Lambda X^{-1}$, τότε έχουμε:

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

Έτσι για να υπολογίσουμε το Ax μπορούμε να επεκτείνουμε το x με βάση τις 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 ήδη αναφέρθηκε, το σύνολο των ιδιοδιανυσμάτων που αντιστοιχούν σε μια ιδιοτιμή, μαζί με το μηδενικό διάνυσμα, αποτελούν έναν υπόχωρο του \mathbb{C}^m γνωστό ως *ιδιόχωρο*. Έστω μια ιδιοτιμή ενός πίνακα A , τότε συμβολίζουμε τον αντίστοιχο ιδιόχωρο με \mathcal{E}_λ . Ο ιδιόχωρος \mathcal{E}_λ αποτελεί ένα παράδειγμα ενός αναλλοίωτου υποχώρου του A , δηλαδή $A\mathcal{E}_\lambda \subset \mathcal{E}_\lambda$ (καθώς, αν $x \in A\mathcal{E}_\lambda$, τότε $x = Ay = \lambda y$ για κάποιο ιδιοδιάνυσμα y . Τότε όμως $Ax = A\lambda y = \lambda Ay = \lambda x$. Άρα τελικά $x \in \mathcal{E}_\lambda$.)

Η διάσταση του \mathcal{E}_λ μπορεί να ερμηνευτεί ως ο μέγιστος αριθμός γραμμικώς ανεξαρτήτων ιδιοδιανυσμάτων που αντιστοιχούν στην ίδια ιδιοτιμή λ που μπορεί να βρεθεί. Αυτός ο αριθμός είναι γνωστός ως *γεωμετρική πολλαπλότητα* του λ . Η γεωμετρική πολλαπλότητα μπορεί να περιγραφεί επίσης σαν τη διάσταση του μηδενόχωρου του $A - \lambda I$ από τη στιγμή που αυτός ο μηδενόχωρος είναι ο ίδιος ο \mathcal{E}_λ .

1.4. Χαρακτηριστικό Πολυώνυμο

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

$$(1.5) \quad p_A(z) = \det(zI - A)$$

Λόγω της συγκεκριμένης θέσης του προσήμου $(-)$, το p είναι μονικό (δηλαδή, ο συντελεστής του μεγιστοβάθμιου όρου, βαθμού m , είναι 1).

ΘΕΩΡΗΜΑ 1.4. Το λ είναι μια ιδιοτιμή του πίνακα A αν και μόνο αν $p_A(\lambda) = 0$.

ΑΠÓΔΕΙΞΗ. Το ζητούμενο προκύπτει από τον ορισμό της ιδιοτιμής: λ είναι ιδιοτιμή αν και μόνο αν υπάρχει μη-μηδενικό διάνυσμα τέτοιο ώστε $\lambda x - Ax = 0$. Αυτό όμως είναι ισοδύναμο με το ότι ο πίνακας $\lambda I - A$ είναι μη αντιστρέψιμος και αυτό ισχύει αν και μόνο αν

$$\det(\lambda I - A) = 0. \quad \square$$

Το Θεώρημα 1.4 έχει μια σημαντική συνέπεια. Ακόμα και αν ένας πίνακας είναι πραγματικός, κάποιες από τις ιδιοτιμές του μπορεί να είναι μιγαδικές. (Για παράδειγμα, ο πίνακας:

$$\begin{bmatrix} 0 & -1 \\ 1 & 0 \end{bmatrix}.$$

είναι πραγματικός αλλά οι ρίζες του χαρακτηριστικού του πολυωνύμου, δηλαδή οι ιδιοτιμές του είναι οι $\pm i$ που είναι μιγαδικές.) Στη φυσική, αυτό σχετίζεται με το φαινόμενο στο οποίο, πραγματικά δυναμικά συστήματα μπορεί να παρουσιάζουν παλινδρομικές κινήσεις όπως και αυξήσεις και μειώσεις. Αλγοριθμικά, σημαίνει ότι ακόμα και αν τα δεδομένα εισόδου σε ένα πίνακα προβλήματος ιδιοτιμών είναι πραγματικά, τα αποτελέσματα, (ιδιοτιμές και ιδιοδιανύσματα), μπορεί να είναι μιγαδικά.

1.5. Αλγεβρική Πολλαπλότητα

Από το θεμελιώδες θεώρημα της Άλγεβρας, μπορούμε να γράψουμε το p_A στη μορφή

$$(1.6) \quad p_A(z) = (z - \lambda_1)(z - \lambda_2) \cdots (z - \lambda_m),$$

για κάποια λ_j . Σύμφωνα με το Θεώρημα 1.4, κάθε λ_j είναι ιδιοτιμή του A και όλες οι ιδιοτιμές του A είναι κάποιο $\lambda_j \in \mathbb{C}$. Γενικά, μια ιδιοτιμή μπορεί να εμφανιστεί περισσότερες από μία φορές. Ορίζουμε ως *αλγεβρική πολλαπλότητα* μιας ιδιοτιμής λ του A , την 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 υπολογισμού του πλήθους των ιδιοτιμών ενός πίνακα.

ΘΕΩΡΗΜΑ 1.5. *Εάν $A \in \mathbb{C}^{m \times m}$, τότε ο A έχει m ιδιοτιμές, μετρημένες με τις αλγεβρικές πολλαπλότητές τους. Πιο συγκεκριμένα, αν οι ρίζες του p_A είναι απλές, τότε ο A έχει m διακριτές ιδιοτιμές.*

Η απόδειξη αυτού του Θεωρήματος είναι άμεση από το Θεμελιώδες Θεώρημα της Άλγεβρας σύμφωνα με το οποίο, κάθε μιγαδικό πολυώνυμο έχει ακριβώς τόσες ρίζες όσες και ο βαθμός του μετρημένες με τις πολλαπλότητές τους.

Σημείωση: Γενικά κάθε πίνακας έχει τουλάχιστον μία ιδιοτιμή. (Η ιδιοτιμή είναι μοναδική όταν η πολλαπλότητά της είναι όση και ο βαθμός του χαρακτηριστικού πολυωνύμου του πίνακα δηλαδή όση και η διάσταση του πίνακα.)

Η αλγεβρική πολλαπλότητα μιας ιδιοτιμής είναι πάντα μεγαλύτερη η ίση από τη γεωμετρική της πολλαπλότητα. Αυτό θα αποδειχθεί αργότερα (ως Θεώρημα 1.7) διότι απαιτείται η χρήση μετασχηματισμών ομοιότητας που παρουσιάζονται αμέσως παρακάτω.

1.6. Μετασχηματισμοί ομοιότητας

Εάν ο πίνακας $X \in \mathbb{C}^{m \times m}$ είναι αντιστρέψιμος, τότε η απεικόνιση $A \rightarrow X^{-1}AX$ καλείται *μετασχηματισμός ομοιότητας* του A . Λέμε ότι δυο πίνακες A και B είναι *όμοιοι* αν υπάρχει ένας μετασχηματισμός ομοιότητας που να σχετίζει τον έναν με τον άλλο, δηλαδή αν υπάρχει αντιστρέψιμος πίνακας $X \in \mathbb{C}^{m \times m}$ τέτοιος ώστε $B = X^{-1}AX$. Όπως περιγράφηκε παραπάνω στην ειδική περίπτωση της διαγωνιοποίησης (1.2), οποιοσδήποτε μετασχηματισμός ομοιότητας, είναι μία διαδικασία αλλαγής βάσης. Οι όμοιοι πίνακες έχουν αρκετές κοινές ιδιότητες.

ΘΕΩΡΗΜΑ 1.6. *Εάν X αντιστρέψιμος, τότε ο A και ο $X^{-1}AX$ έχουν ίδια τα: χαρακτηριστικό πολυώνυμο, ιδιοτιμές αλγεβρικές και γεωμετρικές πολλαπλότητες.*

ΑΠÓΔΕΙΞΗ. Η απόδειξη για το ότι τα χαρακτηριστικά πολυώνυμα είναι τα ίδια, είναι ευθύς υπολογισμός:

$$\begin{aligned} p_{X^{-1}AX}(z) &= \det(zI - X^{-1}AX) = \det(X^{-1}(zI - A)X) \\ &= \det(X^{-1})\det(zI - A)\det(X) = \det(zI - A) = p_A(z). \end{aligned}$$

Εφόσον τα χαρακτηριστικά πολυώνυμα συμπίπτουν, το ότι οι ιδιοτιμές και οι αλγεβρικές πολλαπλότητες θα είναι ίσες, είναι προφανές (ιδιοτιμές είναι οι ρίζες του χαρακτηριστικού πολυωνύμου και οι αλγεβρικές πολλαπλότητες, οι πολλαπλότητες των ριζών των πολυ-ωνύμων). Τέλος για την απόδειξη του ότι οι γεωμετρικές πολλαπλότητες είναι οι ίδιες, αρκεί να επαληθεύσουμε ότι εάν \mathcal{E}_λ είναι ένας ιδιόχωρος του A που αντιστοιχεί στην ιδιοτιμή λ , τότε $X^{-1}\mathcal{E}_\lambda$ είναι ιδιόχωρος του $X^{-1}AX$ και αντιστρόφως. (Έχουμε ότι $x \in \mathcal{E}_\lambda \Rightarrow Ax = \lambda x$. Τότε ισχύει: $X^{-1}x \in X\mathcal{E}_\lambda \Rightarrow (X^{-1}AX)X^{-1}x = X^{-1}Ax = X^{-1}\lambda x = \lambda X^{-1}x$. Δηλαδή \mathcal{E}_λ ιδιόχωρος του A , συνεπάγεται $X^{-1}\mathcal{E}_\lambda$ είναι ιδιόχωρος του $X^{-1}AX$. Αντιστρόφως, έστω $y \in X\mathcal{E}_\lambda$ ιδιόχωρο του $X^{-1}AX$ τότε $y = X^{-1}x$ για $x \in \mathcal{E}_\lambda$ οπότε $Xy = x \in \mathcal{E}_\lambda$. Τότε: $(X^{-1}AX)y = \lambda y \Rightarrow AXy = \lambda Xy \Rightarrow Xy$ ιδιοτιμή του A . Άρα \mathcal{E}_λ ιδιοτιμή του A). \square

Τώρα μπορούμε να συσχετίσουμε την αλγεβρική με τη γεωμετρική πολλαπλότητα.

ΘΕΩΡΗΜΑ 1.7. *Η γεωμετρική πολλαπλότητα είναι το πολύ όση και η αλγεβρική πολλαπλότητα.*

ΑΠÓΔΕΙΞΗ. Έστω n η γεωμετρική πολλαπλότητα της ιδιοτιμής λ ενός πίνακα A . (Προφανώς $n \leq m$ διότι υπάρχουν το πολύ m γραμμικώς ανεξάρτητα διανύσματα του $\mathbb{C}^{m \times m}$ και η γεωμετρική πολλαπλότητα είναι ακριβώς πλήθος γραμμικώς ανεξάρτητων ιδιοδιανυσμάτων.) Σχηματίζουμε τον πίνακα $\tilde{V}_{m \times n}$ του οποίου οι n στήλες σχηματίζουν ορθοκανονική βάση του ιδιοχώρου $\{x : Ax = \lambda x\}$. Επεκτείνοντας τον \tilde{V} σε ένα τετραγωνικό ορθο-μοναδιαίο πίνακα V έχουμε ότι ο V^*AV είναι της μορφής

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

όπου I είναι ο $n \times n$ ταυτοτικός πίνακας, C ένας $n \times (m - n)$ πίνακας και D ένας $(m - n) \times (m - n)$ πίνακας. Από τον ορισμό της ορίζουσας ενός πίνακα, έχουμε:

$$\det(zI - B) = \det(zI - \lambda I)\det(zI - D).$$

Επομένως η αλγεβρική πολλαπλότητα μιας ιδιοτιμής λ του B είναι τουλάχιστον n . Καθώς όμως οι μετασχηματισμοί ομοιότητας διατηρούν τις πολλαπλότητες, το ίδιο ισχύει και για τον A , (δηλαδή η αλγεβρική πολλαπλότητα του A είναι τουλάχιστον n). \square

1.7. Ιδιόμορφες ιδιοτιμές και ιδιόμορφοι πίνακες

Παρόλο που υπάρχουν πίνακες που έχουν ίση αλγεβρική και γεωμετρική πολλαπλότητα, αυτό δεν αληθεύει για όλους τους πίνακες.

ΠΑΡΑΔΕΙΓΜΑ 1.8. . Έστω οι πίνακες

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

Και οι δύο έχουν χαρακτηριστικό πολυώνυμο $(z - 2)^3$ επομένως υπάρχει μόνο μια ιδιοτιμή η $\lambda = 2$ αλγεβρικής πολλαπλότητας 3. Στην περίπτωση του πίνακα A μπορούμε να βρούμε 3 γραμμικώς ανεξάρτητα ιδιοδιανύσματα του A , για παράδειγμα τα e_1, e_2, e_3 , επομένως η γεωμετρική πολλαπλότητα είναι επίσης 3. Στην περίπτωση του πίνακα B όμως, μπορούμε να βρούμε μόνο ένα γραμμικά ανεξάρτητο ιδιοδιάνυσμα (ένα πολλαπλάσιο του e_1) επομένως η γεωμετρική πολλαπλότητα της ιδιοτιμής είναι 1.

ΟΡΙΣΜΟΣ 1.9. Μια ιδιοτιμή της οποίας η αλγεβρική πολλαπλότητα είναι μεγαλύτερη από τη γεωμετρική πολλαπλότητα ονομάζεται *ιδιόμορφη ιδιοτιμή*. Ένας πίνακας που έχει μια ή περισσότερες ιδιόμορφες ιδιοτιμές ονομάζεται *ιδιόμορφος πίνακας*.

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

1.8. Διαγωνιοποίηση

Η κλάση των μη ιδιόμορφων πινάκων είναι ακριβώς η κλάση των πινάκων που υφίστανται την διαγωνιοποίηση (1.1).

ΘΕΩΡΗΜΑ 1.10. Ένας $m \times m$ πίνακας A είναι μη ιδιόμορφος εάν και μόνο εάν διαγωνιοποιείται, $A = X\Lambda X^{-1}$.

ΑΠÓΔΕΙΞΗ. : (\Leftarrow) Δεδομένης της διαγωνιοποίησης $A = X\Lambda X^{-1}$, γνωρίζουμε από το Θεώρημα 1.6 ότι ο Λ είναι όμοιος με τον A , με τις ίδιες ιδιοτιμές και τις ίδιες πολλαπλότητες. Εφόσον ο Λ είναι διαγώνιος πίνακας, είναι μη ιδιόμορφος και επομένως το ίδιο ισχύει και για τον A .

(\Rightarrow) Ένας μη ιδιόμορφος πίνακας πρέπει να έχει m γραμμικώς ανεξάρτητα ιδιοδιανύσματα επειδή τα ιδιοδιανύσματα που αντιστοιχούν σε διαφορετικές ιδιοτιμές είναι γραμμικώς ανεξάρτητα (αν ήταν γραμμικώς εξαρτημένα, τότε θα ανήκαν στον ίδιο ιδιόχωρο, κάτι το οποίο θα ήταν άτοπο) και κάθε ιδιοτιμή μπορεί να έχει τόσα γραμμικώς ανεξάρτητα ιδιοδιανύσματα όσα η αλγεβρική πολλαπλότητα της. Αν αυτά τα m ανεξάρτητα ιδιοδιανύσματα τεθούν ως στήλες ενός πίνακα Q , τότε ο Q είναι αντιστρέψιμος και έχουμε $A = X\Lambda X^{-1}$. \square

Βάσει αυτού του αποτελέσματος, ένας άλλος όρος για έναν μη ιδιόμορφο πίνακα είναι διαγωνιοποιήσιμος.

Στην ερώτηση αν ο διαγωνιοποιήσιμος πίνακας A συμπεριφέρεται υπό μία έννοια όπως ο όμοιός του διαγώνιος Λ , η απάντηση εξαρτάται από το ποια συμπεριφορά μετράμε και στο δείκτη κατάσταση του X , του πίνακα των ιδιοδιανυσμάτων. Εάν ο X είναι κακής κατάστασης δηλαδή έχει μεγάλο δείκτη κατάσταση, τότε ένας μεγάλος αριθμός πληροφοριών μπορεί να χαθεί περνώντας από τον A στον Λ .

1.9. Ορίζουσα και Ίχνος

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

ΘΕΩΡΗΜΑ 1.11. Η ορίζουσα $\det(A)$ και το ίχνος $\text{tr}(A)$ είναι ίσα με το γινόμενο και το άθροισμα των ιδιοτιμών του A αντίστοιχα μετρημένων με τις αλγεβρικές τους πολλαπλότητες:

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

ΑΠΟΔΕΙΞΗ. Από τις σχέσεις (1.5) και (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. Για τον δεύτερο, από την (1.5) προκύπτει ότι ο συντελεστής του z^{m-1} στο p_A είναι το αντίθετο του αθροίσματος των διαγώνιων στοιχείων του A δηλαδή το $-\text{tr}(A)$. Επίσης, από την (1.6) ο συντελεστής του όρου z^{m-1} του p_A είναι το αντίθετο του αθροίσματος των διαγώνιων στοιχείων του A ίσος με $-\sum_{j=1}^m \lambda_j$. Επομένως $\text{tr}(A) = \sum_{j=1}^m \lambda_j$. \square

1.10. Ορθοκανονική διαγωνιοποίηση

Μερικές φορές συμβαίνει ο $m \times m$ πίνακας A να έχει όχι μόνο m γραμμικώς ανεξάρτητα ιδιοδιανύσματα αλλά αυτά να μπορούν να επιλεγούν να είναι ορθογώνια (και συνεπώς και ορθομοναδιαία). Σ' αυτή την περίπτωση ο A είναι ορθοκανονικά διαγωνιοποιήσιμος που σημαίνει ότι υπάρχει ορθομοναδιαίος πίνακας Q τέτοιος ώστε

$$(1.9) \quad A = Q\Lambda Q^*.$$

Οι ερμητιανοί πίνακες είναι ορθομοναδιαία διαγωνιοποιήσιμοι. Το επόμενο αποτέλεσμα προκύπτει από το θεώρημα 1.14 που θα αναφερθεί και αποδειχθεί παρακάτω.

ΘΕΩΡΗΜΑ 1.12. Ένας ερμητιανός πίνακας είναι ορθοκανονικά διαγωνιοποιήσιμος και οι ιδιοτιμές του είναι πραγματικές.

Οι ερμητιανοί πίνακες δεν είναι οι μόνοι πίνακες που είναι ορθοκανονικά διαγωνιοποιήσιμοι. Άλλα παραδείγματα ορθοκανονικά διαγωνιοποιήσιμων πινάκων είναι οι αντηρμητιανοί, οι ορθομοναδιαίοι, οι κυκλικοί και κάθε άλλος πίνακας που είναι άθροισμα ενός από τους παραπάνω και ενός πολλαπλασίου του ταυτοτικού πίνακα. Γενικώς, η κλάση των πινάκων που είναι ορθοκανονικά διαγωνιοποιήσιμοι έχει έναν πιο κομψό χαρακτηρισμό. Ένας πίνακας A λέγεται *καυονικός* αν

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

Το επόμενο αποτέλεσμα είναι πολύ γνωστό.

ΘΕΩΡΗΜΑ 1.13. Ένας πίνακας είναι ορθοκανονικά διαγωνιοποιήσιμος εάν και μόνο εάν είναι κανονικός.

1.11. Κανονική μορφή του Schur

Μια άλλη παραγοντοποίηση πινάκων η οποία είναι πολύ χρήσιμη στην αριθμητική ανάλυση επειδή όλοι οι πίνακες, ακόμα και οι ιδιόμορφοι μπορούν να παραγοντοποιηθούν βάσει αυτής είναι η παραγοντοποίηση Schur. Παραγοντοποίηση Schur ενός πίνακα A είναι η παραγοντοποίηση:

$$(1.10) \quad A = QTQ^*.$$

Όπου Q είναι ορθομοναδιαίος και T άνω τριγωνικός. Σημειώνουμε εδώ ότι οι πίνακες A και T είναι όμοιοι, επομένως οι ιδιοτιμές του A αναγκαστικά εμφανίζονται στη διαγώνιο του T .

ΘΕΩΡΗΜΑ 1.14. Κάθε τετραγωνικός πίνακας επιδέχεται μια Schur παραγοντοποίηση.

ΑΠÓΔΕΙΞΗ. Εφαρμόζουμε επαγωγή στη διάσταση m του A . Η περίπτωση $m = 1$ είναι τετριμμένη, καθώς $A = IAI^*$ με I ορθομοναδιαίο και A άνω τριγωνικός. Υποθέτουμε ότι $m \geq 2$. Έστω x ένα ιδιοδιάνυσμα του A με αντίστοιχη ιδιοτιμή λ . Παίρνουμε το x να έχει νόρμα 1, πάντα μπορούμε να βρούμε ένα τέτοιο x , και έστω U ένας ορθομοναδιαίος πίνακας με πρώτη στήλη το x . Τότε, 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 ορθομοναδιαίος πίνακας, διότι

$$QQ^* = U \begin{bmatrix} 1 & 0 \\ 0 & V \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & V^* \end{bmatrix} U^* = UIU^* = UU^* = I$$

και έχουμε ότι: $Q^*AQ = \begin{bmatrix} \lambda & BV \\ 0 & T \end{bmatrix}$ και αυτή είναι η παραγοντοποίηση Schur που ψάχνουμε. \square

1.12. Παραγοντοποιήσεις πινάκων για την εύρεση ιδιοτιμών

Έως τώρα, παρουσιάστηκαν τρία παραδείγματα παραγοντοποιήσεων για να βρίσκουμε ιδιοτιμές, παραγοντοποιήσεων ενός πίνακα που τον φέρνουν σε μία άλλη μορφή πίνακα στην οποία οι ιδιοτιμές είναι εμφανείς. Μπορούμε να τις συνοψίσουμε ως εξής:

Μια διαγωνιοποίηση $A = X\Lambda X^{-1}$ εάν και μόνο εάν ο A είναι μη ιδιόμορφος (οι αλγεβρικές πολλαπλότητες των ιδιοτιμών του είναι ίσες με τις αντίστοιχες γεωμετρικές)

Μια ορθοκανονική διαγωνιοποίηση $A = Q\Lambda Q^*$ υπάρχει εάν και μόνο εάν ο A είναι κανονικός.

Μια ορθοκανονική τριγωνοποίηση (παραγοντοποίηση Schur)

$A = QTQ^*$ υπάρχει πάντα.

Για να υπολογίσουμε ιδιοτιμές, θα πρέπει επομένως να κατασκευάσουμε μια από αυτές τις παραγοντοποιήσεις. Συνήθως αυτή είναι η Schur παραγοντοποίηση καθώς

εφαρμόζεται σε όλους τους πίνακες χωρίς περιορισμούς. Ακόμα, καθώς εμπλέκονται ορθοκανονικοί μετασχηματισμοί, οι αλγόριθμοι που προκύπτουν, τείνουν να είναι αριθμητικά σταθεροί. Εάν ο A είναι κανονικός, τότε ο πίνακας που προκύπτει από τη Schur παραγοντοποίηση είναι διαγώνιος και ειδικά αν ο A είναι ερμητιανός, τότε εκμεταλλευόμενοι τη συμμετρία κατά τον υπολογισμό τον μετατρέπουμε σε διαγώνιο με τη μισή εργασία από όση χρειάζεται για τη μετατροπή ενός γενικού πίνακα A .

