ΕΛΛΗΝΙΚΗ ΔΗΜΟΚΡΑΤΙΑ
Εθνικόν και Καποδιστριακόν
Πανεπιστήμιον Αθηνών

Αριθμητική Ανάλυση

Ενότητα 4

Αριθμητικός Υπολογισμός Ιδιοτιμών και Ιδιοδιανυσμάτων

Ν. Μ. Μισυρλής

Τμήμα Πληροφορικής και Τηλεπικοινωνιών,

Κεφ.5. Αριθμητικός Υπολογισμός Ιδιοτιμών και Ιδιοδιανυσμάτων

Δίνεται ένας πίνακας $\mathbf{A} \in \mathbb{C}^{\mathbf{n}\times\mathbf{n}}$ και ζητούνται να προσδιορισθούν οι ιδιοτιμές $\lambda_i,\ i = 1(1)n$ και τα αντίστοιχα ιδιοδιανύσματα $\mathbf{x}^{(\mathbf{i})},\ \mathbf{i} = \mathbf{1}(\mathbf{1})\mathbf{n}$ του πίνακα $\mathbf{A}$.

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$$

$$\mathbf{det}(\mathbf{A} - \lambda\mathbf{I}) = \mathbf{0}$$

Θεωρητικός προσδιορισμός των Ιδιοτιμών και Ιδιοδιανυσμάτων

- Επίλυση της πολυωνυμικής εξίσωσης *(εύρεση ιδιοτιμών)*

$$\mathbf{det}(\mathbf{A} - \lambda\mathbf{I}) = \underbrace{(-1)^n\lambda^n + c_1\lambda^{n-1} + \cdots + c_{n-1}\lambda + c_n}_{\mathbf{p}(\lambda)} = \mathbf{0} \qquad (1)$$

- Επίλυση του ομογενούς γραμμικού συστήματος *(εύρεση ιδιοδιανυσμάτων)*

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \qquad (2)$$

Μηχανή Αναζήτησης: Ιδιοτιμές και Γραφήματα

- Η μηχανή αναζήτησης Google χρησιμοποιεί την ακόλουθη μέθοδο για την αναζήτηση ενός ιστοτόπου. Δημιουργεί τον *πίνακα γειτνίασης* A με στοιχεία, τα οποία είτε είναι 0 ή 1, με $a_{ij} = 1$ αν ο ιστότοπος i συνδέεται με τον ιστότοπο j, διαφορετικά $a_{ij} = 0$.
- Για παράδειγμα, μια εταιρεία με επτά υπαλλήλους επιθυμεί την όσο το δυνατόν μεγαλύτερη χρήση των ιστοτόπων. Κάθε υπάλληλος έχει ένα ιστότοπο και κάποιοι υπάλληλοι έχουν συνδέσμους σε ιστοτόπους συναδέλφων τους. Ο *πίνακας γειτνίασης* είναι ο ακόλουθος

	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7
Y_1	0	1	1	0	1	0	1
Y_2	1	0	0	0	1	1	0
Y_3	0	1	0	1	0	1	0
Y_4	1	1	0	0	0	1	0
Y_5	1	0	0	1	0	0	0
Y_6	0	1	0	1	1	0	0
Y_7	0	0	1	0	0	1	1

όπου με $Y_k,\ k = 1, 2, \ldots, 7$ συμβολίζεται ο κάθε υπάλληλος.

	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7
Y_1	0	1	1	0	1	0	1
Y_2	1	0	0	0	1	1	0
Y_3	0	1	0	1	0	1	0
Y_4	1	1	0	0	0	1	0
Y_5	1	0	0	1	0	0	0
Y_6	0	1	0	1	1	0	0
Y_7	0	0	1	0	0	1	1

- Παρατηρώντας τη γραμμή Y_3 διαπιστώνουμε ότι ο ιστότοπος του υπαλλήλου Y_3 συνδέεται με τους ιστοτόπους των συναδέλφων του Y_2, Y_4 και Y_6.
- Για την κατάταξη των ιστοτόπων ανάλογα με την συχνότητα αναφοράς τους απαιτείται ο υπολογισμός της μεγαλύτερης κατά μέτρο ιδιοτιμής και του αντίστοιχου ιδιοδιανύσματος.
- Για τον ανωτέρω πίνακα το ιδιοδιάνυσμα αυτό είναι το $z = [0.4261,\ 0.4746,\ 0.2137,\ 0.3596,\ 0.4416,\ 0.4214,\ 0.2137]^T$. Αν η $k-$ιοστή συνιστώσα είναι η μεγαλύτερη, τότε ο $k-$ιοστός ιστότοπος θεωρείται και ο πλέον σημαντικός, δηλαδή έχει τον υψηλότερο βαθμό.