Γενική περιγραφή αλγορίθμων εύρεσης ιδιοτιμών

Οι αριθμητικές μέθοδοι εύρεσης ιδιοτιμών χωρίζονται σε δύο κατηγορίες, τις άμεσες μεθόδους και τις επαναληπτικές μεθόδους. Οι άμεσες μέθοδοι, όπως για παράδειγμα η μέθοδος QR που θα αναλυθεί σε επόμενο κεφάλαιο, εφαρμόζονται συνήθως σε πυκνούς πίνακες και υπολογίζουν όλες τις ιδιοτιμές και ενδεχομένως και τα ιδιοδιανύσματα (ο υπολογισμός των ιδιοδιανυσμάτων από τις ιδιοτιμές είναι απλός και προκύπτει από τον ορισμό τους). Το πλήθος των πράξεων που απαιτούνται δεν εξαρτάται από τα στοιχεία του πίνακα. Οι επαναληπτικές μέθοδοι, εφαρμόζονται κυρίως σε αραιούς πίνακες ή σε πίνακες που ο πολλαπλασιασμός πίνακα με διάνυσμα είναι η μόνη πράξη που μπορεί να γίνει εύκολα. Συνήθως αυτές οι μέθοδοι βρίσκουν ένα υποσύνολο του φάσματος του πίνακα και των ιδιοδιανυσμάτων του και εκτελούνται τόσες επαναλήψεις ώστε να επιτευχθεί η επιθυμητή ακρίβεια. Η σύγκλιση αυτών των μεθόδων εξαρτάται από τα στοιχεία του πίνακα. Ο διαχωρισμός σε άμεσες και επαναληπτικές μεθόδους δε σημαίνει ότι οι πρώτες δε χρησιμοποιούν επαναλήψεις.

Αυτό και τα επόμενα κεφάλαια περιγράφουν μερικούς από τους άμεσους αλγορίθμους υπολογισμού ιδιοτιμών και ιδιοδιανυσμάτων όπως και κάποιες παραλλαγές τους.

Οι περισσότεροι από τους αλγορίθμους που θα μελετήσουμε λειτουργούν σε δύο φάσεις. Η πρώτη είναι μια προκαταρκτική μεταβολή από γενική σε συγκεκριμένα δομημένη μορφή και η δεύτερη μια αναδρομική διαδικασία για την τελική σύγκλιση. Αυτό το κεφάλαιο, περιγράφει την προσέγγιση σε δυο φάσεις και εξηγεί τα πλεονεκτήματα.

2.1. Μειονεκτήματα προφανών αλγορίθμων

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

Ίσως η πρώτη μέθοδος που θα μπορούσε κανείς να σκεφτεί θα ήταν να υπολογίσει τους συντελεστές του χαρακτηριστικού πολυωνύμου και να χρησιμοποιήσει μεθόδους εύρεσης ριζών για να υπολογίσει τις ιδιοτιμές. Δυστυχώς, αυτή η σκέψη δεν είναι καλή επειδή οι αλγόριθμοι εύρεσης ριζών πολυωνύμων δημιουργούν προβλήματα κακής κατάστασης γενικά ακόμα και αν το πρόβλημα με τις ιδιοτιμές είναι καλής κατάστασης. Για παράδειγμα, έστω ότι κάποιος πίνακας έχει χαρακτηριστικό πολυώνυμο το $p(x) = (x-2)^9$ το οποίο αναπτύσσεται ως

$$p(x) = x^9 - 18x^8 + 144x^7 - 672x^6 + 2016x^5 - 4032x^4 + 5376x^3 - 4608x^2 + 2304x - 512 .$$

Η ακριβής ρίζα του χαρακτηριστικού αυτού πολυωνύμου είναι προφανώς το 2 με πολλαπλότητα 9. Χρησιμοποιώντας τη μέθοδο εύρεσης των τιμών του πολυωνύμου, Horner και τη μέθοδο της διχοτόμησης για την εύρεση των ριζών του πολυωνύμου, παρατηρούμε ότι οι τιμές του πολυωνύμου για τα $x > 2$ και για τα $x < 2$ που είναι κοντά στη ρίζα είναι και θετικές και αρνητικές επομένως με τη μέθοδο της διχοτόμησης, δεν είναι σίγουρο ότι το νέο διάστημα που θα χρησιμοποιηθεί θα περιέχει τη ρίζα. Άρα

θα πάρουμε μία εκτίμηση ρίζας (δηλαδή ιδιοτιμής) που απέχει μακράν από την ακριβή δηλαδή θα έχουμε πολύ μεγάλο σφάλμα. Στα παρακάτω γραφήματα αναπαριστάται γραφικά το πρόβλημα αυτό.

Μια άλλη ιδέα θα ήταν να εκμεταλλευτούμε το γεγονός ότι η ακολουθία: $\frac{x}{\|x\|}, \frac{Ax}{\|Ax\|}, \frac{A^2x}{\|A^2x\|}, \frac{A^3x}{\|A^3x\|} \dots$ συγκλίνει υπό κάποιες προϋποθέσεις, σε ένα ιδιοδιάνυσμα που αντιστοιχεί στη μεγαλύτερη σε απόλυτη τιμή ιδιοτιμή του A . Αυτή η μέθοδος εύρεσης ιδιοτιμής ονομάζεται «Μέθοδος Δυνάμεων». Δυστυχώς, παρόλο που η μέθοδος δυνάμεων είναι πολύ γνωστή μέθοδος, δεν είναι καθόλου αποτελεσματικό εργαλείο για γενική χρήση. Εκτός από μερικούς ειδικούς πίνακες, είναι πολύ αργός αλγόριθμος.

Οι καλύτεροι γενικοί αλγόριθμοι ιδιοτιμών δε βασίζονται σε αυτές τις ιδέες αλλά σε μια διαφορετική αρχή, τον υπολογισμό μιας ανάλυσης του A , όπου οι ιδιοτιμές του πίνακα εμφανίζονται ως στοιχεία ενός από τους πίνακες της ανάλυσης αυτής. Είδαμε τρεις τέτοιες αναλύσεις για την εύρεση ιδιοτιμών στην προηγούμενη ενότητα: τη διαγωνιοποίηση, την ορθοκανονική διαγωνιοποίηση και την ορθοκανονική τριγωνοποίηση (ή αλλιώς παραγοντοποίηση Schur). Πρακτικά, οι ιδιοτιμές υπολογίζονται συνήθως κατασκευάζοντας μια από αυτές τις αναλύσεις. Για να κάνουμε αυτές τις αναλύσεις, εφαρμόζουμε μια ακολουθία από μετασχηματισμούς ομοιότητας του A οι οποίοι εισάγουν μηδενικά σε κατάλληλες θέσεις. Έτσι, βάσει όσων αναφέραμε για τους μετασχηματισμούς ομοιότητας πινάκων, μπορούμε να βρούμε τις ιδιοτιμές ενός άλλου πίνακα, όμοιου με τον αρχικού, οι οποίες θα υπολογίζονται πολύ ευκολότερα και με περισσότερη ακρίβεια. Αρκεί δηλαδή να γίνει η διαγωνιοποίηση ή η τριγωνοποίηση του πίνακα όταν αυτό είναι δυνατό και τότε οι ιδιοτιμές του αρχικού πίνακα θα είναι τα διαγώνια στοιχεία του διαγώνιου ή του τριγωνικού πίνακα που προκύπτει.

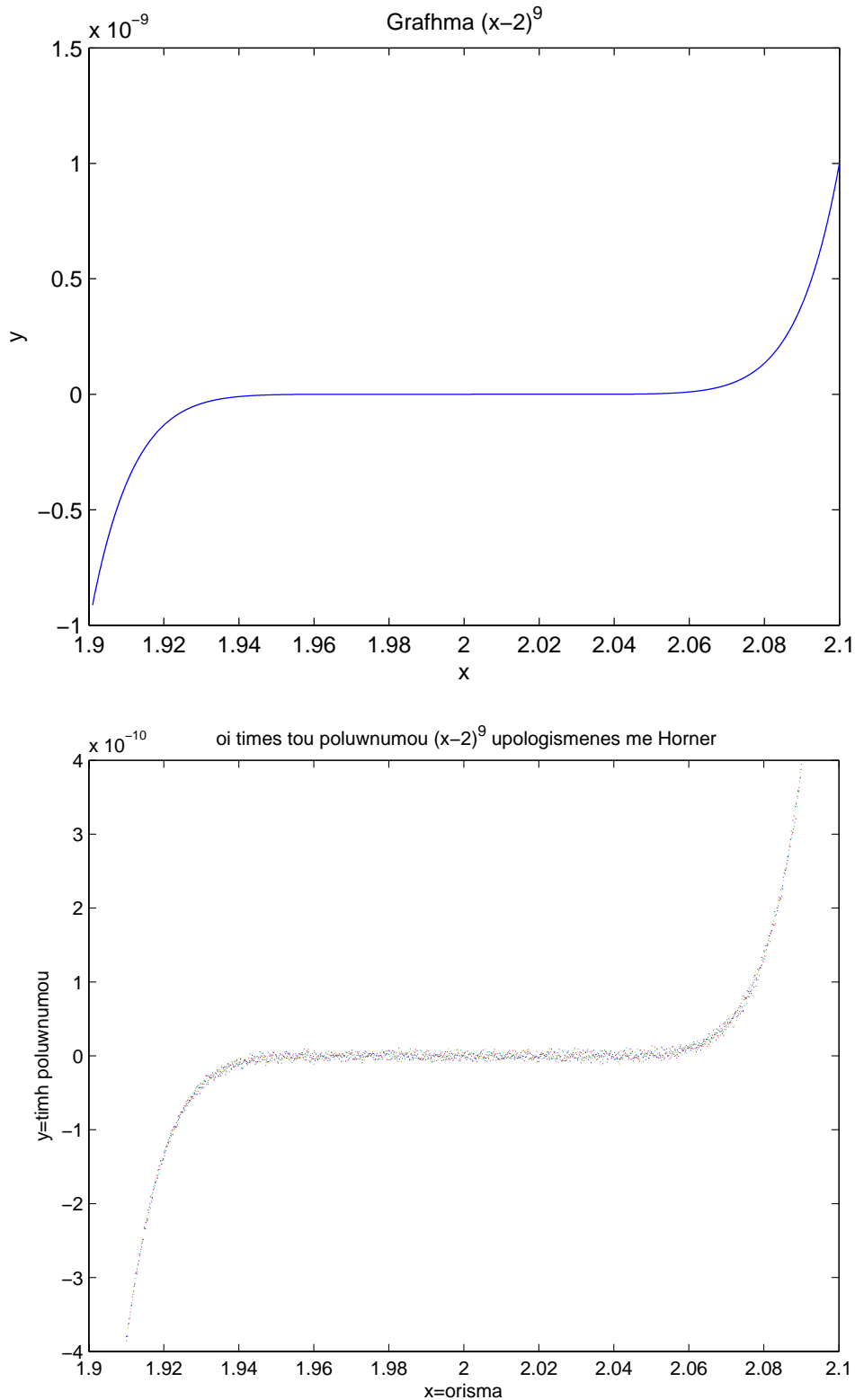
2.2. Μια βασική δυσκολία

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 πρόβλημα ιδιοτιμών. Έστω ότι έχουμε το μονικό πολυώνυμο:

$$(2.1) \quad p(z) = z^m + a_{m-1}z^{m-1} + \dots + a_1z + a_0$$

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

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



ΣΧΗΜΑ 2.1. Γραφική παράσταση των τιμών του πολυωνύμου $(x - 2)^9$ όπως υπολογίζονται στη MatLab με αριθμητική διπλής ακρίβειας.- Επάνω: Τιμές υπολογισμένες από την παραγοντοποιημένη μορφή του πολυωνύμου. Κάτω: Τιμές υπολογισμένες με το σχήμα του Horner από την ανεπτυγμένη μορφή του πολυωνύμου.

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

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

(Επίσης μπορούμε να φθάσουμε στην (2.3) απευθείας χωρίς να χρησιμοποιήσουμε την (2.2) παρατηρώντας ότι αν z είναι ρίζα του p , τότε προκύπτει από τη (2.1) ότι το $(1, z, z^2, \dots, z^{m-1})$ είναι ένα αριστερό ιδιοδιάνυσμα του A με ιδιοτιμή z .)

Τώρα η δυσκολία είναι φανερή. Είναι γνωστό ότι δεν υπάρχει συγκεκριμένος τύπος που να δίνει τις ρίζες ενός τυχαίου πολυωνύμου όταν γνωρίζουμε τους συντελεστές του. Αυτό το αποτέλεσμα αδυναμίας ρητής έκφρασης των ριζών πολυωνύμων είναι ένα από τα κορυφαία επιτεύγματα μελέτης μαθηματικών όπως ο Abel, ο Galois και άλλων το δέκατο ένατο αιώνα. Ο Abel απέδειξε το 1824 ότι δεν υπάρχει τύπος ανάλογος με αυτόν για τις ρίζες ενός πολυωνύμου δευτέρου βαθμού για πολυώνυμα βαθμού μεγαλύτερου ή ίσου του 5.

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

Αυτό το θεώρημα σημαίνει ότι ακόμα και αν μπορούσαμε να χρησιμοποιήσουμε ακριβή αριθμητική (χωρίς σφάλματα στρογγύλευσης στη μηχανή κλπ.) δε θα υπήρχε πρόγραμμα που να βρίσκει ακριβώς τις ρίζες ενός τυχαίου πολυωνύμου σε πεπερασμένου πλήθους βήματα. Έπεται ότι το ίδιο συμβαίνει και στο γενικότερο πρόβλημα υπολογισμού ιδιοτιμών πινάκων.

Αυτό δε σημαίνει βέβαια ότι δεν μπορούμε να γράψουμε ένα καλό πρόγραμμα εύρεσης ιδιοτιμών. Απλώς δεν μπορεί να βασιστεί στις ίδιες τεχνικές που χρησιμοποιούμε για να λύνουμε γραμμικά συστήματα. Μέθοδοι όπως οι ανακλάσεις Householder και η απαλοιφή του Gauss θα μπορούσαν να λύσουν γραμμικά συστήματα εξισώσεων ακριβώς, σε πεπερασμένο πλήθος βημάτων αν μπορούσαν να εφαρμοστούν σε ακριβή αριθμητική. Αντιθέτως,

Κάθε μέθοδος εύρεσης ιδιοτιμών πρέπει να είναι αναδρομική

Σκοπός μιας μεθόδου εύρεσης ιδιοτιμών είναι να δημιουργήσει ακολουθίες αριθμών οι οποίες συγκλίνουν γρήγορα στις ιδιοτιμές. Με αυτό το σκεπτικό, οι υπολογισμοί ιδιοτιμών είναι πιο αντιπροσωπευτικοί του επιστημονικού υπολογισμού από τις λύσεις γραμμικών συστημάτων εξισώσεων.

Η ανάγκη για αναδρομή μπορεί να μοιάζει αποθαρρυντική στην αρχή αλλά οι αλγόριθμοι σε αυτή την περιοχή, συγκλίνουν πάρα πολύ γρήγορα. Στις περισσότερες περιπτώσεις είναι δυνατόν να υπολογίσουμε ακολουθίες αριθμών που διπλασιάζουν η τριπλασιάζουν το πλήθος των ψηφίων ακριβείας σε κάθε βήμα. Έτσι, παρόλο που ο υπολογισμός ιδιοτιμών είναι ένα «άλυτο» πρόβλημα γενικά, πρακτικά διαφέρει από τη λύση γραμμικών συστημάτων μόνο από ένα μικρό σταθερό παράγοντα, πιο κοντά στο 1 από ότι στο 10.

Θεωρητικά μιλώντας, η εξάρτηση από τις πράξεις μετρημένη βάσει του $\epsilon_{μηχανής}$ περιλαμβάνει όρους της τάξης του $\log(|\log(\epsilon_{μηχανής})|)$

2.3. Παραγοντοποίηση Schur και διαγωνοποίηση

Οι περισσότεροι από τους γενικούς αλγόριθμους ιδιοτιμών που χρησιμοποιούνται σήμερα χρησιμοποιούν την παραγοντοποίηση Schur. Η παραγοντοποίηση Schur: $A = QTQ^*$ υπολογίζεται ως εξής: Μετασχηματίζουμε τον A σε μια ακολουθία βασικών ορθοκανονικών μετασχηματισμών ομοιότητας της μορφής $Q \mapsto Q_j^* X Q_j$, έτσι ώστε το γινόμενο:

$$(2.4) \quad Q_j^* \cdots Q_2^* Q_1^* A Q_1 Q_2 \cdots Q_j$$

που ισοδυναμεί με το γινόμενο: $Q^* A Q$ όπου $Q = Q_1 Q_2 \cdots Q_j$ να συγκλίνει σε έναν άνω τριγωνικό πίνακα T καθώς $j \rightarrow \infty$. Τότε, $T = Q^* A Q$ δηλαδή $A = QTQ^*$ που είναι η παραγοντοποίηση Schur

Αν ο A είναι πραγματικός αλλά όχι συμμετρικός, τότε γενικά μπορεί να έχει μιγαδικές ιδιοτιμές σε ζεύγη συζυγών. Σ' αυτήν την περίπτωση η Schur μορφή θα είναι μιγαδική. Έτσι ένας αλγόριθμος που υπολογίζει την παραγοντοποίηση Schur θα πρέπει να μπορεί να έχει μιγαδικά αποτελέσματα από πραγματικά δεδομένα. 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 λέγεται πραγματική παραγοντοποίηση Schur. Εδώ ο T επιτρέπεται να έχει 2×2 block πίνακες στη διαγώνιο έναν για κάθε ζεύγος συζυγών μιγαδικών ιδιοτιμών. Αυτή η περίπτωση είναι πολύ σημαντική στην πράξη.

Από την άλλη πλευρά, υποθέτουμε ότι ο A είναι ερμητιανός. Τότε το γινόμενο $Q_j^* \cdots Q_2^* Q_1^* A Q_1 Q_2 \cdots Q_j$ είναι επίσης ερμητιανός πίνακας και έτσι το όριο της συγκλίνουσας ακολουθίας είναι τριγωνικός και ερμητιανός πίνακας δηλαδή διαγώνιος. Απ' αυτό έπεται ότι οι ίδιοι αλγόριθμοι που υπολογίζουν ορθοκανονική τριγωνοποίηση ενός γενικού πίνακα, υπολογίζουν επίσης ορθοκανονική διαγωνιοποίηση ενός ερμητιανού πίνακα. Στην πράξη, δηλαδή, έτσι χειριζόμαστε την περίπτωση ερμητιανού πίνακα (ως ειδική μορφή ενός γενικού πίνακα) παρόλο που υπάρχουν αρκετές διαφοροποιήσεις που χρησιμοποιούν την ερμητιανή δομή του πίνακα σε κάθε βήμα.

2.4. Δυο φάσεις στον υπολογισμό ιδιοτιμών

Είτε είναι ο πίνακας A ερμητιανός είτε όχι, η ακολουθία (2.4) χωρίζεται συνήθως σε δυο φάσεις. Εφαρμόζουμε μια επαναληπτική διαδικασία και δημιουργούμε μια ακολουθία πινάκων μορφής Hessenberg η οποία συγκλίνει σε ένα πίνακα άνω τριγωνικό. Σχηματικά, η διαδικασία για μη ερμητιανούς πίνακες μοιάζει ως εξής:

$$(2.5) \quad \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} \longrightarrow \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix} \longrightarrow \begin{bmatrix} \times & \times & \times & \times & \times \\ & \times & \times & \times & \times \\ & & \times & \times & \times \\ & & & \times & \times \\ & & & & \times \end{bmatrix}$$

Ο πρώτος πίνακας είναι ο $A \neq A^*$ (δηλαδή ο A δεν είναι συμμετρικός), ο δεύτερος είναι ο H και ο τρίτος ο T . Από τον πρώτο πίνακα φτάνουμε στον δεύτερο μέσω της φάσης 1 ενώ από τον δεύτερο στον τρίτο μέσω της φάσης 2.

Η πρώτη φάση, απαιτεί $O(m^3)$ πράξεις. Η δεύτερη αναδρομική φάση δεν τερματίζει ποτέ και αν την αφήναμε να δουλεύει ατέρμονα, θα απαιτούσε έναν άπειρο πλήθος πράξεων. Παρόλα αυτά γενικά, σύγκλιση με ακρίβεια μηχανής επιτυγχάνεται σε $O(m)$ αναδρομές. Κάθε αναδρομή απαιτεί $O(m^2)$ πράξεις και έτσι συνολικά η φάση αυτή χρειάζεται $O(m^3)$ πράξεις. Αν ο πίνακας δεν είχε μετασχηματιστεί σε Hessenberg, τότε κάθε αναδρομή θα περιελάμβανε έναν ολόκληρο πίνακα κάτι το οποίο θα απαιτούσε $O(m^3)$ πράξεις, δηλαδή συνολικά $O(m^4)$ πράξεις ή και περισσότερες καθώς η σύγκλιση μπορεί μερικές φορές να απαιτεί περισσότερες από $O(m)$ επαναλήψεις.

Εάν ο A είναι ερμητιανός, τότε η προσέγγιση δυο φάσεων γίνεται ακόμα γρηγορότερη. Ο ενδιάμεσος πίνακας είναι σ' αυτή την περίπτωση ερμητιανός Hessenberg πίνακας δηλαδή τριδιαγώνιος. Το τελικό αποτέλεσμα είναι ένας ερμητιανός τριγωνικός πίνακας, δηλαδή διαγώνιος όπως προαναφέρθηκε. Σχηματικά για αυτή την περίπτωση:

$$(2.6) \quad \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} \longrightarrow \begin{bmatrix} \times & \times & & & \\ \times & \times & \times & & \\ & \times & \times & \times & \\ & & \times & \times & \times \\ & & & \times & \times \end{bmatrix} \longrightarrow \begin{bmatrix} \times & & & & \\ & \times & & & \\ & & \times & & \\ & & & \times & \\ & & & & \times \end{bmatrix}$$

Ο πρώτος πίνακας είναι ο $A = A^*$, ο δεύτερος είναι ο T και ο τρίτος ο D . Και εδώ από τον πρώτο πίνακα φτάνουμε στον δεύτερο μέσω της φάσης 1 ενώ από τον δεύτερο στον τρίτο μέσω της φάσης 2.

Στην περίπτωση αυτή του ερμητιανού πίνακα, πρέπει να παρατηρήσουμε ότι κάθε βήμα της φάσης 2 μπορεί να εκτελεστεί μόνο με $O(m)$ πράξεις κάνοντας τη συνολική εργασία να είναι πράξεις. Έτσι, στα προβλήματα ιδιοτιμών ερμητιανών πινάκων, είμαστε στην παράδοση περίπτωση που το άπειρο μέρος του αλγορίθμου είναι στην πράξη ταχύτερο από το πεπερασμένο τμήμα του αλγορίθμου κατά μία τάξη μεγέθους.

Απλοποίηση σε Hessenberg ή τριδιαγώνια μορφή

Περιγράφουμε τώρα την πρώτη από τις δύο φάσεις υπολογισμού που αναφέρθηκαν στην προηγούμενη ενότητα. Απλοποίηση ενός πίνακα σε μορφή Hessenberg μέσω μιας ακολουθίας ορθοκανονικών μετασχηματισμών ομοιότητας. Αν ο αρχικός πίνακας είναι Ερμητιανός, τότε το αποτέλεσμα είναι τριδιαγώνιος πίνακας.

3.1. Μια κακή ιδέα

Για να υπολογίσουμε την παραγοντοποίηση Schur, $A = QTQ^*$, θα θέλαμε να εφαρμόσουμε ορθοκανονικούς μετασχηματισμούς ομοιότητας στον A έτσι ώστε να εισάγουμε μηδενικά κάτω από τη διαγώνιο. Μια πρώτη ιδέα θα ήταν να χρησιμοποιήσουμε μετασχηματισμούς Householder για να εισάγουμε μηδενικά το ένα μετά το άλλο και να επιτύχουμε ευθεία διαγωνοποίηση.

Ο πρώτος μετασχηματισμός Householder $(Q_1)^*$ πολλαπλασιασμένος εξ αριστερών του A , δίδει μηδενικά κάτω από τη διαγώνιο στην πρώτη στήλη του A . Με τη διαδικασία αυτή αλλάζουν όλες οι γραμμές του πίνακα.

Σχηματικά, αυτό περιγράφεται όπως φαίνεται παρακάτω, όπου ο πρώτος πίνακας είναι ο A και ο δεύτερος ο $(Q_1)^*A$.

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} \longrightarrow \begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix}$$

Για την ολοκλήρωση του μετασχηματισμού ομοιότητας απαιτείται ο πολλαπλασιασμός με τον Q_1 εκ δεξιών του $(Q_1)^*A$. Τότε, κάθε στήλη αντικαθίσταται από ένα γραμμικό συνδυασμό όλων των στηλών και ως αποτέλεσμα έχουμε την καταστροφή των μηδενικών που είχαν δημιουργηθεί με τον πρώτο πολλαπλασιασμό και να καταλήγουμε σε ένα πρόβλημα αντίστοιχο του αρχικού. Σχηματικά, αυτό περιγράφεται όπως φαίνεται παρακάτω, όπου ο πρώτος πίνακας είναι ο $(Q_1)^*A$ και ο δεύτερος ο $(Q_1)^*AQ_1$.

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix} \longrightarrow \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix}$$

Φυσικά αυτή η ιδέα ήταν καταδικασμένη λόγω της βασικής δυσκολίας που περιεγράφη στην προηγούμενη ενότητα. Καμία πεπερασμένη διαδικασία δεν μπορεί να υπολογίσει ακριβώς τις ιδιοτιμές ενός πίνακα A . Όμως, αυτή η απλή στρατηγική, που μοιάζει μάταια, μειώνει το μέγεθος των στοιχείων κάτω από τη διαγώνιο.

3.2. Μια καλή ιδέα

Μια παραλλαγή της παραπάνω μεθόδου είναι να επιλέξουμε ένα μετασχηματισμό Householder Q_1^* που πολλαπλασιαζόμενος από αριστερά δίνει μηδενικά στην πρώτη στήλη, από το τρίτο στοιχείο και μετά. Στη συνέχεια, αν ο $Q_1^*)^*A$ πολλαπλασιαστεί με τον Q_1 από τα δεξιά, αφήνει ανεπηρέαστη την πρώτη στήλη, οπότε και τα μηδενικά που δημιουργήθηκαν.

Σχηματικά, αυτό περιγράφεται όπως φαίνεται παρακάτω, όπου ο πρώτος πίνακας είναι ο A , ο δεύτερος είναι ο Q_1^*A αφού έχει γίνει η εφαρμογή του Q_1^* και ο τρίτος ο $Q_1^*AQ_1$ μετά από εφαρμογή του πίνακα Q_1 στον Q_1^*A από δεξιά

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \end{bmatrix} \longrightarrow \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix} \longrightarrow \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix}$$

Αυτή η ιδέα επαναλαμβάνεται για να εισάγει μηδενικά στις επόμενες στήλες. Για παράδειγμα, ο δεύτερος μετασχηματισμός Householder, ο Q_2 , αφήνει την πρώτη και τη δεύτερη στήλη αναλλοίωτες. Παρακάτω είναι η περιγραφή και του δεύτερου αυτού βήματος, όπου ο αρχικός πλέον πίνακας είναι ο $Q_1^*AQ_1$ ο οποίος μετατρέπεται στον δεύτερο πίνακα, $Q_2^*Q_1^*AQ_1$, μετά από εφαρμογή του Q_2^* από αριστερά και στη συνέχεια αυτός μετατρέπεται στον τρίτο πίνακα, $Q_2^*Q_1^*AQ_1Q_2$, ο οποίος είναι όμοιος με τον αρχικό πίνακα, μετά από εφαρμογή του Q_2 από δεξιά.

$$\begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \end{bmatrix} \longrightarrow \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & \times & \times & \times \end{bmatrix} \longrightarrow \begin{bmatrix} \times & \times & \times & \times & \times \\ \times & \times & \times & \times & \times \\ 0 & \times & \times & \times & \times \\ 0 & 0 & \times & \times & \times \\ 0 & 0 & \times & \times & \times \end{bmatrix}$$

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

$$H = Q^*AQ = Q_{m-2}^* \cdots Q_2^*Q_1^*AQ_1Q_2 \cdots Q_{m-2}.$$

$$\begin{bmatrix} \times & \times & \cdots & \times & \times & \times \\ \times & \times & \cdots & \times & \times & \times \\ 0 & \times & \cdots & \times & \times & \times \\ \vdots & \ddots & \ddots & & \vdots & \vdots \\ \vdots & & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & \times & \times \end{bmatrix}$$

Ο αλγόριθμος καταγράφεται παρακάτω:

Αναγωγή τύπου Householder ενός πίνακα A , σε μορφή Hessenberg

$$\left\{ \begin{array}{l} \text{Για } k = 1 \text{ έως } m - 2 \\ x = A_{k+1:m,k} \\ u_k = \text{sign}(x_1) \|x\|_2 e_1 + x \\ u_k = u_k / \|u_k\|_2 \\ A_{k+1:m,k:m} = A_{k+1:m,k:m} - 2u_k(u_k^* A_{k+1:m,k:m}) \\ A_{1:m,k+1:m} = A_{1:m,k+1:m} - 2(A_{1:m,k+1:m} u_k) u_k^* \end{array} \right.$$

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

3.3. Μέτρηση πράξεων

Η μέθοδος αυτή διαχωρίζει τον A σε δύο υποπίνακες. Για τον πρώτο υποπίνακα εφαρμόζει έναν ανακλαστή Householder στα αριστερά του πίνακα A . Ο k -οστός τέτοιος ανακλαστής επηρεάζει τις τελευταίες $m - k$ γραμμές. Από τη στιγμή που ο ανακλαστής εφαρμόζεται, αυτές οι γραμμές έχουν μηδενικά στις πρώτες $k - 1$ στήλες και χρειάζεται να εκτελέσουμε μόνο τις πράξεις που αφορούν τα τελευταία $m - k + 1$ στοιχεία κάθε γραμμής. Για αυτούς τους υπολογισμούς έχουμε συνολικά $\frac{4}{3}m^3$ πράξεις.

Για τον δεύτερο υποπίνακα, εφαρμόζει έναν ανακλαστή Householder στα δεξιά του πίνακα A . Στο k -οστό βήμα, ο ανακλαστής διαμορφώνει γραμμικούς συνδυασμούς των τελευταίων $m - k$ στηλών. Αυτό χρειάζεται περισσότερη δουλειά από τη δημιουργία του πρώτου υποπίνακα διότι δεν υπάρχουν μηδενικά που μπορούν να αγνοηθούν. Επομένως, πρέπει να υπολογίσουμε και τα m στοιχεία κάθε στήλης από αυτές που αλλάζουν δηλαδή συνολικά $m(m - k)$ στοιχεία για κάθε τιμή του k . Για αυτούς τους υπολογισμούς κάνουμε συνολικά $2m^3$ πράξεις.

Συνολικά δηλαδή η μετατροπή ενός $m \times m$ πίνακα σε πίνακα Hessenberg απαιτεί εργασία $\frac{10}{3}m^3$ πράξεων.

3.4. Η περίπτωση Ερμητιανών πινάκων: Μετατροπή σε τριδιαγώνια μορφή

Εάν ο A είναι ερμητιανός πίνακας, ο αλγόριθμος που μόλις περιγράφηκε θα μετατρέψει τον A σε τριδιαγώνια μορφή (τουλάχιστον όταν δεν υπάρχουν λάθη στρογγυλοποίησης). Αυτό συμβαίνει επειδή όταν ο A είναι ερμητιανός, μετατρέπεται από τον αλγόριθμο σε έναν πίνακα $Q^* A Q$ ο οποίος είναι επίσης ερμητιανός αλλά και Hessenberg δηλαδή τριδιαγώνιος.

Καθώς τώρα εισάγονται μηδενικά και σε γραμμές και σε στήλες, μπορούμε να μειώσουμε τις πράξεις, αγνοώντας αυτά τα επιπρόσθετα μηδενικά. Μάυτό το σκεπτικό, στη συγκεκριμένη περίπτωση, ένας ανακλαστής Householder είναι εξίσου φθηνός σε πράξεις όταν εφαρμοστεί στα δεξιά ενός πίνακα και στα αριστερά του ίδιου πίνακα. Το συνολικό κόστος σε πράξεις είναι τότε για την εφαρμογή του ανακλαστή από δεξιά $\frac{4}{3}m^3$. Το ίδιο και για την εφαρμογή του από αριστερά. Επομένως, για ολόκληρη τη μετατροπή χρειαζόμαστε $\frac{8}{3}m^3$ πράξεις.

Αυτό το αποτέλεσμα βασίζεται στην αραιότητα του νέου πίνακα και όχι στη συμμετρικότητά του. Η συμμετρικότητα μας βοηθά στο ότι σε κάθε βήμα του υπολογισμού, ο πίνακας είναι ερμητιανός επομένως αρκεί να υπολογίσουμε τα μισά στοιχεία του πίνακα επομένως το συνολικό κόστος μειώνεται σε $\frac{4}{3}m^3$.

3.5. Σταθερότητα

Ο αλγόριθμος που μόλις περιγράφηκε είναι οπισθοδρομικά σταθερός.

Το παρακάτω θεώρημα μπορεί να αποδειχθεί.

ΘΕΩΡΗΜΑ 3.1. Έστω η απλοποίηση Hessenberg $A = QHQ^*$ ενός πίνακα $A \in \mathbb{C}^{m \times m}$ υπολογισμένη από τον παραπάνω αλγόριθμο σε έναν υπολογιστή που ικανοποιεί τα παρακάτω δυο αξιώματα:

ΑΞΙΩΜΑ 3.2. Για κάθε $x \in \mathbb{R}$ υπάρχει ϵ τέτοιο ώστε: $|\epsilon| \leq \epsilon_{\text{μηχανής}}$ και $fl(x) = x(1 + \epsilon)$ όπου: $fl: \mathbb{R} \rightarrow F$ μια απεικόνιση που αντιστοιχεί πραγματικούς αριθμούς στους πιο κοντινούς τους ρητούς.

ΑΞΙΩΜΑ 3.3. Για κάθε $x, y \in \mathbb{R}$ υπάρχει ϵ τέτοιο ώστε: $|\epsilon| \leq \epsilon_{\text{μηχανής}}$ και $x \otimes y = (x * y)(1 + \epsilon)$ όπου: $fl: \mathbb{R} \rightarrow F$ η ίδια απεικόνιση όπως παραπάνω.

Έστω επίσης \tilde{H} ο πίνακας Hessenberg που υπολογίζεται με αριθμητική κινητής υποδιαστολής και έστω \tilde{Q} ορθομοναδιαίος πίνακας που αντιστοιχεί στα διανύσματα-ανακλαστές \tilde{v}_k υπολογισμένα επίσης με αριθμητική κινητής υποδιαστολής. Τότε έχουμε:

$$(3.1) \quad \tilde{Q}\tilde{H}\tilde{Q}^* = A + \delta A \frac{\|\delta A\|}{\|A\|} = O(\epsilon_{\text{μηχανής}})$$

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

Πηλίο Rayleigh, Αντίστροφη Αναδρομή

Σε αυτό το κεφάλαιο παρουσιάζονται μερικοί κλασικοί αλγόριθμοι εύρεσης ιδιοτιμών. Αυτοί είναι πολύ χρήσιμοι σε συγκεκριμένες περιπτώσεις, ειδικά στην αντίστροφη αναδρομή που είναι μια καθιερωμένη μέθοδος για τον προσδιορισμό ιδιοδιανυσμάτων όταν οι αντίστοιχες ιδιοτιμές είναι γνωστές.

4.1. Περιορισμός σε πραγματικούς συμμετρικούς πίνακες

Οι μέθοδοι που θα παρουσιαστούν σε αυτό το κεφάλαιο εφαρμόζονται είτε σε γενικούς πίνακες είτε σε ερμητιανούς ($A = A^*$). Για ευκολία της παρουσίασης αυτών των αλγορίθμων, θα περιοριστούμε σε πραγματικούς και συμμετρικούς πίνακες και στην ευκλείδεια νόρμα στο χώρο \mathbb{R}^m την οποία θα συμβολίζουμε $\|\cdot\| = \|\cdot\|_2$ και τη φυσική νόρμα που επάγεται από αυτήν την οποία θα συμβολίζουμε ως $\|\cdot\|$.

Έτσι 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 παρακάτω συμβολισμό:

πραγματικές ιδιοτιμές: l_1, \dots, l_m ,

ορθοκανονικά διανύσματα: q_1, \dots, q_m .

Τα ιδιοδιανύσματα θεωρούνται κανονικοποιημένα δηλαδή: $\|q_j\| = 1$ και οι ιδιοτιμές θα διαταχθούν όπως χρειάζεται σε κάθε περίπτωση.

Οι περισσότερες από τις ιδέες που θα περιγραφούν στα επόμενα κεφάλαια ανήκουν στη δεύτερη φάση όπως αυτή περιγράφηκε στο δεύτερο κεφάλαιο. Αυτό σημαίνει ότι από τη στιγμή που εφαρμόζουμε αυτές τις ιδέες, ο A δε θα είναι απλώς πραγματικός και συμμετρικός αλλά τριδιαγώνιος. Αυτή η τριδιαγώνια μορφή έχει μαθηματική σημασία. Για παράδειγμα στην επιλογή εναλλαγών στον αλγόριθμο QR και έχει αλγοριθμική σημασία καθώς μειώνει αρκετά τις πράξεις σε κάθε βήμα από $O(m^3)$ σε $O(m)$.

4.2. Πηλίο Rayleigh

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

$$(4.1) \quad r(x) = \frac{x^T A x}{x^T x}$$

Σημειώνουμε ότι αν το x είναι ιδιοδιάνυσμα με αντίστοιχη ιδιοτιμή λ , τότε το

$$r(x) = \frac{x^T \lambda x}{x^T x} = \lambda.$$

Ένας τρόπος για να χρησιμοποιήσουμε αυτόν τον τύπο είναι να αναρωτηθούμε δεδομένου του διανύσματος x , ποιος αριθμός α συμπεριφέρεται "περισσότερο ως ιδιοτιμή" του x με την έννοια ότι ελαχιστοποιεί την $\|Ax - \alpha x\|$. Έχουμε δηλαδή ένα $m \times 1$ πρόβλημα ελαχίστων τετραγώνων της μορφής $x\alpha \approx Ax$ (όπου x είναι ο πίνακας, α το άγνωστο

διάνυσμα, Ax το δεξί μέλος). Γράφοντας την κανονική εξίσωση για αυτό το σύστημα: $x^T x \alpha = x^T Ax$ προκύπτει ότι $\alpha = r(x)$. Επομένως, το $r(x)$ μπορεί να θεωρηθεί ως μια φυσική εκτίμηση ιδιοτιμής αν το x είναι σχεδόν αλλά όχι ακριβώς ιδιοδιάνυσμα.

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

$$\begin{aligned} \frac{\partial r(x)}{\partial x_j} &= \\ \frac{[\frac{\partial}{\partial x_j}(x^T Ax)]x^T x - (x^T Ax)[\frac{\partial}{\partial x_j}(x^T x)]}{(x^T x)^2} &= \\ \frac{\frac{\partial}{\partial x_j}(x^T Ax)}{x^T x} - \frac{(x^T Ax)\frac{\partial}{\partial x_j}(x^T x)}{(x^T x)^2} &= \\ \frac{2(Ax)_j}{x^T x} - \frac{(x^T Ax)2x_j}{(x^T x)^2} &= \\ \frac{2}{x^T x}(Ax - r(x)x)_j & \end{aligned}$$

Εάν μαζέψουμε όλες τις μερικές παραγώγους σε ένα m -διάνυσμα, έχουμε υπολογίσει την κλίση του $r(x)$, που συμβολίζεται ως: $\nabla r(x)$.

Έχουμε δείξει επομένως ότι:

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

Από αυτόν τον τύπο βλέπουμε ότι για κάθε ιδιοδιάνυσμα x του A , $\nabla r(x) = 0$, διότι $r(x) = \lambda$ και $Ax = \lambda x$. Αντιστρόφως, εάν $\nabla r(x) = 0$ με $x \neq 0$, τότε το x είναι ένα ιδιοδιάνυσμα και το $r(x)$ η αντίστοιχη ιδιοτιμή.

Από γεωμετρική άποψη, τα ιδιοδιανύσματα του A είναι τα τοπικά ακρότατα της συνάρτησης $r(x)$ και οι ιδιοτιμές του A είναι οι τιμές του $r(x)$ σε αυτά τα σημεία. Μπορούμε να αναπαραστήσουμε τα ιδιοδιανύσματα x του A ως σημεία του \mathbb{R}^m που ανήκουν σε ευθείες οι οποίες διέρχονται από την αρχή των αξόνων. Αν τα κανονικοποιήσουμε, δηλαδή $\|x\|_2 = 1$, αυτά θα είναι σημεία στην επιφάνεια της μοναδιαίας σφαίρας του \mathbb{R}^m . Έστω q_j ένα ιδιοδιάνυσμα του A . Από το γεγονός ότι $\nabla r(q_j) = 0$ και από την ομαλότητα της συνάρτησης $r(x)$ παντού εκτός από το σημείο $x = 0$ προκύπτει το ακόλουθο σημαντικό αποτέλεσμα.

$$(4.3) \quad r(x) - r(q_j) = O(\|x - q_j\|^2) \text{ καθώς } x \rightarrow q_j$$

Επομένως το πηλίκο Rayleigh είναι μια τετραγωνικά ακριβής-ορθή εκτίμηση μιας ιδιοτιμής. Από αυτό προκύπτει η δύναμή του.

Ένας πιο σαφής τρόπος να προκύψει η 4.3 είναι να επεκτείνουμε το x ως γραμμικό συνδυασμό των ιδιοδιανυσμάτων q_1, q_2, \dots, q_m του A . Εάν $x = \sum_{j=1}^m a_j q_j$, τότε $r(x) = \frac{x^T Ax}{x^T x} = \frac{\sum_{j=1}^m a_j^2 \lambda_j}{\sum_{j=1}^m a_j^2}$. Επομένως, το $r(x)$ είναι ο σταθμικός μέσος των ιδιοτιμών του A με βάρη ίσα των τετραγώνων των συντεταγμένων του x στη βάση των ιδιοδιανυσμάτων. Εξαιτίας αυτού του τετραγωνισμού των συντεταγμένων, προκύπτει ότι εάν $|\frac{a_j}{a_J}| \leq \epsilon$ για κάθε $j \neq J$, τότε $r(x) - r(q_J) = O(\epsilon^2)$.

4.3. Μέθοδος Δυνάμεων

Τώρα αλλάζουμε λίγο πορεία. Υποθέτουμε ότι $v^{(0)}$ είναι ένα διάνυσμα με $\|v^{(0)}\| = 1$. Η διαδικασία που ακολουθεί, η μέθοδος των δυνάμεων, είχε χαρακτηριστεί νωρίτερα ως μία όχι και τόσο καλή ιδέα λόγω του χρόνου που απαιτεί. Με τη μέθοδο αυτή, περιμένει κανείς να παραχθεί μια ακολουθία διανυσμάτων η οποία συγκλίνει σε ιδιοδιάνυσμα που αντιστοιχεί στη μεγαλύτερη ιδιοτιμή του πίνακα A .

Μέθοδος Δυνάμεων

$$\left\{ \begin{array}{l} v^{(0)} = \text{κάποιο διάνυσμα με } \|v^{(0)}\| = 1 \\ \text{Για } k = 1, 2, \dots \\ w = Av^{(k-1)} \\ v^{(k)} = \frac{w}{\|w\|} \\ \lambda^{(k)} = (v^{(k)})^T Av^{(k)} \end{array} \right.$$

όπου στην πρώτη σειρά της επανάληψης εφαρμόζουμε τον A στο κανονικοποιημένο διάνυσμα που έχουμε από την προηγούμενη επανάληψη, στη δεύτερη σειρά, κανονικοποιούμε το καινούριο διάνυσμα και στην τρίτη, βρίσκουμε το πηλίκο του Rayleigh διότι $\lambda^{(k)} = (v^{(k)})^T Av^{(k)} = \frac{w^T Aw}{\|w\|^2} = \frac{w^T Aw}{w^T w}$

Σε αυτόν αλλά και στους αλγορίθμους που ακολουθούν, δεν ασχολούμαστε με κριτήρια τερματισμού (αναφερόμαστε στις επαναλήψεις του βρόγχου μόνο με το: Για $k = 1, 2, \dots$). Στην πράξη τα κριτήρια τερματισμού είναι πολύ σημαντικά. Κάποια από αυτά θα αναφερθούν στο κεφάλαιο όπου παρουσιάζονται τα αριθμητικά παραδείγματα.

Μπορούμε να αναλύσουμε τη Μέθοδο των Δυνάμεων με τον εξής τρόπο: Γράφουμε το $v^{(0)}$ ως γραμμικό συνδυασμό των ορθοκανονικών ιδιοδιανυσμάτων q_i :

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

Επειδή το $v^{(k)} = A^k v^{(0)}$ έχουμε για κάποια σταθερά c_k (που εξαρτάται από τις νόρμες των διανυσμάτων που έχουν υπολογιστεί ως προσεγγίσεις των ιδιοδιανυσμάτων από τον αλγόριθμο στα προηγούμενα βήματα)

$$\begin{aligned} v^{(k)} &= c_k A^k v^{(0)} \\ (4.4) \quad &= c_k (a_1 \lambda_1^k q_1 + a_2 \lambda_2^k q_2 + \dots + a_m \lambda_m^k q_m) \\ &= c_k \lambda_1^k (a_1 q_1 + a_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k q_2 + \dots + a_m \left(\frac{\lambda_m}{\lambda_1}\right)^k q_m) \end{aligned}$$

Από αυτό προκύπτει το ακόλουθο συμπέρασμα.

ΘΕΩΡΗΜΑ 4.1. Υποθέτουμε ότι $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_m| \geq 0$ και $q_1^T v^{(0)} \neq 0$. Τότε οι επαναλήψεις του αλγορίθμου της μεθόδου των δυνάμεων ικανοποιούν:

$$(4.5) \quad \|v^{(k)} - (\pm q_1)\| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right) \text{ και } |\lambda^{(k)} - \lambda_1| = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^{2k}\right)$$

καθώς $k \rightarrow \infty$.

Το σύμβολο \pm σημαίνει ότι σε κάθε βήμα k μπορεί να γίνει οποιαδήποτε από τις δύο επιλογές προσήμου και το φράγμα να ισχύει Αυτό εξαρτάται από τη σχετική θέση του αρχικού

διανύσματος που χρησιμοποιούμε με το ορθοκανονικό ιδιοδιάνυσμα q_1 που αντιστοιχεί στη μεγαλύτερη και ' απόλυτη τιμή ιδιοτιμή.

Απόδειξη. Για να απλοποιήσουμε την απόδειξη χρησιμοποιούμε μόνο θετικές ιδιοτιμές, και υποθέτουμε ότι το εσωτερικό γινόμενο $q_1^T v^{(0)}$ είναι γνήσια θετικό. Το θεώρημα ισχύει και για τις υπόλοιπες περιπτώσεις αλλά η απόδειξη είναι πιο περίπλοκη και δε θα ασχοληθούμε εδώ με αυτές.

Έστω λοιπόν η τυχαία αρχική προσέγγιση ιδιοδιανύσματος

$$v^{(0)} = a_1 q_1 + a_2 q_2 + \dots + a_m q_m$$

με τη μορφή γραμμικού συνδυασμού ορθοκανονικών ιδιοδιανυσμάτων και με μέτρο

1.