- Οι υπόλοιποι ιστότοποι κατατάσσονται ανάλογα με το μέγεθος της αντίστοιχης συνιστώσας του ιδιοδιανύσματος. Για το ανωτέρω ιδιοδιάνυσμα έχουμε $\max_{1 \leq i \leq 7} z_i = 0.4746$, συνεπώς ο υπάλληλος Y_2 έχει ιστότοπο με το μεγαλύτερο βαθμό, ακολουθεί ο ιστότοπος του Y_5 και μετά αυτός του Y_1.
- Παρατηρήστε ότι αν και ο ιστότοπος του Y_1 έχει τις περισσότερες αναφορές από τον ιστότοπο του Y_2 ή του Y_5 εν τούτοις θεωρείται μικρότερου βαθμού.
- Η μέθοδος της κατάταξης ενός ιστότοπου βασίζεται στο βαθμό του ιστότοπου, ο οποίος είναι υψηλότερος όσο υπάρχουν ιστότοποι με μεγαλύτερο βαθμό που είναι συνδεδεμένοι με αυτόν.
- Η παραδοχή που έχουμε κάνει εδώ είναι ότι το ιδιοδιάνυσμα έχει θετικές συνιστώσες κάτι που συμβαίνει σε μεγάλες κατηγορίες προβλημάτων.

Η μέθοδος των δυνάμεων

- Η μέθοδος αυτή υπολογίζει τη μεγαλύτερη κατά μέτρο ιδιοτιμή ενός τετραγωνικού πίνακα $A \in \mathbb{C}^{n\times n}$ και το αντίστοιχο ιδιοδιάνυσμα.
- Έστω ότι ο πίνακας A έχει τις $\lambda_i,\ i = 1, 2, \ldots, n$ ιδιοτιμές και τα αντίστοιχα ιδιοδιανύσματα είναι τα $x^{(i)},\ i = 1, 2, \ldots, n$.
- Έστω ότι ο A έχει n γραμμικά ανεξάρτητα ιδιοδιανύσματα και συνεπώς αποτελούν μια βάση.
- Επιπλέον υποθέτουμε ότι

$$|\lambda_1| > |\lambda_j|, \qquad j = 2, 3, \ldots, n. \tag{3}$$

- Αν $y^{(0)} \neq 0$ ένα διάνυσμα του $\mathbb{C}^n$, τότε επειδή τα ιδιοδιανύσματα του A αποτελούν μία βάση, το $y^{(0)}$ μπορεί να γραφεί ως εξής

$$y^{(0)} = \sum_{i=1}^{n} \alpha_i x^{(i)} \tag{4}$$

όπου α_i είναι βαθμωτά μεγέθη και όχι όλα μηδέν.

...Η μέθοδος των δυνάμεων...

- Στη συνέχεια θεωρούμε την ακολουθία των διανυσμάτων που ορίζονται από το επαναληπτικό σχήμα

$$y^{(m+1)} = Ay^{(m)}, \quad m = 0, 1, 2, \ldots \tag{5}$$

  όπου το $y^{(0)}$ είναι ένα αυθαίρετο διάνυσμα.

- Από την (5) για $m = 0$ έχουμε

$$y^{(1)} = Ay^{(0)} = A\sum_{i=1}^{n} \alpha_i x^{(i)} = \sum_{i=1}^{n} \alpha_i A x^{(i)} = \sum_{i=1}^{n} \alpha_i \lambda_i x^{(i)}$$

- για $m = 1$ έχουμε

$$y^{(2)} = Ay^{(1)} = \sum_{i=1}^{n} \alpha_i \lambda_i^2 x^{(i)}.$$

...Η μέθοδος των δυνάμεων...

Γενικά

$$y^{(m)} = \sum_{i=1}^{n} \alpha_i \lambda_i^m x^{(i)}, \tag{6}$$

η οποία γράφεται

$$\begin{aligned} y^{(m)} &= \lambda_1^m \left[\alpha_1 x^{(1)} + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^m x^{(i)} \right] \\ &= \lambda_1^m [\alpha_1 x^{(1)} + \varepsilon^{(m)}], \end{aligned} \tag{7}$$

όπου

$$\varepsilon^{(m)} = \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^m x^{(i)}.$$

Προσδιορισμός της ιδιοτιμής λ_1

- Επειδή όμως $|\lambda_i/\lambda_1| < 1, \quad i = 2, 3, \ldots, n$, από την (7) προκύπτει ότι

$$\lim_{m\to\infty} y^{(m)} = \lim_{m\to\infty} \lambda_1^m \alpha_1 x^{(1)} \tag{8}$$

  υποθέτοντας βέβαια ότι $\alpha_1 \neq 0$.