Από το γεγονός ότι: $q_1^T v^{(0)} > 0$ και το ότι: $q_1^T v^{(0)} = a_1$ έχουμε ότι: $a_1 > 0$.

Επίσης, από τη φθίνουσα διάταξη των ιδιοτιμών που έχουμε υποθέσει, έχουμε:

$\lambda_1 > \lambda_2 \geq \dots \geq \lambda_m \geq 0$ και παίρνουμε ότι $\lambda_1 > 0$ και ότι

$$\frac{\lambda_j}{\lambda_1} \leq \frac{\lambda_2}{\lambda_1} < 1 \text{ για κάθε } j \in 2, 3, \dots, m$$

Ισχύει ότι: $\|v^{(0)}\| = 1 \Rightarrow \|v^{(0)}\|^2 = 1 \Rightarrow \sum_{j=1}^m a_j^2 = 1$

Ακόμα, από τον τρόπο κατασκευής του διανύσματος $v^{(k)}$ από τις προσεγγίσεις του ιδιοδιανύσματος που έχουν προκύψει από τα προηγούμενα βήματα της επανάληψης στον αλγόριθμο, ισχύει ότι: $v^{(k)} = c_k A^k v^{(0)}$ με $\|v^{(k)}\| = 1$ άρα $c_k^{-1} = \|A^k v^{(0)}\|$ και $\|v^{(k)}\|^2 = \|c_k A^k v^{(0)}\|^2 = 1 \Rightarrow \sum_{j=1}^m c_k^2 a_j^2 \lambda_j^{2k} = 1$

$$\text{Επομένως, } c_k a_1 \lambda_1^k = \frac{1}{\sqrt{1 + \frac{a_2^2}{a_1^2} \left(\frac{\lambda_2}{\lambda_1}\right)^{2k} + \dots + \frac{a_m^2}{a_1^2} \left(\frac{\lambda_m}{\lambda_1}\right)^{2k}}}$$

Ορίζοντας τώρα το τμήμα της υπορίζου ποσότητας

$$M = \frac{a_2^2}{a_1^2} \left(\frac{\lambda_2}{\lambda_1}\right)^{2k} + \dots + \frac{a_m^2}{a_1^2} \left(\frac{\lambda_m}{\lambda_1}\right)^{2k},$$

$$\text{έχουμε: } c_k a_1 \lambda_1^k = \frac{1}{\sqrt{1+M}}$$

Σύμφωνα με τον τύπο του Taylor, έχουμε πως

$$\frac{1}{\sqrt{1+M}} = 1 - \frac{1}{2}M + O(M^2),$$

οπότε κάνοντας πράξεις, προκύπτει το ακόλουθο:

$$\begin{aligned} \|v^{(k)} - q_1\|^2 &= \\ (c_k a_1 \lambda_1^k - 1)^2 + \sum_{j=2}^m m c_k^2 a_j^2 \lambda_j^{2k} &= \\ 1 - 2c_k a_1 \lambda_1^k + 1 + \sum_{j=2}^m m c_k^2 a_j^2 \lambda_j^{2k} &= \\ 1 - 2c_k a_1 \lambda_1^k + 1 = 2 - 2c_k a_1 \lambda_1^k &= \\ 2 - 2\left(1 - \frac{1}{2}M + O(M^2)\right) &= \\ M + O(M^2) &= O(M). \end{aligned}$$

Όμως, έχουμε ορίσει το

$$M = \frac{a_2^2}{a_1^2} \left(\frac{\lambda_2}{\lambda_1}\right)^{2k} + \dots + \frac{a_m^2}{a_1^2} \left(\frac{\lambda_m}{\lambda_1}\right)^{2k}$$

δηλαδή

$$M = O\left(\left(\frac{\lambda_2}{\lambda_1}\right)^{2k}\right).$$

Άρα, τελικά,

$$\|v^{(k)} - q_1\|^2 = O\left(\left(\frac{\lambda_2}{\lambda_1}\right)^{2k}\right) \Rightarrow \|v^{(k)} - q_1\| = O\left(\left(\frac{\lambda_2}{\lambda_1}\right)^k\right)$$

που είναι και το ζητούμενο.

Για την απόδειξη της δεύτερης σχέσης, χρησιμοποιούμε το πηλίκο του Rayleigh καθώς $\lambda^{(k)} = r(v^{(k)})$ και $\lambda_1 = r(q_1)$. Βάσει αυτού, της σχέσης που μόλις αποδείχθηκε και της σχέσης (4.3) προκύπτει άμεσα και το δεύτερο ζητούμενο. \square

Όσον αφορά τη Μέθοδο των Δυνάμεων, αυτή είναι περιορισμένης χρήσης για διάφορους λόγους. Πρώτον, η μέθοδος αυτή μπορεί να βρει μόνο το ιδιοδιάνυσμα που αντιστοιχεί στη μεγαλύτερη ιδιοτιμή. Δεύτερον, η σύγκλιση είναι γραμμική μειώνοντας το σφάλμα σε κάθε επαναληψη κατά ένα σταθερό όρο μόνο $\approx [\frac{\lambda_2}{\lambda_1}]$. Τέλος, αυτός ο σταθερός όρος 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 αμβλύνουμε τις διαφορές μεταξύ των ιδιοτιμών όσον αφορά τις τιμές τους.

4.4. Μέθοδος Αντίστροφης Αναδρομής

Για κάθε $\mu \in \mathbb{R}$ 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 οι: $(\lambda_j - \mu)^{-1}$, όπου λ_j είναι οι ιδιοτιμές του A . Έτσι, συνδυάζοντας αυτό με τα όσα είπαμε για τον παραπάνω αλγόριθμο, μπορούμε να κάνουμε το εξής: Υποθέτουμε ότι το μ είναι κοντά σε μια ιδιοτιμή λ_j του A . Τότε μπορεί για κάποιο J να ισχύει ότι το $(\lambda_J - \mu)^{-1}$ είναι πολύ μεγαλύτερο από τα $(\lambda_j - \mu)^{-1}$ για κάθε $j \neq J$. Επομένως, αν εφαρμόσουμε τη μέθοδο των δυνάμεων στον $(A - \mu I)^{-1}$, η διαδικασία θα συγκλίνει γρήγορα στο q_j που είναι ιδιοδιάνυσμα και των δύο πινάκων. Αυτή η μέθοδος καλείται αντίστροφη αναδρομή.

Μέθοδος Αντίστροφης αναδρομής

$$\left\{ \begin{array}{l} v^{(0)} = \text{κάποιο διάνυσμα με } \|v^{(0)}\| = 1 \\ \text{Για } k = 1, 2, \dots \\ \text{Λύνουμε το σύστημα: } (A - \mu I)w = v^{(k-1)} \text{ ως προς το } w \\ v^{(k)} = \frac{w}{\|w\|} \\ \lambda^{(k)} = (v^{(k)})^T A v^{(k)} \end{array} \right.$$

Εδώ στην πρώτη σειρά της επανάληψης ουσιαστικά εφαρμόζουμε τον πίνακα

$(A - \mu I)^{-1}$ στο διάνυσμα της προηγούμενης επανάληψης καθώς: $(A - \mu I)w = v^{(k-1)}$ ισοδυναμεί με $w = (A - \mu I)^{-1}v^{(k-1)}$. Στη δεύτερη σειρά κανονικοποιούμε το διάνυσμα που βρήκαμε και στην τρίτη βρίσκουμε το πηλίκο Rayleigh αυτού του διανύσματος.

Μπορεί τώρα να αναρωτηθεί κανείς τί γίνεται όταν το μ είναι ιδιοτιμή του αρχικού πίνακα A οπότε ο πίνακας $A - \mu I$ είναι μη αντιστρέψιμος ή όταν το μ είναι σχεδόν ιδιοτιμή του A οπότε δεν μπορούμε να βρούμε ακριβή λύση του $(A - \mu I)w = v^{(k-1)}$. Η αντίστροφη αναδρομή δουλεύει σ' αυτές τις περιπτώσεις χωρίς πρόβλημα. Όπως η μέθοδος των δυνάμεων, η αντίστροφη αναδρομή συγκλίνει γραμμικά. Όμως αντίθετα με τη μέθοδο των δυνάμεων, μπορούμε να επιλέξουμε ένα ιδιοδιάνυσμα το οποίο βρίσκουμε χρησιμοποιώντας μια εκτίμηση μ της αντίστοιχης ιδιοτιμής. Επιπλέον ο ρυθμός της γραμμικής σύγκλισης μπορεί να ελεγχθεί καθώς εξαρτάται από το μ . Εάν το μ είναι αρκετά πιο κοντά σε μία ιδιοτιμή του A από όσο στις άλλες, τότε η μεγαλύτερη ιδιοτιμή του πίνακα $(A - \mu I)^{-1}$ θα είναι πολύ μεγαλύτερη από τις υπόλοιπες. Κάνοντας χρήση της ιδίας επιχειρηματολογίας με τη μέθοδο των δυνάμεων, προκύπτει το παρακάτω Θεώρημα.

ΘΕΩΡΗΜΑ 4.2. Υποθέτουμε ότι το λ_J είναι η πιο κοντινή στο μ ιδιοτιμή και ότι το λ_K είναι η αμέσως επόμενη. Δηλαδή $|\mu - \lambda_J| < |\mu - \lambda_K| \leq |\mu - \lambda_j|$ για κάθε $j \neq J$. Επιπλέον υποθέτουμε ότι $q_J^T v^{(0)} \neq 0$. Τότε οι επαναλήψεις του αλγορίθμου αντίστροφης αναδρομής ικανοποιούν:

$$\|v^{(k)} - (\pm q_J)\| = O\left(\left|\frac{\mu - \lambda_J}{\mu - \lambda_K}\right|^k\right)$$

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

καθώς $k \rightarrow \infty$. Το σύμβολο \pm σημαίνει το ίδιο με αυτό που ισχυε στο Θεώρημα 4.1 δηλαδή ότι σε κάθε βήμα k μπορεί να γίνει οποιαδήποτε από τις δύο επιλογές προσήμου και το φράγμα να ισχύει.

Η αντίστροφη αναδρομή είναι πολύ χρήσιμη στην αριθμητική γραμμική άλγεβρα γιατί είναι μια κλασική μέθοδος υπολογισμού ενός ή περισσότερων ιδιοδιανυσμάτων ενός πίνακα όταν οι ιδιοτιμές είναι γνωστές. Σε αυτή την περίπτωση, ο αλγόριθμος την αντίστροφης αναδρομής εφαρμόζεται ως έχει χωρίς όμως να υπολογίζεται το πηλίκο του Rayleigh.

4.5. Επαναληπτική Μέθοδος πηλίκου Rayleigh

Μέχρι στιγμής, έχει παρουσιαστεί μία μέθοδος εύρεσης προσεγγίσεων ιδιοτιμών από εκτιμήσεις ιδιοδιανυσμάτων (Πηλίκο του Rayleigh) και μία μέθοδος εύρεσης προσεγγίσεων ιδιοδιανυσμάτων από εκτιμήσεις ιδιοτιμών (Αντίστροφη αναδρομή).

Ο συνδυασμός των δύο αυτών μεθόδων χρησιμοποιείται για να βρούμε συνεχώς καλύτερες εκτιμήσεις ιδιοτιμών και ιδιοδιανυσμάτων ως εξής: Από μια εκτίμηση μιας ιδιοτιμής λ_j μπορούμε μέσω της μεθόδου αντίστροφης αναδρομής να βρούμε μια εκτίμηση για ένα ιδιοδιάνυσμα q_j και μέσω αυτού (με το πηλίκο του Rayleigh) να βρούμε μια νέα εκτίμηση του λ_j . Για να γίνει όμως αυτό, χρειαζόμαστε μία πρώτη εκτίμηση ιδιοδιανύσματος-τος q_j .

Ο αλγόριθμος αυτός καλείται: Επαναληπτική Μέθοδος πηλίκου Rayleigh και αυξάνει συνεχώς το ρυθμό σύγκλισης της αντίστροφης αναδρομής.

Επαναληπτική Μέθοδος πηλίκου Rayleigh

$$\left\{ \begin{array}{l} v^{(0)} = \text{κάποιο διάνυσμα με } \|v^{(0)}\| = 1 \\ \lambda^{(0)} = (v^{(0)})^T A v^{(0)} = \text{το πηλίκο Rayleigh που αντιστοιχεί στο } v^{(0)} \\ \text{Για } k = 1, 2, \dots \\ \quad \text{Λύνουμε το σύστημα: } (A - \lambda^{(k-1)}I)w = v^{(k-1)} \text{ ως προς το } w \\ \quad v^{(k)} = \frac{w}{\|w\|} \\ \quad \lambda^{(k)} = (v^{(k)})^T A v^{(k)} \end{array} \right.$$

Εδώ, στην πρώτη γραμμή της επανάληψης είναι σαν να εφαρμόζουμε τον πίνακα $(A - \lambda^{(k-1)}I)^{-1}$ στο $v^{(k-1)}$ καθώς το $(A - \lambda^{(k-1)}I)w = v^{(k-1)}$ είναι ισοδύναμο με το $w = (A - \lambda^{(k-1)}I)^{-1}v^{(k-1)}$ όταν ο $(A - \lambda^{(k-1)}I)$ είναι αντιστρέψιμος ($\lambda^{(k-1)}$ εκτίμηση ιδιοτιμής) και μπορεί να χρησιμοποιηθεί και όταν ο $(A - \lambda^{(k-1)}I)$ δεν είναι αντιστρέψιμος, δηλαδή όταν το $\lambda^{(k-1)}$ είναι ιδιοτιμή. Στη δεύτερη γραμμή κανονικοποιούμε το διάνυσμα w και στην τρίτη βρίσκουμε το πηλίκο του Rayleigh.

Σε αυτόν τον αλγόριθμο, η σύγκλιση είναι πολύ γρήγορη. Σε κάθε επανάληψη τριπλασιάζεται το πλήθος ψηφίων ακρίβειας όπως φαίνεται στο παρακάτω Θεώρημα.

ΘΕΩΡΗΜΑ 4.3. Η Επαναληπτική Μέθοδος πηλίκου Rayleigh συγκλίνει σε ένα ζεύγος ιδιοτιμής και ιδιοδιανύσματος για κάθε αρχικό διάνυσμα $v^{(0)}$ εκτός από ένα πλήθος τέτοιων διανυσμάτων τάξης 0. Στις περιπτώσεις που η μέθοδος συγκλίνει, η σύγκλιση αυτή είναι κυβική με την έννοια ότι εάν λ_J είναι μία ιδιοτιμή του A και το $v^{(0)}$ είναι αρκετά κοντά σε ένα ιδιοδιάνυσμα q_J , τότε:

$$(4.6) \quad \|v^{(k+1)} - (\pm q_J)\| = O(\|v^{(k)} - (\pm q_J)\|^3)$$

και

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

καθώς $k \rightarrow \infty$.

Το σύμβολο \pm δεν είναι αναγκαστικά το ίδιο και στα δύο μέλη της (4.6).

ΑΠÓΔΕΙΞΗ. Δε θα κάνουμε την απόδειξη του ότι σχεδόν για όλα τα αρχικά διανύσματα η μέθοδος συγκλίνει αλλά θα δείξουμε ότι εάν η μέθοδος συγκλίνει, τότε συγκλίνει κυβικά. Για απλούστευση της απόδειξης, υποθέτουμε ότι η ιδιοτιμή λ_J είναι απλή. Από τη σχέση (4.3), εάν $\|v^{(k)} - q_J\| \leq \epsilon$ για αρκετά μικρό ϵ , τότε το πηλίκου του Rayleigh δίνει μια εκτίμηση της ιδιοτιμής λ_k τέτοια ώστε: $|\lambda_k - \lambda_J| = O(\epsilon^2)$. Χρησιμοποιώντας παρόμοια επιχειρήματα με αυτά που χρησιμοποιούνται για τις αποδείξεις των προηγούμενων, έχουμε ότι εάν πάρουμε ένα βήμα της μεθόδου της οπισθοδρομικής αναδρομής για να βρούμε μία νέα εκτίμηση ιδιοδιανύσματος $v^{(k+1)}$ από τα $v^{(k)}$ και $\lambda^{(k)}$ τότε:

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

Ακόμα, ο σταθερός όρος που βρίσκεται στο O είναι σταθερός για μικρές γειτονίες του λ_J και του q_J . Επομένως, αν αρχικά το $\|v^{(k)} \pm q_J\|$ είναι της τάξης του ϵ , τότε το $|\lambda_k - \lambda_J|$ θα είναι της τάξεως του ϵ^2 . Άρα στο επόμενο βήμα θα είναι το $\|v^{(k+1)} \pm q_J\|$ της τάξης του ϵ^3 , και το $|\lambda_{k+1} - \lambda_J|$ θα είναι της τάξεως του ϵ^6 κ.ο.κ. Οι εκτιμήσεις (4.6)-(4.7) προκύπτουν από αυτό. \square

ΠΑΡΑΔΕΙΓΜΑ 4.4. Δίνουμε ένα αριθμητικό παράδειγμα για να δούμε το πόσο γρήγορη είναι η κυβική σύγκλιση της μεθόδου.

Έστω ο συμμετρικός πίνακας:

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

και έστω $v^{(0)} = (1, 1, 1)^T / \sqrt{3}$ το αρχικό εκτιμώμενο ιδιοδιάνυσμα. Όταν εφαρμόζουμε τη μέθοδο αναδρομής πηλίκου του Rayleigh στον A , στις τρεις πρώτες επαναλήψεις, υπολογίζουμε τις παρακάτω τιμές για την ιδιοτιμή $\lambda^{(k)}$:

$$\lambda^{(0)} = 5$$

$$\lambda^{(1)} = 5.2131\dots$$

$$\lambda^{(2)} = 5.214319743184.$$

Η πραγματική τιμή της ιδιοτιμής που αντιστοιχεί στο πιο κοντινό ιδιοδιάνυσμα στο αρχικό $v^{(0)}$ είναι η $\lambda = 5.214319743377$. Επομένως η αναδρομική μέθοδος πηλίκου του Rayleigh μετά από μόλις τρεις επαναλήψεις έχει πετύχει αποτελέσματα με ακρίβεια δεκα ψηφίων.

4.6. Μέτρηση πράξεων

Θα ολοκληρώσουμε αυτό το κεφάλαιο αναφέροντας κάποια πράγματα για το πόσο ακριβές είναι οι μέθοδοι που αναφέρθηκαν νωρίτερα όσον αφορά το πλήθος πράξεων που απαιτούν.

Αρχικά υποθέτουμε ότι ο $A \in \mathbb{R}^{m \times m}$ είναι ένας πυκνός πίνακας. Τότε κάθε βήμα της μεθόδου των δυνάμεων περιέχει έναν πολλαπλασιασμό πίνακα-διανύσματος ο οποίος απαιτεί $O(m^2)$ πράξεις. Κάθε βήμα της μεθόδου αντίστροφης αναδρομής περιλαμβάνει τη λύση ενός γραμμικού συστήματος η οποία χρειάζεται $O(m^3)$ πράξεις εκτός εάν έχει υπολογιστεί η παραγοντοποίηση LU ή QR του πίνακα A ή έχει χρησιμοποιηθεί κάποια άλλη μέθοδος, οπότε οι πράξεις μειώνονται σε $O(m^2)$. Στην περίπτωση της επαναληπτικής μεθόδου ηλίικου του Rayleigh σε κάθε βήμα, ο πίνακας που πρέπει να αντιστραφεί δηλαδή για τον οποίον θα πρέπει να λύσουμε το γραμμικό σύστημα αλλάζει διότι αλλάζει η εκτίμηση της ιδιοτιμής. Επομένως στην τρίτη μέθοδο δεν είναι εύκολο να μειώσουμε τις $O(m^3)$ πράξεις.

Γενικά όμως όταν ο πίνακας είναι μιας ειδική μορφής, το πλήθος των πράξεων μειώνεται. Στην ειδική περίπτωση τριδιαγώνιου πίνακα και οι τρεις μέθοδοι απαιτούν $O(m)$ πράξεις σε κάθε βήμα. Αν ο πίνακας είναι πίνακας Hessenberg αντί για τριδιαγώνι-ος, τότε οι απαιτούμενες πράξεις σε κάθε βήμα είναι $O(m^2)$.

Αλγόριθμος QR χωρίς εναλλαγές

Ο αλγόριθμος QR είναι από τους πλέον σημαντικούς αλγορίθμους της αριθμητικής ανάλυσης και βασίζεται στην παραγοντοποίηση QR ενός πίνακα. Η παραγοντοποίηση QR ενός πίνακα A είναι η ανάλυση του σε δυο πίνακες Q και R , $A = QR$, όπου ο πρώτος είναι ορθοκανονικός πίνακας και ο δεύτερος ένας άνω τριγωνικός πίνακας αν ο A είναι τετραγωνικός ή πλησιάζει τη μορφή κανονικού πίνακα αν ο A είναι ορθογώνιος. Εδώ δείχνουμε πώς στην πιο απλή του μορφή ο αλγόριθμος αυτός είναι μια σταθερή διαδικασία για τον υπολογισμό παραγοντοποιήσεων QR δυνάμεων ενός πίνακα

$$A, A^2, A^3, \dots$$

5.1. Ο απλός αλγόριθμος QR

Η πιο βασική μορφή του αλγορίθμου QR είναι εξαιρετικά απλή.

Ο απλός αλγόριθμος QR

$$\begin{cases} A^{(0)} = A \\ \text{Για } k = 1, 2, \dots \\ Q^{(k)} R^{(k)} = A^{(k-1)} \\ A^{(k)} = R^{(k)} Q^{(k)} \end{cases}$$

Στην πρώτη γραμμή της επανάληψης, βρίσκουμε την QR παραγοντοποίηση του πίνακα $A^{(k-1)}$ και στη δεύτερη αναδιατάσσουμε τους παράγοντες $Q^{(k)}$ και $R^{(k)}$ για να πάρουμε τον επόμενο πίνακα $A^{(k)}$. Το μόνο που κάνουμε είναι να επαναλαμβάνουμε το εξής: Βρίσκουμε την παραγοντοποίηση QR ενός πίνακα και πολλαπλασιάζουμε τους όρους της παραγοντοποίησης Q και R που έχουμε υπολογίσει με την αντίστροφη διάταξη. Επίσης, με κάποιες επιπλέον υποθέσεις, ο αλγόριθμος αυτός συγκλίνει σε έναν πίνακα A που είναι της μορφής του Schur (άνω τριγωνικός αν ο A είναι τυχαίος και διαγώνιος αν ο A είναι ερμητιανός). Εδώ χάριν απλότητας, θα εξακολουθήσουμε να θεωρούμε τον A πραγματικό και συμμετρικό πίνακα με πραγματικές ιδιοτιμές λ_j και ορθοκανονικά ιδιοδιανύσματα q_j . Επομένως το ενδιαφέρον μας περιορίζεται στη σύγκλιση της ακολουθίας πινάκων $A^{(k)}$ σε διαγώνια μορφή.

Για να είναι χρήσιμος ο αλγόριθμος αυτός δηλαδή η σύγκλιση σε διαγώνια μορφή να μας δίνει τις ιδιοτιμές του αρχικού πίνακα, χρειάζεται ο νέος πίνακας (ο διαγώνιος) που προκύπτει από τη σύγκλιση του αλγορίθμου να είναι όμοιος ως προς τον αρχικό. Δηλαδή οι υπολογισμοί που κάνουμε να είναι αποτελέσματα μετασχηματισμών ομοιότητας. Αυτό μπορεί εύκολα να αποδειχθεί.

Αρχικά αναλύουμε τον αρχικό πίνακα του k βήματος, τον $A^{(k-1)}$ ως εξής:

$A^{(k-1)} = Q^{(k)} R^{(k)}$, οπότε, καθώς ο $Q^{(k)}$ είναι εξ' ορισμού ορθομοναδιαίος, μπορούμε να πολλαπλασιάσουμε από αριστερά με τον αντίστροφο $(Q^{(k)})^T$ του $Q^{(k)}$ και να βρούμε τον $R^{(k)}$. Έχουμε τότε:

$$(Q^{(k)})^T A^{(k-1)} = R^{(k)}.$$

Πολλαπλασιάζοντας τώρα από δεξιά με τον $Q^{(k)}$ και τα δύο μέλη, παίρνουμε το εξής: $(Q^{(k)})^T A^{(k-1)} Q^{(k)} = R^{(k)} Q^{(k)}$. Όμως σύμφωνα με τον αλγόριθμο, $A^{(k)} = R^{(k)} Q^{(k)}$. Άρα τελικά έχουμε: $(Q^{(k)})^T A^{(k-1)} Q^{(k)} = A^{(k)}$ που είναι πραγματικά ένας μετασχηματισμός ομοιότητας.

Όπως και η επαναληπτική μέθοδος ηλιόκου Rayleigh, έτσι και η αλγοριθμική μέθοδος QR συγκλίνει κυβικά. Για να επιτευχθεί αυτό όμως, ο αλγόριθμος πρέπει να εφοδιαστεί με εναλλαγές σε κάθε βήμα. Αυτή είναι μία από τις τρεις αλλαγές που κάνουμε στον απλό αλγόριθμο QR για να τον μετατρέψουμε σε έναν πρακτικό αλγόριθμο που να συγκλίνει με γρήγορους ρυθμούς.

Οι τρεις αλλαγές είναι:

1. Προτού ξεκινήσει η επανάληψη, ο A τριδιαγωνιοποιείται όπως αναφέρθηκε στο προηγούμενο κεφάλαιο.

2. Αντί για τον $A^{(k)}$, ένας άλλος πίνακας ο $A^{(k)} - \mu^{(k)} I$ παραγοντοποιείται σε κάθε βήμα, όπου το $\mu^{(k)}$ είναι μία προσέγγιση ιδιοτιμής.

3. Όποτε είναι δυνατόν και ειδικότερα όταν έχει βρεθεί μια ιδιοτιμή, το πρόβλημα χωρίζεται σε άλλα μικρότερα σπάζοντας τον πίνακα A σε υποπίνακες.

Ένας αλγόριθμος QR που εμπεριέχει τα παραπάνω στοιχεία έχει την παρακάτω γενική μορφή:

Πρακτικός Αλγόριθμος QR

$$\left\{ \begin{array}{l} (Q^{(0)})^T A^{(0)} Q^{(0)} = A \\ \text{Για } k = 1, 2, \dots \\ \bullet \text{ επιλέγουμε ένα } \mu^{(k)} \text{ και παραγοντοποιούμε το:} \\ Q^{(k)} R^{(k)} = A^{(k-1)} - \mu^{(k)} I \\ \bullet A^{(k)} = R^{(k)} Q^{(k)} + \mu^{(k)} I \end{array} \right.$$

Το πρώτο πράγμα, όπως προαναφέρθηκε, είναι να φέρουμε τον πίνακα A σε τριδιαγώνια μορφή. Στην πρώτη σειρά της επανάληψης, επιλέγουμε κάποια εκτίμηση για την ιδιοτιμή. Μπορούμε για παράδειγμα να πάρουμε το κάτω δεξιά στοιχείο του πίνακα: $A_{m,m}^{(k-1)}$.

Στην επόμενη σειρά γίνεται η παραγοντοποίηση QR του πίνακα $A^{(k-1)} - \mu^{(k)} I$. Η τρίτη σειρά είναι η αναδιάταξη των παραγόντων της ανάλυσης QR. Σε αυτό το σημείο, εάν κάποιο στοιχείο της υπερδιαγωνίου $A_{j,j+1}^{(k)}$ είναι κοντά στο 0, θέτουμε $A_{j,j+1} = A_{j+1,j} = 0$ ώστε να διαχωρίσουμε τον $A^{(k)}$ ως

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

Για τα τμήματα του αλγορίθμου QR είτε έχει ήδη γίνει αναφορά στο προηγούμενο κεφάλαιο (τριδιαγωνιοποίηση) είτε θα γίνει αναφορά στο επόμενο κεφάλαιο (εναλλαγές) είτε δε θα γίνει καθόλου αναφορά (διαχωρισμός πίνακα σε υποπίνακες και διαχωρισμός

προβλήματος σε υποπροβλήματα). Προς το παρόν θα ασχοληθούμε με το πώς ο απλός αλγόριθμος QR βρίσκει τις ιδιοτιμές του πίνακα.

5.2. Μη κανονικοποιημένες παράλληλες Επαναλήψεις

Για να δείξουμε το πώς ο αλγόριθμος QR βρίσκει τις ιδιοτιμές του πίνακα, θα προσπαθήσουμε να τον συσχετίσουμε με μια άλλη μέθοδο, τη μέθοδο των παράλληλων επαναλήψεων, της οποίας η συμπεριφορά είναι πιο προφανής. Η ιδέα της μεθόδου των παράλληλων επαναλήψεων είναι να εφαρμόσει τη μέθοδο των δυνάμεων ταυτοχρόνως σε περισσότερα από ένα διανύσματα σε κάθε βήμα. (Ένας ισοδύναμος όρος είναι μέθοδος των δυνάμεων σε block). Έστω ότι ξεκινάμε με ένα σύνολο n γραμμικώς ανεξάρτητων διανυσμάτων $v_1^{(0)}, \dots, v_n^{(0)}$. Φαίνεται εύλογο ότι, όπως το $A^k v_1^{(0)}$ συγκλίνει στο ιδιοδιάνυσμα με τη μεγαλύτερη σε απόλυτη τιμή ιδιοτιμή του A καθώς $k \rightarrow \infty$ (υπό κάποιες κατάλληλες προϋποθέσεις), έτσι και ο χώρος $\langle A^k v_1^{(0)}, \dots, A^k v_n^{(0)} \rangle$ που παράγεται από τα διανύσματα $A^k v_1^{(0)}, \dots, A^k v_n^{(0)}$ θα συγκλίνει (πάλι υπό τις ίδιες προϋποθέσεις) στο χώρο $\langle q_1, \dots, q_n \rangle$ που παράγεται από τα ιδιοδιανύσματα q_1, \dots, q_n του A που αντιστοιχούν στις n μεγαλύτερες κατ' απόλυτη τιμή ιδιοτιμές του πίνακα A .

Σε συμβολισμό πινάκων, μπορούμε να το προσεγγίσουμε ως εξής. Ορίζουμε $V^{(0)}$ να είναι ο αρχικός $m \times n$ πίνακας

$$(5.1) \quad V^{(0)} = \left[v_1^{(0)} \mid \dots \mid v_n^{(0)} \right]$$

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

$$(5.2) \quad V^{(k)} = A^k V^{(0)} = \left[v_1^{(k)} \mid \dots \mid v_n^{(k)} \right].$$

Επειδή το ενδιαφέρον μας είναι στο χώρο στηλών του $V^{(k)}$, θα υπολογίσουμε μια βάση αυτού του χώρου χρησιμοποιώντας την ανάλυση QR του $V^{(k)}$:

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

Εδώ οι $Q^{(k)}$ και $R^{(k)}$ έχουν διαστάσεις $m \times n$ και $n \times n$ αντίστοιχα. Πάλι εδώ μπορεί να αποδειχθεί ότι καθώς $k \rightarrow \infty$ οι διαδοχικές στήλες του $Q^{(k)}$ συγκλίνουν, υπό ορισμένες προϋποθέσεις, στα ιδιοδιανύσματα $\pm q_1, \pm q_2, \dots, \pm q_n$. Αυτό μπορεί να γίνει όπως και στο προηγούμενο κεφάλαιο. Γράφουμε τα $v_j^{(0)}$ και $v_j^{(k)}$ ως γραμμικούς συνδυασμούς των ιδιοδιανυσμάτων του A . Έχουμε:

$$\begin{aligned} v_j^{(0)} &= a_{1j} q_1 + \dots + a_{mj} q_m \\ v_j^{(k)} &= \lambda_1^k a_{1j} q_1 + \dots + \lambda_m^k a_{mj} q_m \end{aligned}$$

Και εδώ έχουμε απλή σύγκλιση αν ισχύουν δύο προϋποθέσεις.

Η πρώτη είναι ότι οι $n+1$ μεγαλύτερες ιδιοτιμές κατά απόλυτη τιμή είναι έχουν διακριτές απόλυτες τιμές:

$$(5.4) \quad |\lambda_1| > |\lambda_2| > \dots > |\lambda_n| > |\lambda_{n+1}| \geq |\lambda_{n+2}| \geq \dots \geq |\lambda_m|.$$

Η δεύτερη είναι ότι ο πίνακας που παράγουν τα a_{ij} είναι αντιστρέψιμος. Ορίζουμε ως \hat{Q} τον $m \times n$ πίνακα του οποίου οι στήλες είναι τα ιδιοδιανύσματα q_1, q_2, \dots, q_n .

(Δηλαδή ο \hat{Q} πίνακας των ιδιοδιανυσμάτων είναι διαφορετικός από τους πίνακες $\hat{Q}^{(k)}$ που προκύπτουν από την παραγοντοποίηση QR.) Υποθέτουμε και ότι:

(5.5) Όλοι οι πρώτοι υποπίνακες του $\hat{Q}^T V^{(0)}$ είναι αντιστρέψιμοι

Λέγοντας πρώτοι 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 το να πούμε ότι ο $\hat{Q}^T V^{(0)}$ επιδέχεται παραγοντοποίηση LU.)

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

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

για κάθε j με $1 \leq j \leq n$, όπου το $C < 1$ είναι η σταθερά $\max_{1 \leq k \leq n} \frac{|\lambda_{k+1}|}{|\lambda_k|}$. Όπως και στα προηγούμενα θεωρήματα, έτσι και σε αυτό το πρόσημο \pm σημαίνει ότι σε κάθε βήμα k μπορούμε να πάρουμε ένα από τα δύο πρόσημα και τα φράγματα να ισχύουν.

ΑΠÓΔΕΙΞΗ. Επεκτείνουμε τον πίνακα των ιδιοδιανυσμάτων \hat{Q} σε έναν γεμάτο $m \times m$ ορθογώνιο πίνακα Q ιδιοδιανυσμάτων του A και έστω Λ ο αντίστοιχος διαγώνιος πίνακας των ιδιοτιμών. Τότε: $A = Q\Lambda Q^T$. Όπως ο \hat{Q} είναι το πάνω αριστερά $m \times n$ τμήμα του Q , του πίνακα των ιδιοδιανυσμάτων του A , ορίζουμε ως $\hat{\Lambda}$ το αρχικό (πάνω αριστερά) τμήμα του Λ που είναι επίσης διαγώνιος. 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 $k \rightarrow \infty$. Εάν η (5.5) ισχύει, τότε ειδικότερα ο $\hat{Q}^T V^{(0)}$ είναι αντιστρέψιμος οπότε μπορούμε να πολλαπλασιάσουμε τον όρο $O(|\lambda_{n+1}|^k)$ από δεξιά με τον $(\hat{Q}^T V^{(0)})^{-1} \hat{Q}^T V^{(0)}$ και να μετατρέψουμε την εξίσωση σε

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

Καθώς ο $\hat{Q}^T V^{(0)}$ είναι αντιστρέψιμος, ο χώρος στηλών αυτού του πίνακα είναι ο ίδιος με το 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 φανερό ότι ο χώρος στηλών του συγκλίνει στο χώρο στηλών του πίνακα \hat{Q} .

Στην πραγματικότητα, εδώ υποθέσαμε, όχι μόνο ότι ο πίνακας $\hat{Q}^T V^{(0)}$ είναι αντιστρέψιμος, αλλά και ότι όλοι οι κύριοι υποπίνακες του είναι αντιστρέψιμοι. Το ίδιο ισχύει για τα κύρια υποσύνολα στηλών των $V^{(k)}$ και \hat{Q} : οι πρώτες στήλες, οι πρώτες και οι δεύτερες κοκ. Σε κάθε περίπτωση, καταλήγουμε στο συμπέρασμα 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 QR παραγοντοποίησης (5.3), προκύπτει το ζητούμενο. \square

5.3. Ταυτόχρονες επαναλήψεις

Καθώς το $k \rightarrow \infty$ τα διανύσματα $v_1^{(k)}, \dots, v_n^{(k)}$ του αλγορίθμου (5.1) -(5.3) συγκλίνουν όλα σε πολλαπλάσια του ιδιοδιανύσματος q_1 (το ιδιοδιάνυσμα που αντιστοιχεί στη μεγαλύτερη κατ' απόλυτη τιμή ιδιοτιμή του A). Έτσι, αν και ο χώρος που παράγουν κάποια από αυτά

$\langle v_1^{(k)}, \dots, v_j^{(k)} \rangle$ συγκλίνει σε κάτι χρήσιμο, τα διανύσματα αυτά αποτελούν μια προβληματική βάση του χώρου. Εάν εκτελούσαμε ταυτόχρονες επαναλήψεις, όπως περιγράφηκαν παραπάνω, με αριθμητική απλής ακρίβειας, τότε η πληροφορία που θέλουμε να εξάγουμε (δηλαδή ο πίνακας που θα μας δώσει τις ιδιοτιμές) γρήγορα θα αλλοιωθεί από λάθη που θα γίνονται κατά τις επαναλήψεις.

Αυτό που μπορεί να γίνει είναι να ορθοκανονικοποιούμε σε κάθε βήμα της επανάληψης και όχι μία φορά στην αρχή για όλα τα βήματα της επανάληψης. Έτσι, αντί να κατασκευάσουμε τους $V^{(k)}$ όπως παραπάνω, θα κατασκευάσουμε τους $Z^{(k)}$ οι οποίοι θα έχουν τον ίδιο χώρο στηλών (άρα θα συγκλίνουν στα ιδιοδιανύσματα του A που αντιστοιχούν στις μεγαλύτερες κατ' απόλυτη τιμή ιδιοτιμές).

Μέθοδος Ταυτόχρονων Επαναλήψεων

$$\left\{ \begin{array}{l} \text{Επιλέγουμε } \hat{Q}^{(0)} \in \mathbb{R}^{m \times n} \text{ με ορθοκανονικές στήλες} \\ \text{Για } k = 1, 2, \dots \\ Z = A\hat{Q}^{(k-1)} \\ \hat{Q}^{(k)}\hat{R}^{(k)} = Z \end{array} \right.$$

Από τη μορφή αυτού του αλγορίθμου, είναι φανερό ότι οι χώροι στηλών των $\hat{Q}^{(k)}$ και $\hat{R}^{(k)}$ είναι οι ίδιοι καθώς είναι και οι δύο ίσοι με το χώρο στηλών του $A^k\hat{Q}^{(0)}$. Έτσι, ο νέος αυτός αλγόριθμος συγκλίνει με τις ίδιες συνθήκες όπως και ο προηγούμενος.

ΘΕΩΡΗΜΑ 5.2. *Ο αλγόριθμος των ταυτόχρονων επαναλήψεων, που μόλις αναφέρθηκε, δημιουργεί τους ίδιους πίνακες $\hat{Q}^{(k)}$ όπως η επαναληπτική μέθοδος που περιγράφουν οι εξισώσεις (5.1)-(5.3) όπως στο Θεώρημα 5.1 με την προϋπόθεση ότι οι αρχικοί πίνακες $\hat{Q}^{(0)}$ είναι οι ίδιοι και κάτω από τις ίδιες υποθέσεις (5.4) και (5.5) συγκλίνει όπως σε εκείνο το Θεώρημα.*

5.4. Ισοδυναμία Ταυτόχρονων Επαναλήψεων και Αλγορίθμου QR

Τώρα μπορούμε να εξηγήσουμε τον αλγόριθμο QR. Αυτός είναι ισοδύναμος με τη μέθοδο των παράλληλων επαναλήψεων για ένα σύνολο $n = m$ αρχικών διανυσμάτων (που το πλήθος τους είναι ίσο με τη διάστασή τους) που σχηματίζουν ένα ταυτοτικό πίνακα $\hat{Q}^{(0)} = I$. Καθώς οι πίνακες $\hat{Q}^{(k)}$ είναι τώρα τετραγωνικοί, οι QR παραγοντοποιήσεις είναι πλήρεις, οπότε μπορούμε να επανέλθουμε στη γραφή $Q^{(k)}$ και $R^{(k)}$ αντί για $\hat{Q}^{(k)}$ και $\hat{R}^{(k)}$. Τώρα όμως θα χρησιμοποιήσουμε ένα νέο συμβολισμό για τους πίνακες Q του αλγορίθμου QR, τον $\check{Q}^{(k)}$ για να τους ξεχωρίζουμε από αυτούς που προκίπτουν από τη μέθοδο των παράλληλων επαναλήψεων.

Εδώ υπάρχουν τρεις εξισώσεις που περιγράφουν τη μέθοδο των παράλληλων επαναλήψεων με $\check{Q}^{(0)} = I$ οι οποίες ακολουθούνται από μια τρίτη εξίσωση που θα χρησιμοποιήσουμε ως ορισμό του $m \times m$ πίνακα $A^{(k)}$.

Παράλληλες επαναλήψεις:

$$(5.7) \quad \check{Q}^{(0)} = I$$

$$(5.8) \quad Z = A\check{Q}^{(k-1)}$$

$$(5.9) \quad Z = \check{Q}^{(k)}R^{(k)}$$

$$(5.10) \quad A^{(k)} = (\check{Q}^{(k)})^T A Q^{(k)}$$

Και εδώ υπάρχουν άλλες τρεις εξισώσεις οι οποίες καθορίζουν τον απλό αλγόριθμο QR και οι οποίες επίσης ακολουθούνται από τον ορισμό του $m \times m$ πίνακα $\check{Q}^{(k)}$.

Απλός αλγόριθμος QR χωρίς επαναλήψεις:

$$(5.11) \quad A^{(0)} = A$$

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

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

$$(5.14) \quad \check{Q}^{(k)} = Q^{(1)} Q^{(2)} \dots Q^{(k)}$$

Επιπλέον, και για τους δύο αλγόριθμους, ορίζουμε επιπλέον ένα πίνακα $m \times m$ τον $\check{R}^{(k)}$ ως εξής:

$$(5.15) \quad \check{R}^{(k)} = R^{(k)} R^{(k-1)} \dots R^{(1)}$$

Μπορούμε τώρα να δούμε την ισοδυναμία των δύο αλγορίθμων:

ΘΕΩΡΗΜΑ 5.3. *Οι δύο μέθοδοι που περιγράφονται από τις εξισώσεις (5.7)-(5.10) και (5.11)-(5.14) δημιουργούν την ίδια ακοιλουθία πινάκων $\check{R}^{(k)}$, $\check{Q}^{(k)}$ και $A^{(k)}$ που ορίζονται από την παραγοντοποίηση QR της κοστής δύναμης του πίνακα A ,*

$$(5.16) \quad A^k = \check{Q}^{(k)} \check{R}^{(k)}$$

και από την

$$(5.17) \quad A^{(k)} = (\check{Q}^{(k)})^T A \check{Q}^{(k)}$$

ΑΠÓΔΕΙΞΗ. Δουλεύουμε με επαγωγή ως προς το k .

Για $k = 0$, και για τις δύο μεθόδους ισχύει ότι $A^0 = \check{Q}^{(0)} = \check{R}^{(0)} = I$ και $A^{(0)} = A$ απ'όπου οι εξισώσεις του θεωρήματος είναι προφανείς. Έστω ότι οι ισότητες ισχύουν στο τυχαίο βήμα $k - 1$ για $k \geq 1$. Θα δείξουμε ότι ισχύουν και στην περίπτωση που $k \geq 1$. Στην περίπτωση της μεθόδου των παράλληλων επαναλήψεων η εξίσωση (5.17) είναι η ίδια με την εξίσωση (5.10), οπότε αρκεί να δείξουμε ότι ισχύει και η εξίσωση (5.16). Έχουμε $A^k = A A^{k-1} = A \check{Q}^{(k-1)} \check{R}^{(k-1)} = Z \check{R}^{(k-1)} = \check{Q}^{(k)} R^{(k)} \check{R}^{(k-1)} = \check{Q}^{(k)} \check{R}^{(k)}$.

Η πρώτη ισότητα είναι ο αναδρομικός ορισμός της δύναμης πίνακα, η δεύτερη προκύπτει από την επαγωγική υπόθεση, η τρίτη είναι αποτέλεσμα της (5.8), η επόμενη προκύπτει από την (5.9) και η τελευταία από την (5.15)

Στην περίπτωση της μεθόδου QR αποδεικνύουμε την εξίσωση (5.16) ως εξής:

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

Η πρώτη ισότητα είναι ο αναδρομικός ορισμός της δύναμης πίνακα, η δεύτερη προκύπτει από την επαγωγική υπόθεση (5.16), η τρίτη είναι αποτέλεσμα της επαγωγικής υπόθεσης (5.17), η επόμενη προκύπτει από την (5.12) και η τελευταία από τις (5.14) και (5.15).

Η εξίσωση (5.17) μπορεί να αποδειχθεί ως εξής:

$$A^{(k)} = R^{(k)} Q^{(k)} = (Q^{(k)})^T A^{(k-1)} Q^{(k)} = (Q^{(k)})^T (Q^{(k-1)})^T A Q^{(k-1)} Q^{(k)} = (\check{Q}^{(k)})^T A \check{Q}^{(k-1)}.$$

Η πρώτη ισότητα είναι η εξίσωση (5.13), η δεύτερη προκύπτει από την (5.12), η τρίτη είναι αποτέλεσμα της επαγωγικής υπόθεσης για την (5.17) και η τελευταία έπεται από την εξίσωση (5.14) \square

5.5. Σύγκλιση του αλγορίθμου QR

Τώρα μπορούμε να αναφερθούμε στη σύγκλιση του αλγορίθμου QR. Η εξίσωση (5.16) μας δείχνει ότι η μέθοδος QR μπορεί να βρει τα ιδιοδιανύσματα του πίνακα A καθώς παράγει ορθοκανονικές βάσεις διαδοχικών δυνάμεων A^k του πίνακα A . Η εξίσωση (5.17) μας δείχνει πως ο αλγόριθμος βρίσκει τις ιδιοτιμές διότι από αυτήν προκύπτει ότι τα διαγώνια στοιχεία του πίνακα $A^{(k)}$ είναι τα πηλίκα του Rayleigh του A που αντιστοιχούν στις στήλες του $\tilde{Q}^{(k)}$ και καθώς οι στήλες αυτές συγκλίνουν στα ιδιοδιανύσματα, τα πηλίκα του Rayleigh συγκλίνουν με τετραγωνική ταχύτητα (βλ.(4.3)) στις ιδιοτιμές. Επιπλέον, η σχέση (5.17) δείχνει ότι τα μη-διαγώνια στοιχεία του $A^{(k)}$ αντιστοιχούν σε γενικευμένα πηλίκα του Rayleigh που περιλαμβάνουν προσεγγίσεις ιδιοδιανυσμάτων του A διακριτών μεταξύ τους. Αυτές οι προσεγγίσεις πρέπει να είναι ορθογώνιες και άρα τα μη-διαγώνια στοιχεία του $A^{(k)}$ πρέπει να συγκλίνουν στο 0.

Από το Θεώρημα 5.2 προκύπτει το παρακάτω θεώρημα.