- Αν όπου m θέσουμε $m-1$ στην (8) έχουμε

$$\lim_{m\to\infty} y^{(m-1)} = \lim_{m\to\infty} \lambda_1^{m-1} \alpha_1 x^{(1)}. \tag{9}$$

- Διαιρώντας τις αντίστοιχες συνιστώσες των $y^{(m)}$ και $y^{(m-1)}$ λαμβάνουμε από τις (8) και (9) ότι

$$\lim_{m\to\infty} \frac{y_j^{(m)}}{y_j^{(m-1)}} = \lambda_1 \quad j = 1, 2, \ldots, n. \tag{10}$$

Προσδιορισμός του ιδιοδιανύσματος $\mathbf{x}^{(1)}$

Έχοντας υπολογίσει την λ_1 από την (8) προκύπτει ότι

$$\lim_{m\to\infty} \frac{y^{(m)}}{\lambda_1^m} = \alpha_1 x^{(1)}, \qquad (11)$$

δηλαδή το ιδιοδιάνυσμα που αντιστοιχεί στην λ_1 τη μεγαλύτερη κατά μέτρο ιδιοτιμή.

Με άλλα λόγια δημιουργούμε την ακολουθία των διανυσμάτων $y^{(0)}, y^{(1)}, \ldots, y^{(m)}$ μέχρις ότου οι λόγοι των αντίστοιχων συνιστωσών δύο διαδοχικών διανυσμάτων τείνουν προς την ίδια σταθερή τιμή, η οποία είναι μία προσέγγιση της ιδιοτιμής λ_1 (βλ.(10)). Το διάνυσμα $y^{(m)}$ είναι μία μη κανονικοποιημένη προσέγγιση του αντίστοιχου ιδιοδιανύσματος.

Η ταχύτητα σύγκλισης της μεθόδου των δυνάμεων

- Εξαρτάται από τις σταθερές α_i και τους λόγους $|\lambda_2/\lambda_1|, |\lambda_3/\lambda_1|, \ldots, |\lambda_n/\lambda_1|$ (βλ.(7)).
  Έτσι όσο μικρότεροι είναι αυτοί οι λόγοι τόσο ταχύτερη είναι η σύγκλιση της μεθόδου.
- Ιδιαίτερα αν $|\lambda_2| / |\lambda_1|$ είναι κοντά στη μονάδα, τότε η σύγκλιση της μεθόδου είναι πιθανό να είναι πάρα πολύ αργή.
- Θεωρητικά, αν τύχει και η εκλογή του $y^{(0)}$ είναι τέτοια ώστε $\alpha_1 = 0$ και $|\lambda_2| > |\lambda_j|$, $j \geq 3$, τότε η μέθοδος θα συγκλίνει στην λ_2 και σε ένα πολλαπλάσιο του $x^{(2)}$.
- Στην πράξη όμως δεν έχουμε δυσκολίες αν $\alpha_1 = 0$, γιατί τα σφάλματα στρογγύλευσης δημιουργούν μία μικρή (που με την αύξηση του αριθμού των επαναλήψεων μεγαλώνει) τιμή του α_1 αρκετά ικανοποιητική, ώστε να έχουμε σύγκλιση τελικά στην λ_1.

Παρατήρηση

- Για $|\lambda_1| > 1$ τότε από την (8) έχουμε ότι $\lim_{m\to\infty} y_j^{(m)} = \pm\infty, j = 1, 2, \ldots, n,$ ενώ για $|\lambda_1| < 1, \quad \lim_{m\to\infty} y_j^{(m)} = 0.$
- Έτσι εκτός αν $|\lambda_1| \simeq 1$ θα πρέπει να εκτελούμε πράξεις με απόλυτα πάρα πολύ μεγάλους ή πάρα πολύ μικρούς αριθμούς, πράγμα που σημαίνει αύξηση των σφαλμάτων στρογγύλευσης στους υπολογισμούς.
- Το πρόβλημα αυτό αποφεύγεται με μία τροποποίηση της μεθόδου των δυνάμεων.

Τροποποίηση της μεθόδου των δυνάμεων

Συνίσταται από τα ακόλουθα τρία διαδοχικά βήματα σε κάθε επανάληψη όπου ουσιαστικά κανονικοποιείται το $y^{(m)}$

$$\begin{aligned} y_{j_m}^{(m)} &= \max_j \left| y_j^{(m)} \right| = \| y^{(m)} \|_\infty, \\ z^{(m)} &= \frac{1}{y_{j_m}^{(m)}} y^{(m)} \\ y^{(m+1)} &= A z^{(m)}, \quad m = 0, 1, 2, \ldots \end{aligned} \tag{12}$$

Εργαζόμενοι με ανάλογο τρόπο όπως και στην προηγούμενη παράγραφο, η αντίστοιχη έκφραση της $y^{(m)}$ θα δίνεται από τη σχέση

$$y^{(m)} = \frac{1}{y_{j_0}^{(0)} y_{j_1}^{(1)} \ldots y_{j_{m-1}}^{(m-1)}} \lambda_1^m \left[\alpha_1 x^{(1)} + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^m x^{(i)} \right]. \tag{13}$$

....Τροποποίηση της μεθόδου των δυνάμεων

Επίσης έχουμε ότι

$$
\begin{aligned}
z^{(m-1)} &= \frac{1}{y_{j_{m-1}}^{(m-1)}} y^{(m-1)} \\
&= \frac{1}{y_{j_{m-1}}^{(m-1)}} \left[\frac{1}{y_{j_0}^{(0)} y_{j_1}^{(1)} \ldots y_{j_{m-2}}^{(m-2)}} \lambda_1^{m-1} \left(\alpha_1 x^{(1)} + \sum_{i=2}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^{m-1} x^{(i)} \right) \right]. \quad (14)
\end{aligned}
$$