ΘΕΩΡΗΜΑ 5.4. *Έστω ότι ο απλός αλγόριθμος QR εφαρμόζεται σε ένα πραγματικό συμμετρικό πίνακα A για του οποίου τις ιδιοτιμές ισχύει ότι: $|\lambda_1| > |\lambda_2| > \dots > |\lambda_m|$ και του οποίου ο πίνακας ιδιοδιανυσμάτων Q έχει όλους τους κύριους υποπίνακες αντιστρέψιμους. Τότε καθώς το $k \rightarrow \infty$ ο $A^{(k)}$ συγκλίνει γραμμικά στον $\text{diag}(\lambda_1, \dots, \lambda_m)$ με σταθερά $\max_j |\lambda_{j+1}|/|\lambda_j|$ και ο πίνακας $\tilde{Q}^{(k)}$ (με πρόσχημα όπως χρειάζεται κάθε φορά) συγκλίνει στον Q με τον ίδιο ρυθμό.*

Αλγόριθμος QR με εναλλαγές

Αυτό που κάνει την επαναληπτική μέθοδο QR να είναι πάρα πολύ χρήσιμη και δυνατή είναι η εισαγωγή εναλλαγών του πίνακα A που τον μετατρέπουν σε ένα πίνακα της μορφής $A - \mu I$ σε κάθε βήμα της επανάληψης. Εδώ θα εξηγήσουμε πώς αυτή η ιδέα οδηγεί σε κυβική σύγκλιση της μεθόδου με τη βοήθεια της επαναληπτικής μεθόδου πηλίκου Rayleigh με την οποία θα τη συνδέσουμε.

6.1. Σύνδεση με αντίστροφη αναδρομή

Εξακολουθούμε να υποθέτουμε ότι ο $A \in \mathbb{R}^{m \times m}$ είναι πραγματικός και συμμετρικός με πραγματικές ιδιοτιμές λ_j και ορθοκανονικά διανύσματα q_j .

Όπως έχουμε ήδη δει, ο απλός αλγόριθμος QR είναι ισοδύναμος με τη μέθοδο των παράλληλων επαναλήψεων εφαρμοζόμενη στον ταυτοτικό πίνακα. Ειδικότερα η πρώτη στήλη του αποτελέσματος συγκλίνει από τη μέθοδο των δυνάμεων εφαρμοζόμενη στο e_1 . Μπορούμε να δούμε αυτό και με ένα δεύτερο τρόπο. Ο απλός αλγόριθμος QR είναι επίσης ισοδύναμος με τις ταυτόχρονες αντίστροφες επαναλήψεις εφαρμοζόμενες σε έναν αλλαγμένο πίνακα P και εδώ η m -οστή στήλη του αποτελέσματος προκύπτει σύμφωνα με την αντίστροφη αναδρομή εφαρμοζόμενη στο e_m .

Μπορούμε τώρα να το αποδείξουμε αυτό ως εξής: Έστω $Q^{(k)}$, να είναι, όπως και στο προηγούμενο κεφάλαιο ο ορθοκανονικός πίνακας που προκύπτει στο k βήμα από τον αλγόριθμο QR. Στο προηγούμενο κεφάλαιο δείξαμε ότι το γινόμενο αυτών των πινάκων,

$\check{Q}^{(k)} = \prod_{j=1}^k Q^{(j)} = [q_1^{(k)} | q_2^{(k)} | \dots | q_m^{(k)}]$ είναι ο ίδιος ορθογώνιος πίνακας ο οποίος εμφανίζεται στο k βήμα της μεθόδου των παράλληλων επαναλήψεων. Ένας άλλος τρόπος να το δούμε αυτό είναι να δούμε ότι ο $\check{Q}^{(k)}$ είναι ο ορθογώνιος παράγοντας της παραγοντοποίησης QR

$A^{(k)} = \check{Q}^{(k)} \check{R}^{(k)}$. Τώρα εξετάζουμε τι συμβαίνει όταν αντιστρέψουμε τον τύπο αυτό. Υπολογίζουμε τα

$$(6.1) \quad A^{-k} = (\check{R}^{(k)})^{-1} (\check{Q}^{(k)})^T = \check{Q}^{(k)} (\check{R}^{(k)})^{-T}$$

Η πρώτη ισότητα προκύπτει από το γεγονός ότι ο Q είναι ορθοκανονικός και η δεύτερη ισότητα προκύπτει από το γεγονός ότι ο A άρα και ο A^{-1} είναι συμμετρικός. Ορίζουμε ως P τον $m \times m$ πίνακα μετάθεσης που αντιστρέφει τη σειρά των γραμμών και των στηλών:

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

Καθώς $P^2 = I$, η (6.1) μετατρέπεται σε:

$$(6.2) \quad A^{-k}P = [\check{Q}^{(k)}P][P(\check{R}^{(k)})^{-T}P]$$

$$(6.3) \quad A^{-k}P = [\check{Q}^{(k)}P][P(\check{R}^{(k)})^{-T}P].$$

Ο πρώτος παράγοντας σε αυτό το γινόμενο, ο $\check{Q}^{(k)}P$ είναι ορθογώνιος καθώς

$$\check{Q}^{(k)}P(\check{Q}^{(k)}P)^T = \check{Q}^{(k)}PP^T(\check{Q}^{(k)})^T = \check{Q}^{(k)}(\check{Q}^{(k)})^T = I.$$

Ο δεύτερος, $P(\check{R}^{(k)})^{-T}P$ είναι άνω τριγωνικός διότι ο $(\check{R}^{(k)})^{-T}$ είναι άνω τριγωνικός και όταν εφαρμοστεί ο P από δεξιά γίνεται κάτω τριγωνικός ενώ όταν ξαναεφαρμοστεί ο P από αριστερά, ξαναγίνεται άνω τριγωνικός. Έτσι μπορούμε να δούμε την (6.3) ως μια παραγοντοποίηση QR του πίνακα $A^{-k}P$. Δηλαδή χρησιμοποιούμε τη μέθοδο των παράλληλων επαναλήψεων στον A^{-1} εφαρμοζόμενο στον πίνακα P , την οποία μέθοδο ονομάζουμε μέθοδο παράλληλων αντίστροφων επαναλήψεων του A . Συγκεκριμένα, η πρώτη στήλη του πίνακα $\check{Q}^{(k)}P$ που είναι η τελευταία στήλη του πίνακα $\check{Q}^{(k)}$ είναι το αποτέλεσμα της εφαρμογής k βημάτων από τη μέθοδο της αντίστροφης αναδρομής στο διάνυσμα e_m .

6.2. Σύνδεση με την εναλλασσόμενη αντίστροφη αναδρομή

Με βάση τα παραπάνω, ο αλγόριθμος QR είναι ταυτόχρονα και παράλληλες επαναλήψεις και παράλληλες αντίστροφες επαναλήψεις. Όπως είδαμε στο κεφάλαιο 4, το ότι η μέθοδος των αντίστροφων επαναλήψεων μπορεί να γίνει χρησιμοποιώντας τυχαίες εναλλαγές δημιουργεί μια μεγάλη διαφορά. Όσο καλύτερα μπορούμε να εκτιμήσουμε μία μεταβλητή $\mu \approx \lambda_j$ τόσο γρηγορότερα θα φτάνουμε σε καλύτερες εκτιμήσεις χρησιμοποιώντας τον πίνακα $A - \mu I$. Ο πρακτικός αλγόριθμος QR δείχνει το πως μπορούμε να εισάγουμε τις εναλλαγές στα βήματα της επανάληψης. Αυτό είναι ακριβώς η εισαγωγή εναλλαγών στις ανίστοιχες μεθόδους των παράλληλων επαναλήψεων και των παράλληλων αντίστροφων επαναλήψεων και επομένως το αποτέλεσμα μιας τέτοιας εισαγωγής εναλλαγών θα είναι το ίδιο.

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

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

Από αυτό προκύπτει ότι

$$(6.4) \quad A^{(k)} = (Q^{(k)})^T A^{(k-1)} Q^{(k)}$$

και επαγωγικά προκύπτει ότι: $A^{(k)} = (\check{Q}^{(k)})^T A \check{Q}^{(k)}$ που δεν είναι άλλη από τη σχέση (5.17). Όμως η (5.16) δεν ισχύει πλέον. Αντί για αυτήν έχουμε την παραγοντοποίηση:

$$(6.5) \quad (A - \mu^{(k)}I)(A - \mu^{(k-1)}I) \cdots (A - \mu^{(1)}I) = \check{Q}^{(k)}\check{R}^{(k)}$$

που είναι μια παραλλαγή με εναλλαγές της μεθόδου των παράλληλων επαναλήψεων. Δηλαδή το γινόμενο $\check{Q}^{(k)} = \prod_{j=1}^k Q^{(j)}$ είναι ένας ορθογώνιος πίνακας που προκύπτει με τη μέθοδο αυτή από τον πίνακα:

$$\prod_{j=k}^1 (A - \mu^{(j)}I).$$

Η πρώτη στήλη του πίνακα $\check{Q}^{(k)}$ είναι το αποτέλεσμα της εφαρμογής της μεθόδου της εναλλασσόμενης αντίστροφης αναδρομής στο διάνυσμα e_1 με τις εναλλαγές $\mu^{(j)}$ και

η τελευταία στήλη είναι το αποτέλεσμα της εφαρμογής k βημάτων της μεθόδου της εναλλασσόμενης αντίστροφης αναδρομής στο διάνυσμα e_m με τις ίδιες εναλλαγές. Εάν οι εναλλαγές αυτές είναι καλές εκτιμήσεις των ιδιοτιμών, τότε η τελευταία στήλη του πίνακα $\check{Q}^{(k)}$ συγκλίνει γρήγορα σε ένα ιδιοδιάνυσμα.

6.3. Σύνδεση με την επαναληπτική μέθοδο του πηλίκου Rayleigh

Η μέθοδος της εναλλαγμένης αντίστροφης αναδρομής είναι ένα πολύ ισχυρό εργαλείο που ανακαλύψαμε μέσω του αλγορίθμου QR. Για να ολοκληρώσουμε την ιδέα ότι αυτή η μέθοδος είναι όντως μια καλή μέθοδος εύρεσης ιδιοτιμών-ιδιοδιανυσμάτων, χρειάζεται να βρούμε ένα τρόπο επιλογής των $\mu^{(k)}$ τέτοιον ώστε η μέθοδος να συγκλίνει γρήγορα στην τελευταία στήλη του πίνακα $\check{Q}^{(k)}$.

Θα ξεκινήσουμε χρησιμοποιώντας το πηλίκο του Rayleigh στην τελευταία στήλη του πίνακα $\check{Q}^{(k)}$ που συγκλίνει σε ιδιοδιάνυσμα και επομένως το πηλίκο του Rayleigh σε ιδιοτιμή. Έχουμε

$$(6.6) \quad \mu^{(k)} = \frac{(q_m^{(k)})^T A q_m^{(k)}}{(q_m^{(k)})^T q_m^{(k)}} = (q_m^{(k)})^T A q_m^{(k)}$$

Εάν αυτή η επιλογή του $\mu^{(k)}$ είναι σταθερή καθ' όλη τη διάρκεια των επαναλήψεων, τότε οι εκτιμήσεις $\mu^{(k)}$ των ιδιοτιμών και $q_m^{(k)}$ των ιδιοδιανυσμάτων είναι οι ίδιες με εκείνες που υπολογίζονται και από την επαναληπτική μέθοδο του πηλίκου Rayleigh που ξεκινά με αρχικό διάνυσμα το e_m . Επομένως και η μέθοδος της εναλλαγμένης αντίστροφης αναδρομής συγκλίνει κυβικά με την έννοια ότι το διάνυσμα $q_m^{(k)}$ συγκλίνει κυβικά σε ένα ιδιοδιάνυσμα.

Στον αλγόριθμο QR, το πηλίκο του Rayleigh $r(q_m^{(k)})$ εμφανίζεται ως το στοιχείο της m γραμμής και m στήλης του πίνακα $A^{(k)}$ καθώς, χρησιμοποιώντας την (6.4), έχουμε

$$(6.7) \quad A_{m,m}^{(k)} = e_m^T A^{(k)} e_m = e_m^T \check{Q}^{(k)T} A \check{Q}^{(k)} e_m = (q_m^{(k)})^T A q_m^{(k)}$$

Επομένως, μπορούμε να πάρουμε στην (6.6) $\mu^{(k)} = A_{m,m}^{(k)}$. Αυτό είναι γνωστό ως η εναλλαγή του πηλίκου Rayleigh.

6.4. Εναλλαγές Wilkinson

Παρ'όλο που οι εναλλαγές του πηλίκου Rayleigh συγκλίνουν κυβικά στη περίπτωση που το είδαμε, η σύγκλιση δεν είναι εξασφαλισμένη για όλες τις αρχικές συνθήκες. Ένα παράδειγμα μη σύγκλισης δίδεται παρακάτω.

ΠΑΡΑΔΕΙΓΜΑ 6.1. Έστω ο πίνακας

$$(6.8) \quad A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}.$$

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 εναλλογή του πηλίκου Rayleigh, $\mu = A_{m,m}$ δεν έχει ούτε αυτή αποτέλεσμα καθώς το $A_{m,m} = 0$. Δηλαδή μπορεί και ο αλγόριθμος QR με τις εναλλογές του πηλίκου Rayleigh να αποτύχει.

Το πρόβλημα προκύπτει λόγω της συμμετρίας των ιδιοτιμών. Η μία ιδιοτιμή είναι +1 και η άλλη είναι -1. Επομένως όταν προσπαθούμε να βελτιώσουμε την εκτίμηση της ιδιοτιμής που έχουμε, $\mu = 0$, αυτή τείνει εξίσου και στις δύο ιδιοτιμές και έτσι η εκτίμηση παραμένει αμετάβλητη. Αυτό που χρειάζεται είναι μία εκτίμηση ιδιοτιμής που να σπάει τη συμμετρία. Για να βρούμε μια τέτοια ιδιοτιμή, κάνουμε τα εξής: Ορίζουμε ως B τον 2×2 κάτω δεξιά υποπίνακα του $A^{(k)}$:

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

Εναλλαγή Wilkinson ορίζουμε την ιδιοτιμή του B που είναι πιο κοντά στο a_m . Σε περίπτωση που κάποιες ιδιοτιμές είναι εξίσου κοντά, τότε επιλέγουμε μία τυχαία. Ένας αριθμητικά σταθερός τύπος για τις εναλλογές Wilkinson είναι ο:

$$(6.9) \quad \mu = a_m - \text{sign}(\delta) b_{m-1}^2 / (|\delta| + \sqrt{\delta^2 + b_{m-1}^2})$$

όπου $\delta = (a_{m-1} - a_m)/2$. Εάν $\delta = 0$, τότε το $\text{sign}(\delta)$ μπορεί να επιλεγεί τυχαία είτε ως +1 είτε ως -1.

Όπως οι εναλλογές πηλίκου Rayleigh, έτσι και η μέθοδος με τις εναλλογές Wilkinson επιτυγχάνει κυβική σύγκλιση στην περίπτωση που την παρουσιάσαμε. Επίσης για τη μέθοδο αποδεικνύεται ότι στη χειρότερη περίπτωση συγκλίνει με τετραγωνικό ρυθμό. Δηλαδή, με ακριβή αριθμητική, ο αλγόριθμος QR με εναλλογές Wilkinson συγκλίνει πάντα.

Στο παράδειγμα που είχαμε, με τις εναλλογές Wilkinson επιλέγουμε τυχαία είτε $\text{sign}(\delta) = 1$ είτε $\text{sign}(\delta) = -1$, η συμμετρία χαλάει και η μέθοδος συγκλίνει σε ένα βήμα.

6.5. Σταθερότητα και ακρίβεια

Εδώ ολοκληρώνουμε την ανάλυση των μηχανισμών του αλγορίθμου QR, χωρίς να αναφερθούμε σε πρακτικές λεπτομέρειες. Μένει να πούμε λίγα λόγια για τη σταθερότητα και την ακρίβεια.

Όπως μπορεί κανείς να περιμένει από τη χρήση ορθογώνιων πινάκων στον αλγόριθμο QR, η μέθοδος είναι οπισθοδρομικά σταθερή. Όπως και σε προηγούμενο κεφάλαιο, ο ευκολότερος τρόπος να βρούμε ένα τύπο για το αποτέλεσμα είναι να ορίσουμε ως \tilde{A} τον πίνακα που είναι διαγωνιοποίηση του A όπως υπολογίζεται με απλής ακρίβειας αριθμητική και ως \tilde{Q} τον ορθογώνιο πίνακα που σχετίζεται με το γινόμενο όλων των υπολογισμένων ανακλαστών Householder (ή περιστροφών Givens) όπως υπολογίζονται ακριβώς από τον αλγόριθμο. Το παρακάτω θεώρημα μπορεί να αποδειχθεί:

ΘΕΩΡΗΜΑ 6.2. Έστω ένας πραγματικός, συμμετρικός, τριδιαγώνιος πίνακας $A \in \mathbb{R}^{m \times m}$ που έχει διαγωνιοποιηθεί από τον πρακτικό αλγόριθμο QR σε υπολογιστή που ικανοποιεί τα

αξιώματα 3.2 και 3.3 και έστω $\tilde{\Lambda}$ και \tilde{Q} να είναι όπως ορίστηκαν παραπάνω. Τότε έχουμε:

$$(6.10) \quad \begin{aligned} \tilde{Q}\tilde{\Lambda}\tilde{Q}^* &= A + \delta A \\ \frac{\|\delta A\|}{\|A\|} &= O(\epsilon_{\text{μηχανής}}) \end{aligned}$$

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

Ο αλγόριθμος QR δίνει ακριβώς την ίδια λύση σε ένα ελαφρώς διαταραγμένο πρόβλημα. Συνδυάζοντας τα θεωρήματα 3.1 και 6.2, βλέπουμε ότι η απλοποίηση σε τριδιαγώνια μορφή που προκύπτει από τον αλγόριθμο QR είναι ένας οπισθοδρομικά σταθερός αλγόριθμος υπολογισμού ιδιοτιμών πινάκων. Για να δούμε πώς αυτό σχετίζεται με την ακρίβεια των υπολογισμένων ιδιοτιμών, συνδυάζουμε αυτό το αποτέλεσμα με το ότι για τις ιδιοτιμές λ_j ενός πραγματικού και συμμετρικού πίνακα A (ειδική μορφή κανονικού πίνακα) και για τις ιδιοτιμές $\tilde{\lambda}_j$ του διαταραγμένου πίνακα $A + \delta A$ ισχύει:

$$(6.11) \quad |\tilde{\lambda}_j - \lambda_j| \leq \|\delta A\|_2.$$

Το αποτέλεσμα είναι ότι οι ιδιοτιμές που υπολογίζονται με αυτόν τον τρόπο $\tilde{\lambda}_j$ ικανοποιούν το:

$$(6.12) \quad \frac{|\tilde{\lambda}_j - \lambda_j|}{\|A\|} = O(\epsilon_{\text{μηχανής}}).$$

Αυτό δεν είναι καθόλου κακό αποτέλεσμα αν αναλογιστεί κανείς ότι το κόστος σε πράξεις αυτού του αλγορίθμου είναι μόνο $\sim \frac{4}{3}m^3$ πράξεις δηλαδή τα $\frac{2}{3}$ των πράξεων που απαιτούνται για τον υπολογισμό του γινομένου δύο πινάκων $m \times m$.

Αριθμητικά Παραδείγματα-Παρατηρήσεις

Για την κατασκευή των αριθμητικών παραδειγμάτων, δημιουργήθηκαν προγράμματα στο περιβάλλον της MatLab για όλες τις μεθόδους που αναφέρθηκαν νωρίτερα στην εργασία δηλαδή για τις εξής:

1. Μέθοδος των Δυνάμεων
2. Μέθοδος του Πηλίκου Rayleigh
3. Μέθοδος της Αντίστροφης Επανάληψης
4. Απλή Μέθοδος QR
5. "Πρακτική" Μέθοδος QR και
6. Μέθοδος των Ταυτόχρονων Επαναλήψεων.

Αρχικά τα προγράμματα αυτά ήταν αυτοτελή ενώ αργότερα μετατράπηκαν σε ένα ενιαίο για να ικανοποιούν τις ανάγκες που θα αναφερθούν στη συνέχεια. Για την πέμπτη μέθοδο, χρησιμοποιήθηκε ένα εξωτερικό πρόγραμμα κατασκευής του πίνακα μορφής Hessenberg που αντιστοιχεί στον αρχικό πίνακα του οποίου τις ιδιοτιμές αναζητούμε. Σκοπός των προγραμμάτων αυτών που δημιουργήθηκαν ήταν ο έλεγχος του ότι οι αλγόριθμοι τους οποίους μελετάμε όντως συγκλίνουν στις ιδιοτιμές και τα ιδιοδιανύσματα του αρχικού πίνακα υπό τις προϋποθέσεις που αναλύθηκαν κατά την παρουσίασή τους στα αντίστοιχα κεφάλαια. Για το λόγο αυτό χρησιμοποιήθηκαν πίνακες των οποίων ήταν γνωστές οι ιδιοτιμές.

Η διαδικασία που ακολουθήθηκε στην κατασκευή των προγραμμάτων ήταν η εξής:

Αρχικά, επιλέγουμε ένα τυχαίο διαγώνιο πίνακα Λ του οποίου τα διαγώνια στοιχεία είναι οι ιδιοτιμές με αλγεβρική πολλαπλότητα το πλήθος των εμφανίσεών τους στη διαγώνιο και εξυπηρετεί το σκοπό για τον οποίο τον χρησιμοποιούμε. Λόγου χάρη, στο Παράδειγμα 2 χρησιμοποιήθηκε πίνακας με μεγαλύτερη κατ' απόλυτη τιμή αρνητική ενώ στο Παράδειγμα 3 πίνακας με μεγαλύτερη κατ' απόλυτη τιμή ιδιοτιμή θετική και δεύτερη κατά φθίνουσα σειρά κατ' απόλυτη τιμή ιδιοτιμή τέτοια ώστε ο λόγος: $\frac{\lambda_2}{\lambda_1}$ να είναι πολύ μεγαλύτερος από του παραδείγματος 2 και να επαληθεύσουμε ότι η σύγκλιση σύμφωνα με όσα ειπώθηκαν στο 4ο κεφάλαιο να είναι πιο αργή.

Έπειτα, με την εντολή $S = randn(m)$ κατασκευάζεται ένας τυχαίος πίνακας $S \in \mathbb{R}^{m \times m}$, όπου m είναι η διάσταση του αρχικού διαγώνιου πίνακα και των τετραγωνικών πινάκων που κατασκευάζονται από αυτόν. Ελέγχοντας αν ο πίνακας αυτός είναι αντιστρέψιμος, χρησιμοποιούμε τον μετασχηματισμό ομοιότητας $A = S\Lambda S^{-1}$. Με τον τρόπο αυτό κατασκευάζεται ένας τυχαίος πίνακας ίδιων διαστάσεων με το διαγώνιο ο οποίος έχει τυχαία στοιχεία αλλά ίδιες ιδιοτιμές με τον αρχικό πίνακα, οι οποίες είναι γνωστές διότι από αυτές τον κατασκευάσαμε. Οι μέθοδοι εφαρμόζονται, τώρα, στον καινούριο αυτό πίνακα. Όμως η εντολή αυτή δημιουργεί κάθε φορά που εκτελούμε κάποιο πρόγραμμα ένα διαφορετικό πίνακα. Επειδή απαιτείται, όμως, για τη σύγκριση των μεθόδων, αυτός

ο πίνακας να είναι ο ίδιος για όλες τις μεθόδους, δημιουργήθηκε ένα πρόγραμμα που με τον ίδιο αρχικό πίνακα βρίσκει τις ιδιοτιμές και με τις 6 μεθόδους.