Από τις (13) και (14) έχουμε

$$
\lim_{m\to\infty} \frac{y_{j_{m-1}}^{(m)}}{z_{j_{m-1}}^{(m-1)}} = \lambda_1
$$

αλλά $z_{j_{m-1}}^{(m-1)} = 1$ (το $z^{(m-1)}$ είναι κανονικοποιημένο) και η ανωτέρω σχέση γράφεται

$$
\lim_{m\to\infty} y_{j_{m-1}}^{(m)} = \lambda_1. \quad (15)
$$

....Τροποποίηση της μεθόδου των δυνάμεων

- Η ακολουθία λοιπόν που δημιουργείται από τις συνιστώσες j_{m-1} του διανύσματος $y^{(m)}$ τείνει στη μεγαλύτερη κατά μέτρο ιδιοτιμή. Υπενθυμίζεται ότι η συνιστώσα j_{m-1} του διανύσματος $y^{(m)}$ είναι εκείνη που αντιστοιχεί στην μεγαλύτερη κατά μέτρο απόλυτη τιμή συνιστώσα του προηγούμενου διανύσματος $y^{(m-1)}$.
- Για την εύρεση του αντίστοιχου ιδιοδιανύσματος $x^{(1)}$ παρατηρούμε ότι, αν ο δείκτης j_m από μία ορισμένη τιμή του m και μετά παραμένει σταθερός, τότε είναι

$$\lim_{m-\infty} z^{(m)} = cx^{(1)} \tag{16}$$

όπου c σταθερά τέτοια ώστε η μεγαλύτερη κατά μέτρο συνιστώσα του $cx^{(1)}$ να είναι μονάδα.
- Άρα η ακολουθία των διανυσμάτων $z^{(m)}$ συγκλίνει προς το κανονικοποιημένο ιδιοδιάνυσμα που αντιστοιχεί στην λ_1.

Ταχύτητα Σύγκλισης της τροποποιημένης μεθόδου των δυνάμεων

- Η ταχύτητα σύγκλισης της ακολουθίας $\left\{ y_{j_{m-1}}^{(m)} \right\}_{m=1}^{\infty}$ προς την λ_1 προσδιορίζεται, όπως αναφέρθηκε, από τους λόγους $|\lambda_j/\lambda_1|^m$ για $j = 2, 3, \ldots, n$ και ιδιαίτερα από τον λόγο $|\lambda_2/\lambda_1|^m$. Με άλλα λόγια η τάξη σύγκλισης είναι $O((\lambda_2/\lambda_1)^m)$.
- Επομένως για μεγάλα m έχουμε

$$\left| y_{j_{m-1}}^{(m)} - \lambda_1 \right| \simeq k \left| \frac{\lambda_2}{\lambda_1} \right|^m,$$

όπου k σταθερά, πράγμα που σημαίνει ότι

$$\lim_{m\to\infty} \frac{\left| y_{j_m}^{(m+1)} - \lambda_1 \right|}{\left| y_{j_{m-1}}^{(m)} - \lambda_1 \right|} \simeq \left| \frac{\lambda_2}{\lambda_1} \right| < 1$$

ή

$$\varepsilon^{(m+1)} \simeq \left| \frac{\lambda_2}{\lambda_1} \right| \varepsilon^{(m)},$$

όπου $\varepsilon^{(m)} = \left| y_{j_{m-1}}^{(m)} - \lambda_1 \right|$, δηλαδή η ταχύτητα σύγκλισης είναι γραμμική.

Παράδειγμα

Δίνεται ο πίνακας $A = \begin{bmatrix} -2 & -2 & 3 \\ -10 & -1 & 6 \\ 10 & -2 & -9 \end{bmatrix}$.

με ιδιοτιμές $\lambda_1 = -12$, $\lambda_2 = -3$ και $\lambda_3 = 3$. Να εκτελεστούν τρεις επαναλήψεις της μεθόδου των δυνάμεων για τον υπολογισμό της μεγαλύτερης κατά μέτρο ιδιοτιμής και του αντιστοίχου ιδιοδιανύσματος του πίνακα A. Λάβετε σαν αρχικό διάνυσμα το $y^{(0)} = [1, 0, 0]^T$.

Λύση

Για $m = 0$ η (12) δίνει διαδοχικά

$$y_{j_0}^{(0)} = ||y^{(0)}||_\infty = \max\{|1|, |0|, |0|\} = 1, \text{ άρα } j_0 = 1$$

$$z^{(0)} = \frac{1}{y_1^{(0)}} y^{(0)} = \frac{1}{1}[1,\ 0,\ 0]^T$$

συνεπώς

$$y^{(1)} = Az^{(0)} = [-2,\ -10,\ 10]^T.$$

Η πρώτη προσέγγιση της μεγαλύτερης κατά μέτρο ιδιοτιμής λ είναι

$$\lambda_1 = y_{j_0}^{(1)} = y_1^{(1)} = -2.$$

...Λύση

Για $m = 1$ η (12) δίνει

$$y_{j_1}^{(1)} = ||y^{(1)}||_\infty = \max\{|-2|, |-10|, |10|\} = 10, \text{ άρα } j_1 = 2$$

και

$$z^{(1)} = \frac{1}{y_2^{(1)}} y^{(1)} = [1/5,\ 1,\ -1]^T.$$

Επιπλέον

$$y^{(2)} = Az^{(1)} = [-27/5,\ -9,\ 9]^T$$

και η δεύτερη προσέγγιση της λ είναι

$$\lambda_2 = y_2^{(2)} = -9$$

...Λύση

Για $m = 2$ η (12) δίνει

$$y_{j_2}^{(2)} = ||y^{(2)}||_\infty = \max\{|-27/5|, |-9|, |9|\} = 9, \text{ άρα } j_2 = 2$$

και

$$z^{(2)} = \frac{1}{y_2^{(2)}} y^{(2)} = [3/5,\ 1,\ -1]^T.$$

Οπότε

$$y^{(3)} = Az^{(2)} = [-31/5,\ -13,\ 13]^T$$

και η τρίτη προσέγγιση της λ είναι

$$\lambda_3 = y_2^{(3)} = -13.$$

Το αντίστοιχο ιδιοδιάνυσμα είναι το

$$z^{(3)} = \frac{1}{y_2^{(3)}} y^{(3)} = [31/65,\ 1,\ -1]^T.$$

Ο αλγόριθμος της μεθόδου των δυνάμεων

Σαν αρχικό διάνυσμα $y^{(0)}$ λαμβάνουμε συνήθως το $y^{(0)} = [1, 1, \dots, 1]^T$.

1 Διάβασε τη διάσταση n του πίνακα A, τα στοιχεία a_{ij}, $1 \leq i, j \leq n$, το αρχικό διάνυσμα y_i, $1 \leq i \leq n$, την ανεκτικότητα ε και το μέγιστο αριθμό επαναλήψεων M.

2 Να τεθεί

$$m = 0$$

$$\lambda_0 = 0$$

3 Να βρεθεί ένας ακέραιος p τέτοιος ώστε

$$|y_p| = \max_{1 \leq i \leq n} |y_i|$$

4 Για $i = 1, 2, \dots, n$ να υπολογιστεί

$$z_i = \frac{1}{y_p} y_i$$

....Ο αλγόριθμος της μεθόδου των δυνάμεων

5 Όσο ισχύει $m \leq M$ να εκτελούνται τα βήματα 5.1-5.7

5.1 Για $i = 1, 2, \dots, n$ να τεθεί

$$y_i = \sum_{j=1}^{n} a_{ij} z_j$$

5.2 Να τεθεί

$$\lambda_1 = y_p$$

5.3 Αν $y_p = 0$ τότε τύπωσε ``Ο A έχει ιδιοτιμή 0, επίλεξε νέο αρχικό διάνυσμα και άρχισε πάλι τη διαδικασία''. Τέλος.

5.4 Να βρεθεί ένας ακέραιος p τέτοιος ώστε

$$|y_p| = \max_{1 \leq i \leq n} |y_i|$$

5.5 Για $i = 1, 2, \dots, n$ να υπολογισθεί

$$z_i = \frac{1}{y_p} y_i$$

....Ο αλγόριθμος της μεθόδου των δυνάμεων

5.6 Αν

$$|\lambda_0 - \lambda_1| < \varepsilon$$

τότε τύπωσε (λ_1, z). Τέλος.

5.7 Να τεθεί

$$m = m + 1$$

$$\lambda_0 = \lambda_1$$

6 Τύπωσε ``Όχι σύγκλιση μετά από M επαναλήψεις''. Τέλος.

Υπολογισμός της μικρότερης κατά μέτρο ιδιοτιμής

Αν

$$|\lambda_n| < |\lambda_{n-1}| \leq \cdots \leq |\lambda_1|$$

τότε ο υπολογισμός της μικρότερης κατά μέτρο ιδιοτιμής λ_n γίνεται ως εξής :

Επειδή $\mathbf{Ax} = \lambda\mathbf{x}$ και $\mathbf{A}^{-1}\mathbf{x} = \frac{\mathbf{1}}{\lambda}\mathbf{x}$ και $\frac{\mathbf{1}}{|\lambda_\mathbf{n}|} > \frac{1}{|\lambda_i|}$
εφαρμόζεται η μέθοδος των δυνάμεων

$$\mathbf{y}^{(\mathbf{m+1})} = \mathbf{A}^{-1}\mathbf{y}^{(\mathbf{m})} , \qquad \mathbf{m} = \mathbf{0}, \mathbf{1}, \mathbf{2}, \cdots$$