Οι αλλαγές που ήταν απαραίτητο να γίνουν για τη μελέτη των αποτελεσμάτων που δίδουν οι μέθοδοι σε διαφορετικές περιπτώσεις που αφορούν τους πίνακες στους οποίους εφαρμόστηκαν, έγιναν στον αρχικό διαγώνιο πίνακα.

Επίσης όπως αναφέρθηκε σε προηγούμενο κεφάλαιο, χρησιμοποιήθηκαν διάφορα κριτήρια τερματισμού για την κατασκευή των προγραμμάτων έτσι ώστε να μη εκτελούνται οι επαναλήψεις συνεχώς.

Για τις τρεις πρώτες μεθόδους, χρησιμοποιήθηκαν δύο διαφορετικά κριτήρια τερματισμού, από τα οποία, τα αποτελέσματα που προέκυψαν δε διέφεραν πολύ μεταξύ τους και έτσι δε θα αναφέρουμε σε κάθε παράδειγμα ποιο χρησιμοποιήθηκε. Αυτά ήταν:

1. $s \leq 10^{-10}$ όπου $s = \|Av^{(k)} - \lambda^{(k)}v^{(k)}\|$ με $v^{(k)}$ η προσέγγιση του ιδιοδιανύσματος στο k βήμα και $\lambda^{(k)}$ η προσέγγιση της ιδιοτιμής που ψάχνουμε στο k βήμα της επανάληψης για τον αρχικό πίνακα A .

2. $s \leq 10^{-10}$ όπου $s = |\lambda^{(k)} - \lambda^{(k-1)}|$ με $\lambda^{(j)}$ η προσέγγιση της ιδιοτιμής που ψάχνουμε στο j βήμα της επανάληψης για τον αρχικό πίνακα.

Για τις τρεις επόμενες, χρησιμοποιήθηκε το εξής κριτήριο τερματισμού:

$s \leq 10^{-10}$ όπου $s = |\max_j A_{j,j}^{(k)} - \max_j A_{j,j}^{(k-1)}|$ όπου A είναι ο πίνακας που θα συγκλίνει στον πίνακα με διαγώνια στοιχεία τις ιδιοτιμές στις Μεθόδους QR (απλή και με αλλαγές) ενώ στη Μέθοδο των Ταυτόχρονων Επαναλήψεων είναι ο πίνακας R της ανάλυσης QR του πίνακα Z που έχει προκύψει από το προηγούμενο βήμα ως το γινόμενο $Z = AQ$ όπου ο A είναι ο αρχικός πίνακας.

Επίσης για τη Μέθοδο του Πηλίκου Rayleigh απαιτήθηκε η γνώση ενός αρχικού διανύσματος η χρήση του οποίου θα μας οδηγούσε στην προσέγγιση του επιθυμητού ιδιοδιανύσματος. Για το λόγο αυτό, αφού βρέθηκε μια προσέγγιση του ιδιοδιανύσματος που αντιστοιχεί στη μεγαλύτερη κατ' απόλυτη τιμή ιδιοτιμή από τη Μέθοδο των Δυνάμεων, υπέστη μια μικρή μεταβολή με τη χρήση της εντολής της MatLab $round(n)$ που στρογγυλοποιεί στον κοντινότερο ακέραιο τον αριθμό n . Επίσης, για τη Μέθοδο της Αντίστροφης Επανάληψης, η αρχική προσέγγιση της ιδιοτιμής που χρησιμοποιήθηκε, επιλέχθηκε κατάλληλα όπως θα φανεί στα παραδείγματα.

7.1. Παράδειγμα 1

Σε αυτό το παράδειγμα χρησιμοποιούμε ένα διαγώνιο πίνακα Λ με τέσσερα διαφορετικά στοιχεία όπου το στοιχείο με μεγαλύτερη απόλυτη τιμή είναι το 19 που είναι θετικός αριθμός.

$$\Lambda = \begin{bmatrix} 10 & 0 & 0 & 0 \\ 0 & 8 & 0 & 0 \\ 0 & 0 & 19 & 0 \\ 0 & 0 & 0 & -6 \end{bmatrix}^T$$

Ο τυχαίος πίνακας που προκύπτει κατά την εκτέλεση του προγράμματος και που είναι όμοιος με το διαγώνιο Λ είναι ο

$$A = \begin{bmatrix} 13.88731491965147 & -2.71774034567642 & -0.87309625757393 & -0.70025003398880 \\ 23.78700609841531 & -5.61136181702393 & -7.73285436145971 & -4.23756313127312 \\ 11.90517348642658 & -6.11329669361690 & -2.24889505126712 & -13.03884374299309 \\ -0.22775335979871 & -1.32031564671054 & 9.51092749293886 & 24.97294194863958 \end{bmatrix}^T$$

Στα δύο παρακάτω πινάκια φαίνονται τα αποτελέσματα που προκύπτουν για τις τρεις πρώτες μεθόδους οι οποίες αποκαλύπτουν μία μόνο ιδιοτιμή και ένα ιδιοδιάνυσμα του πίνακα A . Στη Μέθοδο της Αντίστροφης Επανάληψης χρησιμοποιήθηκε μια προσέγγιση της ιδιοτιμής, το 15 που βρίσκεται πιο κοντά στο 19 απ' ότι στις άλλες ιδιοτιμές του πίνακα. Για το λόγο αυτό η μέθοδος συγκλίνει στην ιδιοτιμή 19. Επίσης, στη Μέθοδο του Πηλίκου Rayleigh χρησιμοποιήθηκε αρχικό διάνυσμα κοντά στο ιδιοδιάνυσμα που αντιστοιχεί στο 19.

	Μέθοδος Δυνάμεων	Μέθοδος Αντίστροφης Επανάληψης ($\mu=15$)
$v^{(0)}$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$
k	40	114
$\lambda^{(k)}$	19.00000000001979	19.00000000002504
$ \lambda^{(k)} - \lambda_J $	$1.979216790459759 \cdot 10^{-11}$	$2.504307872186473 \cdot 10^{-11}$
$v^{(k)}$	-0.02408178782029 -0.00174342447586 -0.53226167143867 0.84623551159461	-0.02408178781915 -0.00174342447435 -0.53226167143835 0.84623551159485

	Μέθοδος Πηλίκου Rayleigh
$v^{(0)}$	$[0 \ 0 \ -0.5 \ 0.8]^T$
k	3
$\lambda^{(k)}$	19.000000000000001
$ \lambda^{(k)} - \lambda_J $	$7.105427357601002 \cdot 10^{-15}$
$v^{(k)}$	0.02408178782456 0.00174342448151 0.53226167143990 -0.84623551159370

Όπως παρατηρούμε, οι τρεις Μέθοδοι συγκλίνουν στην ιδιοτιμή 19 που είναι η μεγαλύτερη κατ' απόλυτη τιμή με πολύ μικρό σφάλμα. Η Μέθοδος του Πηλίκου Rayleigh είναι κατά πολύ ταχύτερη από τις άλλες δύο υπό την έννοια ότι απαιτεί πολύ λιγότερες επαναλήψεις αν και κάθε επανάληψη έχει μεγαλύτερο κόστος πράξεων.

Στα επόμενα δύο πινάκια, καταγράφονται οι πίνακες που περιέχουν τις ιδιοτιμές του αρχικού πίνακα όπως υπολογίζονται από τη Μέθοδο QR , την "Πρακτική" Μέθοδο QR και τη Μέθοδο των Ταυτόχρονων Επαναλήψεων καθώς και τα ιδιοδιανύσματα από τις δύο Μεθόδους από αυτές, καθώς αυτές οι Μέθοδοι υπολογίζουν όλες τις ιδιοτιμές και όλα τα ιδιοδιανύσματα.

Πίνακας που περιέχει τις ιδιοτιμές ως διαγώνια στοιχεία του				
Απλή Μέθοδος QR (k=41)				
19.00000000006627	2.72738945643850	-18.31068372527997	14.98108994333810	$]^T$
-0.00000000021878	9.99996837674994	-2.86412398547390	-28.44869690938315	
0.00000000000001	-0.00002207396444	8.00000957998395	0.97510587619723	
-0.00000000000000	-0.0000000280502	0.00031642513195	-5.99997795680021	
“Πρακτική” Μέθοδος QR (k=38)				
19.00000000006477	-2.72668806075191	15.90673507963452	17.51267537958841	$]^T$
0.00000000021366	10.00189933703330	28.26182116665033	-4.32913237529602	
0.00000000000000	-0.00106193266733	-6.01479819497286	0.72447552478991	
0.00000000000000	0.00000000000000	-0.25005087649958	8.01289885787482	
Μέθοδος Ταυτόχρονων Επαναλήψεων (k=60)				
19.00000000000000	-2.72758863722166	-18.31031401991736	-14.98150554390954	$]^T$
0.00000000000001	9.9999954431718	2.86474793823167	-28.44864249369735	
0.00000000000000	0.00000031813159	8.00000054880853	-0.97447864471344	
-0.00000000000001	0.00000000000017	0.00000133790152	-6.00000009312570	
Πίνακας που περιέχει τα ιδιοδιανύσματα ως στήλες του				
Απλή Μέθοδος QR (k=41)				
0.02408178781024	-0.58905617127646	0.20033618351147	0.78249492532095	$]^T$
0.00174342446255	-0.78035181139139	-0.39215479505096	-0.48709611746334	
0.53226167143579	-0.16892084900862	0.75777504340902	-0.33755035694997	
-0.84623551159673	-0.12461791327696	0.48151533088892	-0.19104664540905	
Μέθοδος Ταυτόχρονων Επαναλήψεων (k=60)				
-0.02408178782456	-0.58905835083549	-0.20034753613351	0.78249037794172	$]^T$
-0.00174342448152	-0.78034754523215	0.39217433949331	-0.48708721658880	
-0.53226167143990	-0.16892909205308	-0.75776554413414	-0.33756755634572	
0.84623551159370	-0.12462315121701	-0.48150963887684	-0.19105757448260	

Και από εδώ παρατηρούμε ότι η “Πρακτική” Μέθοδος QR η οποία έρχεται σε αντιστοιχία με τη Μέθοδο του Πηλίκου Rayleigh συγκλίνει γρηγορότερα στον πίνακα με τις ιδιοτιμές.

7.2. Παράδειγμα 2

Αυτό το παράδειγμα διαφέρει από το προηγούμενο μονάχα ως προς το ότι η μεγαλύτερη κατ' απόλυτη τιμή ιδιοτιμή, αυτή τη φορά, είναι το -15 που είναι αρνητικός αριθμός. Και εδώ, όπως φαίνεται στο επόμενο πινάκιο, η Μέθοδος των Δυνάμεων συγκλίνει όπως αναφέραμε στο Θεώρημα 4.1 στο αντίστοιχο κεφάλαιο στην ιδιοτιμή με μεγαλύτερη απόλυτη τιμή.

Πάλι, Λ είναι ο αρχικός διαγώνιος πίνακας, ενώ A ο πίνακας για τον οποίο βγάζουμε συμπεράσματα κατά την εκτέλεση του προγράμματος.

$$\Lambda = \begin{bmatrix} 10 & 0 & 0 & 0 \\ 0 & -15 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 6 \end{bmatrix}^T$$

$$A = \begin{bmatrix} -11.03932279982248 & -4.33838687219361 & 23.05304652463481 & -14.31658239163972 \\ -17.16754531427765 & -4.20283992257430 & 26.37873919144006 & -22.48782854570656 \\ -1.36132634695765 & -1.31346707670160 & 9.49199650953302 & -2.98821820018655 \\ 1.80850826485135 & -1.81882897144945 & -6.50701775858703 & 6.75016621286376 \end{bmatrix}^T$$

Όπως βλέπουμε, χρησιμοποιήθηκε αρχική προσέγγιση ιδιοτιμής το $\mu=-13$ για τη Μέθοδο της Αντίστροφης Επανάληψης και αρχικό διάνυσμα στη Μέθοδο του Πηλίκου του Rayleigh όπως και στο προηγούμενο παράδειγμα.

	Μέθοδος Δυνάμεων	Μέθοδος Αντίστροφης Επανάληψης ($\mu=-13$)								
$v^{(0)}$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$								
k	63	14								
$\lambda^{(k)}$	-15.00000000013161	-14.99999999999209								
$ \lambda^{(k)} - \lambda_J $	$1.316067255174858 \cdot 10^{-10}$	$7.908340649009915 \cdot 10^{-12}$								
$v^{(k)}$	<table border="1"> <tr><td>-0.57702309961456</td></tr> <tr><td>-0.81150612151015</td></tr> <tr><td>-0.08097359110818</td></tr> <tr><td>-0.04410708339719</td></tr> </table>	-0.57702309961456	-0.81150612151015	-0.08097359110818	-0.04410708339719	<table border="1"> <tr><td>0.57702309961299</td></tr> <tr><td>0.81150612151130</td></tr> <tr><td>0.08097359110937</td></tr> <tr><td>0.04410708339445</td></tr> </table>	0.57702309961299	0.81150612151130	0.08097359110937	0.04410708339445
-0.57702309961456										
-0.81150612151015										
-0.08097359110818										
-0.04410708339719										
0.57702309961299										
0.81150612151130										
0.08097359110937										
0.04410708339445										

	Μέθοδος Πηλίκου Rayleigh				
$v^{(0)}$	$[-0.6 \ -0.8 \ -0.1 \ 0]^T$				
k	5				
$\lambda^{(k)}$	-15.00000000000000				
$ \lambda^{(k)} - \lambda_J $	$3.552713678800501 \cdot 10^{-15}$				
$v^{(k)}$	<table border="1"> <tr><td>-0.57702309961442</td></tr> <tr><td>-0.81150612151036</td></tr> <tr><td>-0.08097359110906</td></tr> <tr><td>-0.04410708339344</td></tr> </table>	-0.57702309961442	-0.81150612151036	-0.08097359110906	-0.04410708339344
-0.57702309961442					
-0.81150612151036					
-0.08097359110906					
-0.04410708339344					

Πίνακας που περιέχει τις ιδιοτιμές ως διαγώνια στοιχεία του				
Απλή Μέθοδος QR (k=63)				
$\left[\begin{array}{cccc} -14.99999999996445 & 34.12180216765140 & 15.07836849184527 & -27.78726134847484 \\ 0.00000000002605 & 9.99999999996420 & 3.78796717379190 & -6.42324404489867 \\ 0.00000000000000 & 0.00000000000026 & 6.00000000000024 & -2.77884999496669 \\ 0 & 0 & 0 & -0.00000000000000 \end{array} \right]^T$				
“Πρακτική” Μέθοδος QR (k=45)				
$\left[\begin{array}{cccc} -14.99999999997485 & -34.12180216758958 & 15.07836849196663 & 27.78726134847285 \\ -0.00000000001841 & 9.99999999996103 & -3.78796717381088 & -6.42324404489719 \\ 0.00000000000000 & -0.00000000001462 & 6.00000000001385 & 2.77884999498976 \\ -0.00000000000000 & & & \end{array} \right]^T$				
Μέθοδος Ταυτόχρονων Επαναλήψεων (k=96)				
$\left[\begin{array}{cccc} -15.00000000000000 & 34.12180216762624 & 15.07836849183910 & -27.78726134846811 \\ -0.00000000000000 & 10.00000000000000 & 3.78796717380736 & -6.42324404492779 \\ -0.00000000000000 & 0.00000000000000 & 6.00000000000000 & -2.77884999496628 \\ 0.00000000000000 & 0.00000000000000 & -0.00000000000000 & -0.00000000000000 \end{array} \right]^T$				

Πίνακας που περιέχει τα ιδιοδιανύσματα ως στήλες του				
Απλή Μέθοδος QR (k=63)				
$\left[\begin{array}{cccc} -0.57702309961438 & 0.03657227880269 & -0.49857911944592 & -0.64585267096074 \\ -0.81150612151042 & -0.05604906988453 & 0.43523028081219 & 0.38586386094105 \\ -0.08097359110929 & -0.22800755150885 & -0.73840273272353 & 0.62945789239967 \\ -0.04410708339243 & 0.97135653940415 & -0.12944059882088 & 0.19433519992093 \end{array} \right]^T$				
Μέθοδος Ταυτόχρονων Επαναλήψεων (k=96)				
$\left[\begin{array}{cccc} -0.57702309961442 & 0.03657227880206 & -0.49857911944592 & -0.64585267096074 \\ -0.81150612151036 & -0.05604906988535 & 0.43523028081219 & 0.38586386094105 \\ -0.08097359110906 & -0.22800755150898 & -0.73840273272351 & 0.62945789239967 \\ -0.04410708339344 & 0.97135653940409 & -0.12944059882094 & 0.19433519992093 \end{array} \right]^T$				

Τα συμπεράσματα για όλες τις Μεθόδους είναι τα ίδια με τα προηγούμενα.

7.3. Παράδειγμα 3

Σκοπός αυτού του παραδείγματος είναι να μελετήσουμε τη συμπεριφορά των αλγορίθμων στην περίπτωση που οι δύο κατ' απόλυτη τιμή μεγαλύτερες ιδιοτιμές είναι πολύ κοντά. Για το λόγο αυτό χρησιμοποιούμε τον παρακάτω διαγώνιο πίνακα :

$$\Lambda = \begin{bmatrix} 10.00000000000000 & 0 & 0 & 0 \\ 0 & 8.00000000000000 & 0 & 0 \\ 0 & 0 & 10.10000000000000 & 0 \\ 0 & 0 & 0 & -6.00000000000000 \end{bmatrix}^T$$

Ο πίνακας A που προκύπτει έχει και αυτός τις ίδιες ιδιοτιμές.

$$A = \begin{bmatrix} 17.80391985165965 & 14.90627053237478 & 21.77082075153276 & -7.96643862279957 \\ -1.74546820877299 & 6.92972964985555 & -4.85189415785974 & 1.47667948157718 \\ -8.78464286601563 & -16.53907717047621 & -14.41735046038064 & 8.70052125997953 \\ -4.31719732091809 & -5.83132899111769 & -10.75951486441894 & 11.78370095886544 \end{bmatrix}^T$$

	Μέθοδος Δυνάμεων	Μέθοδος Αντίστροφης Επανάληψης (μ=13)
$v^{(0)}$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$
k	2025	595
$\lambda^{(k)}$	10.10000000001952	10.10000000001909
$ \lambda^{(k)} - \lambda_J $	$1.951505623765115 \cdot 10^{-11}$	$1.909050695303449 \cdot 10^{-11}$
$v^{(k)}$	0.66893357786816 0.50398541122803 -0.51148846796144 0.19210965830036	0.66893357785529 0.50398541124079 -0.51148846796120 0.19210965831231

	Μέθοδος Πηλίκου Rayleigh
$v^{(0)}$	$[0.7 \ 0.5 \ -0.5 \ 0.2]^T$
k	5
$\lambda^{(k)}$	10.099999999999999
$ \lambda^{(k)} - \lambda_J $	$5.329070518200751 \cdot 10^{-15}$
$v^{(k)}$	0.66893357727699 0.50398541181433 -0.51148846795060 0.19210965884956

Βάσει των παραπάνω αποτελεσμάτων, παρατηρούμε ότι οι Μέθοδοι είναι σταθερές μεταξύ τους υπό την έννοια ότι η τρίτη Μέθοδος είναι ταχύτερη από τις άλλες δύο όπως και στα προηγούμενα αποτελέσματα. Το ίδιο μπορούμε να παρατηρήσουμε και για τις Μεθόδους που βρίσκουν όλες τις ιδιοτιμές του πίνακα.

Πίνακας που περιέχει τις ιδιοτιμές ως διαγώνια στοιχεία του				
Απλή Μέθοδος QR (k=1444)				
10.10000000997868	-0.01956746884806	0.13937425996586	-29.93554103045473	T
0.00000005099612	9.99999999002139	-2.07784714022218	28.00835816400451	
-0.00000000000000	0.00000000000000	7.99999999999999	10.39143695534997	
0.00000000000000	0.00000000000000	-0.00000000000000	-6.00000000000000	
“Πρακτική” Μέθοδος QR (k=1310)				
10.10000000896599	-0.01956747402319	-24.00159402936331	17.89076407559378	T
0.00000004582097	9.99999999103395	23.62754904239677	-15.18303621947145	
0	-0.00000000000000	-5.92989903491995	10.48457372858432	
0	0	0.09313677323415	7.92989903491995	
Μέθοδος Ταυτόχρονων Επαναλήψεων (k=2306)				
10.10000000000187	-0.01956751983457	0.13937320054386	29.93552674996126	T
0.000000000000961	9.99999999999812	-2.07784721128400	-28.00837342709184	
0.00000000000000	0	8.00000000000000	-10.39143695534997	
0.000000000000001	-0.00000000000000	-0.00000000000000	-5.99999999999998	

Πίνακας που περιέχει τα ιδιοδιανύσματα ως στήλες του				
Απλή Μέθοδος QR (k=1444)				
0.66893387950980	-0.59265810508444	0.36342051460044	-0.26307672814084	T
0.50398511206822	0.58778229490663	-0.35054509343733	-0.52690522684342	
-0.51148847349220	0.01086652518445	0.41429406069328	-0.75274291188249	
0.19210937806960	0.55059083074677	0.75723450681333	0.29417618130376	
Μέθοδος Ταυτόχρονων Επαναλήψεων (k=2306)				
-0.66893357733392	0.59265844615048	-0.36342051460044	-0.26307672814084	T
-0.50398541175787	-0.58778203794208	0.35054509343733	-0.52690522684342	
0.51148846795164	-0.01086678597463	-0.41429406069328	-0.75274291188249	
-0.19210965879668	-0.55059073279681	-0.75723450681333	0.29417618130376	

Η διαφορά έγκειται στο γεγονός ότι για την ίδια απαιτούμενη ακρίβεια με πριν το πλήθος των επαναλήψεων σε όλες σχεδόν τις μεθόδους είναι αυξημένο.