ή

$$\mathbf{Ay}^{(\mathbf{m+1})} = \mathbf{y}^{(\mathbf{m})} , \qquad \mathbf{m} = \mathbf{0}, \mathbf{1}, \mathbf{2}, \cdots$$

δηλ. η επίλυση των γραμμικών συστημάτων

$$\mathbf{Ay}^{(\mathbf{1})} = \mathbf{y}^{(\mathbf{0})} , \qquad \mathbf{Ay}^{(\mathbf{2})} = \mathbf{y}^{(\mathbf{1})} , \qquad \mathbf{Ay}^{(\mathbf{3})} = \mathbf{y}^{(\mathbf{2})} , \cdots$$

Τεχνικές επιτάχυνσης της μεθόδου των δυνάμεων

Επιτάχυνση της σύγκλισης - Η μέθοδος του Aitken

Αν η σύγκλιση μιας επαναληπτικής μεθόδου είναι γραμμική, τότε μπορεί να επιταχυνθεί με τη χρήση της μεθόδου του Aitken, η οποία έχει τον ακόλουθο τύπο

$$x_n' = x_n - \frac{(\Delta x_n)^2}{\Delta^2 x_n} \tag{17}$$

όπου Δ είναι ο τελεστής των προς τα εμπρός διαφορών και ορίζεται σαν $\Delta x_n = x_{n+1} - x_n$ με $\Delta^2 x_n = \Delta x_{n+1} - \Delta x_n = x_{n+2} - 2x_{n+1} + x_n$.

Επομένως είναι δυνατόν να τροποποιηθεί ο αλγόριθμος της μεθόδου των δυνάμεων έτσι ώστε να παράγεται η ακολουθία

$$\tilde{\lambda}_1 = \lambda_1 - \frac{(\Delta \lambda_1)^2}{\Delta^2 \lambda_1}. \tag{18}$$

Η μέθοδος των πηλίκων του Rayleigh

Αν ο πίνακας A είναι πραγματικός και συμμετρικός, τότε είναι δυνατόν να επιταχυνθεί η σύγκλιση προς τη μεγαλύτερη κατά απόλυτη τιμή ιδιοτιμή με τη χρήση της μεθόδου των πηλίκων του Rayleigh.

Ορισμός. Για κάθε διάνυσμα $x \neq 0$ η ποσότητα

$$\frac{(x, Ax)}{(x, x)}$$

καλείται πηλίκο του Rayleigh που αντιστοιχεί στο x.

Ένα βασικό αποτέλεσμα που δείχνει τη σπουδαιότητα των πηλίκων του Rayleigh είναι το ακόλουθο

Θεώρημα

Αν ο $A \in \mathbb{R}^{n \times n}$ είναι συμμετρικός και $x \neq 0$ είναι ένα αυθαίρετο διάνυσμα, τότε

$$\lambda_1 = \max_{x \neq 0} \frac{(x, Ax)}{(x, x)} = \frac{(x^{(1)}, Ax^{(1)})}{(x^{(1)}, x^{(1)})}$$

και

$$\lambda_n = \min_{x \neq 0} \frac{(x, Ax)}{(x, x)} = \frac{(x^{(n)}, Ax^{(n)})}{(x^{(n)}, x^{(n)})} \tag{19}$$

όπου λ_1, λ_n είναι η μεγαλύτερη και η μικρότερη ιδιοτιμή, αντίστοιχα και $x^{(1)}, x^{(n)}$ ιδιοδιανύσματα του A που αντιστοιχούν στις λ_1 και λ_n.

- Παρατηρούμε ότι ο υπολογισμός της λ_1 είναι ένα πρόβλημα βελτιστοποίησης. Το ενδιαφέρον μας είναι η χρήση των πηλίκων του Rayleigh για την επιτάχυνση της σύγκλισης της μεθόδου των δυνάμεων.
- Έστω το βασικό επαναληπτικό σχήμα της μεθόδου των δυνάμεων

$$y^{(m+1)} = Ay^{(m)}$$

τότε χρησιμοποιώντας την (6), προκύπτει

$$(y^{(m)}, y^{(m+1)}) = (y^{(m)}, Ay^{(m)}) = \sum_{i=1}^{n} \alpha_i^2 \lambda_i^{2m+1} \tag{20}$$

καθόσον τα ιδιοδιανύσματα του A είναι ορθογώνια, αφού ο A είναι συμμετρικός, δηλαδή

$$(x^{(i)}, x^{(j)}) = \delta_{ij} = \begin{cases} 1, & \text{αν } i = j \\ 0, & \text{αν } i \neq j \end{cases}$$

- Επίσης

$$(y^{(m)}, y^{(m)}) = \sum_{i=1}^{n} \alpha_i^2 \lambda_i^{2m}. \tag{21}$$

- Από τις (20) και (21) έχουμε

$$\frac{(y^{(m)}, Ay^{(m)})}{(y^{(m)}, y^{(m)})} = \lambda_1 + O((\lambda_i/\lambda_1)^{2m}) \tag{22}$$

η οποία συγκρινόμενη με την (7) δείχνει ότι το πηλίκο του Rayleigh που αντιστοιχεί στο $y^{(m)}$ γενικά θα συγκλίνει ταχύτερα $\left(O(\lambda_i/\lambda_1)^{2m}\right)$ από τη μέθοδο των δυνάμεων $\left(O(\lambda_i/\lambda_1)^{m}\right)$.

- Για την κανονικοποίηση του ιδιοδιανύσματος χρησιμοποιείται η ευκλείδεια norm, έτσι

$$z^{(m)} = \frac{1}{||y^{(m)}||_2} y^{(m)} = \frac{1}{(y^{(m)}, y^{(m)})^{1/2}} y^{(m)}. \tag{23}$$

- Επίσης, στην τροποποιημένη μέθοδο των δυνάμεων έχουμε

$$y^{(m+1)} = Az^{(m)} \tag{24}$$

οπότε το πρώτο μέλος της (22), λόγω των (23) και (24), γράφεται διαδοχικά

$$\begin{aligned} \frac{(y^{(m)}, Ay^{(m)})}{(y^{(m)}, y^{(m)})} &= \left(\frac{y^{(m)}}{(y^{(m)}, y^{(m)})^{1/2}}, \frac{Ay^{(m)}}{(y^{(m)}, y^{(m)})^{1/2}} \right) \\ &= (z^{(m)}, Az^{(m)}) = (z^{(m)}, y^{(m+1)}). \end{aligned} \tag{25}$$

Λόγω των (22) και (25) έχουμε τελικά ότι

$$\lim_{m\to\infty}(z^{(m)}, y^{(m+1)}) = \lambda_1. \tag{26}$$

Παράδειγμα

Δίνεται ο πίνακας

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

με ιδιοτιμές $\lambda_1 = 4, \quad \lambda_2 = 3, \quad \lambda_3 = 1$. Να εκτελεστούν δύο επαναλήψεις της μεθόδου των πηλίκων Rayleigh για τον υπολογισμό της μεγαλύτερης κατά μέτρο ιδιοτιμής του πίνακα A. Λάβετε ως αρχικό διάνυσμα το $[1, 1, 1]^T$.

Λύση

1η επανάληψη $y^{(0)} = [1, 1, 1]^T$ και έστω $\lambda_0 = 0$. Λόγω της (23) έχουμε

$$||y^{(0)}||_2 = \sqrt{(y^{(0)}, y^{(0)})} = \sqrt{1^2 + 1^2 + 1^2} = \sqrt{3}$$

και

$$z^{(0)} = \frac{1}{||y^{(0)}||_2} y^{(0)} = \left[\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}\right]^T.$$

Επίσης, από την (24)

$$y^{(1)} = Az^{(0)} = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{3}} \\ 0 \\ \frac{2}{\sqrt{3}} \end{bmatrix}$$

οπότε η (26) δίνει

$$\lambda_1 = (z^{(0)}, y^{(1)}) = \frac{1}{\sqrt{3}} \cdot \frac{2}{\sqrt{3}} + \frac{1}{\sqrt{3}} \cdot 0 + \frac{1}{\sqrt{3}} \cdot \frac{2}{\sqrt{3}}$$
$$= \frac{2}{3} + 0 + \frac{2}{3} = \frac{4}{3}.$$

Παρατηρούμε ότι

$$|\lambda_0 - \lambda_1| = |0 - \frac{4}{3}| = 1.33.$$

2η επανάληψη

$$||y^{(1)}||_2 = \sqrt{\frac{4}{3} + 0 + \frac{4}{3}} = \sqrt{\frac{8}{3}}$$

$$z^{(1)} = \frac{1}{||y^{(1)}||_2} y^{(1)} = \sqrt{\frac{3}{8}} \left[\frac{2}{\sqrt{3}}, 0, \frac{2}{\sqrt{3}} \right]^T = \left[\frac{1}{\sqrt{2}}, 0, \frac{1}{\sqrt{2}} \right]^T$$

$$y^{(2)} = Az^{(1)} = \begin{bmatrix} 3 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{2}} \\ 0 \\ \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \frac{3}{\sqrt{2}} \\ -\sqrt{2} \\ \frac{3}{\sqrt{2}} \end{bmatrix}$$

και

$$\lambda_2 = (z^{(1)}, y^{(2)}) = \frac{1}{\sqrt{2}} \cdot \frac{3}{\sqrt{2}} + 0 \cdot (-\sqrt{2}) + \frac{1}{\sqrt{2}} \cdot \frac{3}{\sqrt{2}} = \frac{3}{2} = 1.5.$$

Τέλος, παρατηρούμε ότι

$$|\lambda_2 - \lambda_1| = |1.5 - 1.33| = 0.17.$$

Ο αλγόριθμος της μεθόδου των πηλίκων του Rayleigh

1 Διάβασε την τάξη n του πίνακα A, τα στοιχεία a_{ij}, $1 \leq i, j \leq n$, το αρχικό διάνυσμα y_i, $1 \leq i \leq n$, την ανεκτικότητα ε και το μέγιστο αριθμό επαναλήψεων M.