7.4. Παράδειγμα 4

Το αντίθετο συμβαίνει όταν ο λόγος $\frac{|\lambda_2|}{|\lambda_1|}$ είναι πολύ μικρός. Για να το δούμε αυτό, κατασκευάσαμε το επόμενο παράδειγμα.

$$\Lambda = \begin{bmatrix} 100 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 6 & 0 \\ 0 & 0 & 0 & -6 \end{bmatrix}^T$$

$$A = \begin{bmatrix} 41.023832511728 & 14.871831727460 & 24.241560577479 & -13.136199582132 \\ 177.186284586920 & 52.216874397871 & 116.599937502275 & -61.026951144163 \\ -13.744731458919 & -4.068042046265 & -9.072821834156 & 3.299519222683 \\ -25.685963657157 & -9.012991084899 & -15.236496850209 & 15.832114924557 \end{bmatrix}^T$$

	Μέθοδος Δυνάμεων	Μέθοδος Αντίστροφης Επανάληψης (μ=60)
$v^{(0)}$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$
k	10	91
$\lambda^{(k)}$	$1.000000000000639 \cdot 10^2$	$1.000000000000154 \cdot 10^2$
$ \lambda^{(k)} - \lambda_J $	$6.387779194483301 \cdot 10^{-11}$	$1.539035565656377 \cdot 10^{-11}$
$v^{(k)}$	0.24729096412891 0.95219018226769 -0.07165064304500 -0.16446039402983	0.24729096412806 0.95219018226775 -0.07165064304470 -0.16446039403093

	Μέθοδος Πηλίκου Rayleigh
$v^{(0)}$	$[0.2 \ 1 \ -0.1 \ -0.2]^T$
k	5
$\lambda^{(k)}$	99.99999999999997
$ \lambda^{(k)} - \lambda_J $	$2.842170943040401 \cdot 10^{-14}$
$v^{(k)}$	0.24729096412839 0.95219018226780 -0.07165064304491 -0.16446039403006

Πίνακας που περιέχει τις ιδιοτιμές ως διαγώνια στοιχεία του				
Απλή Μέθοδος QR (k=11)				
$\begin{bmatrix} 99.999999999996 & -170.243179289388 & -5.877854172185 & -128.341959614707 \\ -0.000000000002 & -5.839627573454 & -1.538465775602 & -7.707141200508 \\ 0.000000000001 & -1.234183973036 & 5.839627573458 & 2.791650844388 \\ -0.000000000000 & 0.000000000000 & -0.000000000000 & 0.000000000000 \end{bmatrix}^T$				
“Πρακτική” Μέθοδος QR (k=12)				
$\begin{bmatrix} 1.000000000002 & 169.851942236363 & -12.946311496698 & 128.341959620132 \\ -0.000000000001 & -6.001502808249 & -0.223631443166 & -7.351972129158 \\ 0 & 0.080650363167 & 6.001502810042 & 3.625168369783 \\ 0 & -0.000000000000 & -0.000000003052 & -0.00000001794 \end{bmatrix}^T$				
Μέθοδος Ταυτόχρονων Επαναλήψεων (k=16)				
$\begin{bmatrix} 100.000000000000 & -167.779255074715 & 29.451839175191 & -128.341959614706 \\ -0.000000000000 & -5.901186741240 & 0.942912475888 & -6.963318620838 \\ -0.000000000000 & 1.247194278455 & 5.901186741240 & 4.324989445800 \\ -0.000000000000 & 0.000000000000 & 0.000000000000 & 0.000000000000 \end{bmatrix}^T$				

Πίνακας που περιέχει τα ιδιοδιανύσματα ως στήλες του			
Απλή Μέθοδος QR (k=11)			
$\begin{bmatrix} -0.24729096412838 & 0.84266859532101 & 0.46720693841526 & 0.10234497650991 \\ -0.95219018226780 & -0.29498204320094 & 0.01079506309199 & 0.07875860330561 \\ 0.07165064304491 & 0.02099336896880 & -0.21332845534347 & 0.97412341823373 \\ 0.16446039403005 & -0.44994945394795 & 0.85795810704850 & 0.18548895987950 \end{bmatrix}^T$			
Μέθοδος Ταυτόχρονων Επαναλήψεων (k=16)			
$\begin{bmatrix} -0.24729096412839 & 0.92106505948452 & 0.28286364390006 & 0.10234497650991 \\ -0.95219018226780 & -0.28637462695436 & 0.07155775454080 & 0.07875860330561 \\ 0.07165064304491 & -0.02357212120140 & -0.21305892729937 & 0.97412341823373 \\ 0.16446039403006 & -0.26281758756902 & 0.93245565053932 & 0.18548895987950 \end{bmatrix}^T$			

Είναι φανερό πως ενώ οι αποκλίσεις των ιδιοτιμών που υπολογίζονται από τις πραγματικές είναι περίπου οι ίδιες με τις αποκλίσεις που προέκυψαν στο προηγούμενο παράδειγμα, εντούτοις, το πλήθος των επαναλήψεων μειώθηκε πάρα πολύ.

Επίσης, στον παραπάνω πίνακα υπάρχουν δυο ιδιοτιμές με ίδια απόλυτη τιμή. Αυτό δεν επηρέασε τα αποτελέσματα των Μεθόδων και αυτό διότι δεν είχαν τη μεγαλύτερη απόλυτη τιμή (για τη Μέθοδο των Δυνάμεων) και διότι με τη Μέθοδο του Πηλίκου Rayleigh και τη Μέθοδο της Αντίστροφης Επανάληψης δεν προσεγγίζουμε κάποια από αυτές τις δυο ιδιοτιμές.

7.5. Παράδειγμα 5

Το παράδειγμα αυτό έχει κατασκευαστεί ώστε να δείξει ότι η Μέθοδος της Αντίστροφης Επανάληψης και η Μέθοδος του Πηλίκου Rayleigh, δε συγκλίνουν πάντα στη μεγαλύτερη κατ' απόλυτη τιμή ιδιοτιμή αλλά στην κοντινότερη στο μ ιδιοτιμή και στην ιδιοτιμή που είναι πιο κοντά στο πηλίκο Rayleigh του αρχικού διανύσματος που χρησιμοποιούμε αντίστοιχα.

Για το λόγο αυτό χρησιμοποιήσαμε πάλι σαν διαγώνιο πίνακα τον:

$$\Lambda = \begin{bmatrix} 10.00000000000000 & 0 & 0 & 0 \\ 0 & 8.00000000000000 & 0 & 0 \\ 0 & 0 & 10.10000000000000 & 0 \\ 0 & 0 & 0 & -6.00000000000000 \end{bmatrix}^T$$

και προκύπτει κατά την εκτέλεση του προγράμματος ο πίνακας:

$$A = \begin{bmatrix} 8.52105191407043 & -0.34059910630167 & 0.50010384212671 & -1.46815272773727 \\ 1.75722304792264 & 4.50731647367719 & 21.21894685809730 & 2.31712091953819 \\ -1.65563856436586 & 1.80117084210486 & 2.71229561166551 & -1.77193885750268 \\ -3.37387126664346 & 4.13860388730768 & -16.54040299918964 & 6.35933600058686 \end{bmatrix}^T$$

με τις ίδιες ιδιοτιμές.

	Μέθοδος Δυνάμεων	Μέθοδος Αντίστροφης Επανάληψης ($\mu=10.002$)
$v^{(0)}$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$
\mathbf{k}	2131	6
$\lambda^{(k)}$	10.1000000000112	9.99999999999995
$ \lambda^{(k)} - \lambda_J $	$1.120881165661558 \cdot 10^{-12}$	0.10000000000005
$v^{(k)}$	-0.54205853638073 0.66514297063861 0.16717375079030 0.48560303622766	0.44384686394794 0.68773565164912 0.20007414146390 -0.53850717054819
Μέθοδος Πηλίκου Rayleigh		
$v^{(0)}$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$	
\mathbf{k}	6	
$\lambda^{(k)}$	8.000000000000000	
$ \lambda^{(k)} - \lambda_J $	2.100000000000000	
$v^{(k)}$	-0.57702309961442 -0.81150612151036 -0.08097359110906 -0.04410708339344	

Από τα παραπάνω πινάκια, παρατηρούμε ότι η Μέθοδος των Δυνάμεων απαιτεί πολύ περισσότερες επαναλήψεις από τις άλλες δύο Μεθόδους. Αυτό συμβαίνει διότι τώρα ο λόγος $\frac{\lambda_2}{\lambda_1}$ είναι πολύ μεγάλος για αυτή τη μέθοδο αλλά οι αντιστοιχοί λόγοι για τους πίνακες που χρησιμοποιεί η παραλλαγή της Μεθόδου των Δυνάμεων στις άλλες δύο Μεθόδους είναι πολύ μικρότεροι. Αυτό οφείλεται στο γεγονός ότι η Μέθοδος της Αντίστροφης Επανάληψης θα προσεγγίσει την ιδιοτιμή 10 που είναι πιο κοντά στο 10.002 και ο λόγος θα είναι $\frac{10.002-10}{10.1-10.002} = \frac{0.002}{0.008} = 0.25$ αρκετά μικρότερος από αυτόν της Μεθόδου των Δυνάμεων. Επίσης, το αρχικό διάνυσμα που χρησιμοποιήθηκε στη Μέθοδο του Πηλίκου Rayleigh είναι τέτοιο ώστε το πηλίκο Rayleigh που αντιστοιχεί σε αυτό να είναι πολύ κοντινότερο σε άλλη ιδιοτιμή. Στη συγκεκριμένη περίπτωση, στο 8.

Πίνακας που περιέχει τις ιδιοτιμές ως διαγώνια στοιχεία του				
Απλή Μέθοδος QR (k=1053)				
$]^T$				
10.10000000993534	0.00111938476948	-1.04907701940435	-4.62577523630907	
-0.00000088757409	9.99999999006472	-3.08303766112806	-23.08591769625436	
0.00000000000000	0.00000000000000	8.00000000000001	6.79647905125826	
0.00000000000000	-0.00000000000000	0.00000000000000	-5.99999999999999	
“Πρακτική” Μέθοδος QR (k=958)				
$]^T$				
10.10000000891866	0.00111947566819	-3.57158208453044	-3.12121138686903	
-0.00000079667538	9.9999999108143	-18.84774637595378	-13.68310541022476	
0	-0.00000000000000	-5.68032851720965	-7.38838272803482	
0	0.00000000000000	-0.59190367677641	7.68032851720965	
Μέθοδος Ταυτόχρονων Επαναλήψεων (k=2070)				
$]^T$				
10.10000000000040	-0.00112027230781	-1.04904965622478	4.62557033977902	
0.00000000003576	9.9999999999960	3.08304697196633	-23.08595875086919	
0.00000000000000	-0.00000000000000	8.00000000000000	-6.79647905125821	
0.00000000000000	-0.00000000000000	-0.00000000000000	-6.00000000000000	
Πίνακας που περιέχει τα ιδιοδιανύσματα ως στήλες				
Απλή Μέθοδος QR (k=1053)				
$]^T$				
-0.54205465097887	0.43780699399295	0.69407834834214	-0.18098905415993	
0.66514914059986	0.69522431509945	0.14162897218290	0.23276813934091	
0.16717554312175	0.20195801331096	-0.24052947576684	-0.93456453491921	
0.48559830507682	-0.53310519465192	0.66358273929360	-0.19912580358481	
Μέθοδος Ταυτόχρονων Επαναλήψεων (k=2070)				
$]^T$				
-0.54205853666199	-0.43780218303339	0.69407834834214	0.18098905415993	
0.66514297019198	-0.69523021852504	0.14162897218290	-0.23276813934091	
0.16717375066056	-0.20195949704988	-0.24052947576684	0.93456453491921	
0.48560303657014	0.53310088476026	0.66358273929359	0.19912580358481	

Όπως παρατηρούμε, οι τρεις Μέθοδοι που βρίσκουν όλες τις ιδιοτιμές, βρίσκουν πάλι σε μεγάλο πλήθος επαναλήψεων τα σωστά αποτελέσματα (με την ακρίβεια που έχουμε θέσει ως κριτήριο τερματισμού στα προγράμματα).

7.6. Παράδειγμα 6

Με τον ίδιο διαγώνιο πίνακα με πριν θελήσαμε να ελέγξουμε τη συμπεριφορά κυρίως της Μεθόδου της Αντίστροφης Επανάληψης όταν η αρχική προσέγγιση της ιδιοτιμής που βρίσκει αυτή η Μέθοδος, ισαπέχει από δυο ιδιοτιμές του πίνακα. Έτσι, ενώ έχουμε τις ιδιοτιμές 10 και 10.1, επιλέγουμε $\mu=10.05$.

$$\Lambda = \begin{bmatrix} 10.00000000000000 & 0 & 0 & 0 \\ 0 & 8.00000000000000 & 0 & 0 \\ 0 & 0 & 10.10000000000000 & 0 \\ 0 & 0 & 0 & -6.00000000000000 \end{bmatrix}^T$$

$$A = \begin{bmatrix} 8.89334152505066 & -7.31951052844119 & 1.95071102212389 & -4.74043604036856 \\ -1.08959907912976 & 3.64147393735673 & 1.77637464448734 & -4.12106392348052 \\ 1.30577784069595 & 4.97876977287589 & 7.58191352980218 & 2.53699404368284 \\ -1.23875734976491 & -10.87453348613229 & 1.60272293737561 & 1.98327100779043 \end{bmatrix}^T$$

	Μέθοδος Δυνάμεων	Μέθοδος Πηλίκου Rayleigh
$v^{(0)}$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$	$[0.9 \ -0.1 \ 0.3 \ 0.1]^T$
k	2511	5
$\lambda^{(k)}$	10.09999999995248	10.10000000000000
$ \lambda^{(k)} - \lambda_J $	$4.751576909711730 \cdot 10^{-11}$	$1.776356839400251 \cdot 10^{-15}$
$v^{(k)}$	0.93204261098466 -0.11324392255936 0.33575434221916 0.07577207264717	0.93204261102249 -0.11324392313248 0.33575434176720 0.07577207332794

Πίνακας που περιέχει τις ιδιοτιμές ως διαγώνια στοιχεία του				
Απλή Μέθοδος QR (k=1178)				
10.10000000991910	-0.04757351643033	-0.38010368947319	-6.50932933612158	$]^T$
0.0000002085018	9.9999999008087	0.28305700496673	7.83530817893945	
0.00000000000000	0.00000000000000	8.00000000000001	2.05833703016838	
-0.00000000000000	-0.00000000000000	0.00000000000000	-6.00000000000000	
“Πρακτική” Μέθοδος QR (k=1070)				
10.10000000895879	0.04757351844907	-6.32340352938184	-1.59072718021537	$]^T$
-0.0000001883144	9.9999999104124	-7.64413797531597	-1.74336806352441	
0.00000000000000	-0.00000000000000	-5.88851029547080	-2.64397925991877	
-0.00000000000000	-0.00000000000000	-0.58564222975039	7.88851029547079	
Μέθοδος Ταυτόχρονων Επανάληψεων (k=2162)				
10.10000000000055	-0.04757353727934	-0.38010363045857	6.50932770253666	$]^T$
0.0000000000117	9.9999999999945	0.28305708421459	-7.83530953607019	
0.00000000000000	0.00000000000000	8.00000000000001	-2.05833703016838	
0.00000000000000	-0.00000000000000	0.00000000000000	-6.00000000000000	
Πίνακας που περιέχει τα ιδιοδιανύσματα ως στήλες του				
Απλή Μέθοδος QR (k=1178)				
0.93204261891966	-0.03787545336083	-0.33944725689065	0.12098581048866	$]^T$
-0.11324404277174	0.57380444204780	-0.08755441766689	0.80638605705329	
0.33575424742133	0.45249444388814	0.80459421344521	-0.18747270542509	
0.07577221544034	-0.68158835859285	0.47930985325456	0.54768416571146	
Μέθοδος Ταυτόχρονων Επανάληψεων (k=2162)				
-0.93204261102293	0.03787564768251	0.33944725689065	0.12098581048866	$]^T$
0.11324392313917	-0.57380446565805	0.08755441766689	0.80638605705329	
-0.33575434176193	-0.45249437388668	-0.80459421344521	-0.18747270542508	
-0.07577207333588	0.68158837439060	-0.47930985325456	0.54768416571146	

Αποτέλεσμα αυτού του πειράματος ήταν ότι καμία Μέθοδος δεν επηρεάστηκε εκτός από την Μέθοδο της Αντίστροφης Επανάληψης η οποία λειτουργούσε ατέρμονα.

7.7. Παράδειγμα 7

Αυτό το παράδειγμα αποτελείται από δυο υποπαράδειγματα τα οποία δείχνουν την προβληματική κατάσταση που υπάρχει όταν προσπαθούμε να προσεγγίσουμε μια διπλή ιδιοτιμή. Στην πρώτη περίπτωση έχει χρησιμοποιηθεί πίνακας με 1 γραμμικώς ανεξάρτητα ιδιοδιανύσματα να αντιστοιχεί στην μεγαλύτερη κατ' απόλυτη τιμή ιδιοτιμή πολλαπλότητας 2, ενώ στη δεύτερη δύο.

Υποπαράδειγμα 1

$$A = \begin{bmatrix} 3 & -4 & 0 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}^T$$

	Μέθοδος Δυνάμεων	Μέθοδος Αντίστροφης Επανάληψης ($\mu=3$)								
$v^{(0)}$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$								
k	44722	89444								
$\lambda^{(k)}$	1.00002236065955	1.00002236065956								
$ \lambda^{(k)} - \lambda_J $	$2.236065955041333 \cdot 10^{-5}$	$2.236065955729671 \cdot 10^{-5}$								
$v^{(k)}$	<table border="1"> <tr><td>-0.89442919098916</td></tr> <tr><td>-0.44720959549910</td></tr> <tr><td>0</td></tr> <tr><td>0</td></tr> </table>	-0.89442919098916	-0.44720959549910	0	0	<table border="1"> <tr><td>-0.89442919098916</td></tr> <tr><td>-0.44720959549910</td></tr> <tr><td>0.00000000000000</td></tr> <tr><td>0.00000000000000</td></tr> </table>	-0.89442919098916	-0.44720959549910	0.00000000000000	0.00000000000000
-0.89442919098916										
-0.44720959549910										
0										
0										
-0.89442919098916										
-0.44720959549910										
0.00000000000000										
0.00000000000000										

Πίνακας που περιέχει τις ιδιοτιμές ως διαγώνια στοιχεία του

Απλή Μέθοδος QR (k=100001)				
$\left[\begin{array}{cccc} 1.00000999987959 & 4.9999999998298 & 0 & 0 \\ -0.0000000002000 & 0.99999000012046 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{array} \right]^T$				

Πίνακας που περιέχει τα ιδιοδιανύσματα ως στήλες του

Απλή Μέθοδος QR (k=100001)				
$\left[\begin{array}{cccc} -0.89442808541483 & -0.44721180666623 & 0 & 0 \\ -0.44721180666636 & 0.89442808541457 & 0 & 0 \\ 0 & 0 & 1.00000000000000 & 0 \\ 0 & 0 & 0 & 1.00000000000000 \end{array} \right]^T$				

Υποπαράδειγμα 2

$$A = \begin{bmatrix} 2 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^T$$

	Μέθοδος Δυνάμεων	Μέθοδος Αντίστροφης Επανάληψης (μ=3)								
$v^{(0)}$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$	$\frac{1}{2} [1 \ 1 \ 1 \ 1]^T$								
k	33	33								
$\lambda^{(k)}$	2.000000000000000	2.000000000000000								
$ \lambda^{(k)} - \lambda_J $	$4.440892098500626 \cdot 10^{-16}$	$4.440892098500626 \cdot 10^{-16}$								
$v^{(k)}$	<table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>0.70710678118655</td></tr> <tr><td>0.70710678118655</td></tr> <tr><td>0</td></tr> <tr><td>0.000000000008232</td></tr> </table>	0.70710678118655	0.70710678118655	0	0.000000000008232	<table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>-0.70710678118655</td></tr> <tr><td>-0.70710678118655</td></tr> <tr><td>-0.000000000000000</td></tr> <tr><td>-0.000000000008232</td></tr> </table>	-0.70710678118655	-0.70710678118655	-0.000000000000000	-0.000000000008232
0.70710678118655										
0.70710678118655										
0										
0.000000000008232										
-0.70710678118655										
-0.70710678118655										
-0.000000000000000										
-0.000000000008232										

	Μέθοδος Πηλίκου Rayleigh				
$v^{(0)}$	$[0.7 \ 0.7 \ 0 \ 0]^T$				
k	1				
$\lambda^{(k)}$	2.000000000000000				
$ \lambda^{(k)} - \lambda_J $	$4.440892098500626 \cdot 10^{-16}$				
$v^{(k)}$	<table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>-0.70710678118655</td></tr> <tr><td>-0.70710678118655</td></tr> <tr><td>0</td></tr> <tr><td>0</td></tr> </table>	-0.70710678118655	-0.70710678118655	0	0
-0.70710678118655					
-0.70710678118655					
0					
0					

Πίνακας που περιέχει τις ιδιοτιμές ως διαγώνια στοιχεία του					
Απλή Μέθοδος QR (k=1)					
	<table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>$2 \ 0 \ 0 \ 0$</td></tr> <tr><td>$0 \ 2 \ 0 \ 0$</td></tr> <tr><td>$0 \ 0 \ 0 \ 0$</td></tr> <tr><td>$0 \ 0 \ 0 \ 1$</td></tr> </table>	$2 \ 0 \ 0 \ 0$	$0 \ 2 \ 0 \ 0$	$0 \ 0 \ 0 \ 0$	$0 \ 0 \ 0 \ 1$
$2 \ 0 \ 0 \ 0$					
$0 \ 2 \ 0 \ 0$					
$0 \ 0 \ 0 \ 0$					
$0 \ 0 \ 0 \ 1$					
Μέθοδος Ταυτόχρονων Επαναλήψεων (k=1)					
	<table border="1" style="margin-left: auto; margin-right: auto;"> <tr><td>$2 \ 0 \ 0 \ 0$</td></tr> <tr><td>$0 \ 2 \ 0 \ 0$</td></tr> <tr><td>$0 \ 0 \ 0 \ 0$</td></tr> <tr><td>$0 \ 0 \ 0 \ 1$</td></tr> </table>	$2 \ 0 \ 0 \ 0$	$0 \ 2 \ 0 \ 0$	$0 \ 0 \ 0 \ 0$	$0 \ 0 \ 0 \ 1$
$2 \ 0 \ 0 \ 0$					
$0 \ 2 \ 0 \ 0$					
$0 \ 0 \ 0 \ 0$					
$0 \ 0 \ 0 \ 1$					

Πίνακας που περιέχει τα ιδιοδιανύσματα ως στήλες του	
Απλή Μέθοδος QR (k=1)	
$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^T$	
Μέθοδος Ταυτόχρονων Επαναλήψεων(k=1)	
$\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{bmatrix}^T$	

Παρατηρούμε ότι στην περίπτωση των δύο γραμμικώς ανεξάρτητων ιδιοδιανυσμάτων, περισσότερες μέθοδοι λειτουργούν σωστά και σε πολύ λιγότερες επαναλήψεις απ' ότι στην περίπτωση όπου υπάρχει ένα γραμμικώς ανεξάρτητο ιδιοδιάνυσμα. Παρ' όλα αυτά και οι δύο περιπτώσεις είναι προβληματικές.

7.8. Παρατηρήσεις

Για το λόγο αυτό είναι απαραίτητη η γνώση ορισμένων βασικών χαρακτηριστικών - ιδιοτήτων των πινάκων των οποίων τις ιδιοτιμές ψάχνουμε να βρούμε ότως ώστε να γίνει σωστή επιλογή της κατάλληλης αριθμητικής μεθόδου.

Υπάρχουν και άλλες Μέθοδοι υπολογισμού ιδιοτιμών και ιδιοδιανυσμάτων πινάκων οι οποίες δημιουργήθηκαν για να αντιμετωπιστούν προβλήματα όπως αυτά που παρουσιάστηκαν παραπάνω, αλλά δε θα αναφερθούν σε αυτή την εργασία.

Βιβλιογραφία

- [1] *NUMERICAL LINEAR ALGEBRA*, Lloyd N. Trefethen - David Bau III, Siam, Philadelphia, 1997
- [2] *Applied Numerical Linear Algebra*, J.W. Demmel, Siam, Philadelphia, 1997
- [3] *Accuracy and Stability of Numerical Algorithms*, Nicholas J. Higham, Siam
- [4] *Numerical Computation 1 - Methods, Software, and Analysis*, Christoph W. Ueberhuber, Springer, Verlag Berlin and Heidelberg, 1997
- [5] *Numerical Computation 2 - Methods, Software, and Analysis*, Christoph W. Ueberhuber, Springer, Verlag Berlin and Heidelberg, 1997
- [6] *Afternotes on Numerical Analysis*, G.W. Stewart, Siam, 1996
- [7] *Elements of Numerical Analysis*, James Singer, Academic Press
- [8] *Το Θεμελιώδες Θεώρημα της Άλγεβρας*, Benjamin Fine- Gerhard Rosenberg, Leader Books, Αθήνα, 2001
- [9] *Γραμμική Άλγεβρα και Εφαρμογές*, Gilbert Strang, Πανεπιστημιακές Εκδόσεις Κρήτης, Ηράκλειο, 2004
- [10] *Εισαγωγή στην Αριθμητική Ανάλυση*, Γ.Δ. Ακρίβης- Β.Α. Δουγαλής, Πανεπιστημιακές Εκδόσεις Κρήτης, Ηράκλειο, 2005
- [11] *Σημειώσεις Αριθμητικής Γραμμικής Άλγεβρας*, Β. Δουγαλής- Δ.Νούτσος- Α. Χατζηδήμος, Ηράκλειο, 2006