2 Να τεθεί

$$k = 0$$

$$\lambda_0 = 0$$

3 Για $i = 1, 2, \ldots, n$ να υπολογιστεί η ποσότης (υλοποίηση της (23))

$$z_i = y_i / \parallel y \parallel_2$$

4 Όσο ισχύει $k \leq M$ να εκτελούνται τα βήματα 4.1 - 4.6

4.1 Για $i = 1, 2, \ldots, n$ να υπολογισθεί ($y = Az$)

$$y_i = \sum_{j=1}^{n} a_{ij} z_j$$

4.2 Για $i = 1, 2, \ldots, n$ να υπολογισθεί $\left(\lambda = (z, y) = \frac{(y,Az)}{(y,y)}\right)$

$$\lambda = \sum_{i=1}^{n} z_i y_i$$

4.3 Αν $\| \; y \; \|_2 = 0$ τότε τύπωσε ``Ο A έχει ιδιοτιμή 0, επίλεξε νέο αρχικό διάνυσμα και άρχισε πάλι τη διαδικασία''. Τέλος.

4.4 Για $i = 1, 2, \ldots, n$ να υπολογισθεί η ποσότητα

$$z_i = y_i / \| \; y \; \|_2$$

4.5 Αν

$$|\lambda - \lambda_0| < \varepsilon$$

τότε τύπωσε (λ, z). Τέλος.

4.6 Να τεθεί

$$k = k + 1$$
$$\lambda_0 = \lambda$$

5 Τύπωσε ``Όχι σύγκλιση μετά από M επαναλήψεις''. Τέλος.

Μετατόπιση της αρχής των αξόνων

Πρόταση

Οι πίνακες $\mathbf{A}$ και $\mathbf{A} - \mathbf{qI}$ έχουν τα ίδια ιδιοδιανύσματα και αν λ_i είναι ιδιοτιμή του $\mathbf{A}$ τότε η αντίστοιχη ιδιοτιμή του $\mathbf{A} - \mathbf{qI}$ είναι η $\lambda_i - \mathbf{q}$.

Απόδειξη

Αν $\mathbf{Ax}^{(i)} = \lambda_i \mathbf{x}^{(i)}$ τότε

$$(\mathbf{A} - \mathbf{qI})\mathbf{x}^{(i)} = \mathbf{Ax}^{(i)} - \mathbf{qIx}^{(i)} = (\lambda_i - \mathbf{q})\mathbf{x}^{(i)}$$

Αφαιρώντας λοιπόν την ποσότητα $\mathbf{q}$ από τα διαγώνια στοιχεία του $\mathbf{A}$ έχει σαν αποτέλεσμα την αφαίρεση της $\mathbf{q}$ από τις ιδιοτιμές.

....Μετατόπιση της αρχής των αξόνων

Υποθέτουμε ότι ο $\mathbf{A} \in \mathbb{R}^{\mathbf{n},\mathbf{n}}$ έχει **n** γραμμικά ανεξάρτητα ιδιοδιανύσματα και όλες οι ιδιοτιμές του είναι πραγματικές και ικανοποιούν τη σχέση

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3 \geq \ldots \geq |\lambda_{n-1}| \geq |\lambda_n| \tag{27}$$

Αν αφαιρέσουμε την ποσότητα $\mathbf{q} \in \mathbf{R}$ με $\mathbf{q} \in (|\lambda_n|, |\lambda_1|)$ από τα διαγώνια στοιχεία του **A**, τότε ανεξάρτητα από την τιμή της **q**, η μεγαλύτερη κατά μέτρο ιδιοτιμή του $\mathbf{A} - \mathbf{qI}$ θα είναι πάντα η $\lambda_1 - \mathbf{q}$ ή η $\lambda_n - \mathbf{q}$.

Ας υποθέσουμε ότι θέλουμε να προσδιορίσουμε την λ_1. Οι ιδιοτιμές του $\mathbf{A} - \mathbf{qI}$ είναι οι $\mu_i = \lambda_i - \mathbf{q}$.

Μετατόπιση της αρχής των αξόνων

Ταχύτητα σύγκλισης της μεθόδου των δυνάμεων

Με τη χρήση του πίνακα $\mathbf{A} - \mathbf{qI}$ αντί του $\mathbf{A}$, εξαρτάται από την ποσότητα

$$\max_{i \neq 1} \left| \frac{\lambda_i - \mathbf{q}}{\lambda_1 - \mathbf{q}} \right| \tag{28}$$

- Όσο μικρότερη είναι η ανωτέρω ποσότητα, τόσο ταχύτερη η σύγκλιση της μεθόδου. Αρκεί δηλαδή να εκλέξουμε το q τέτοιο ώστε να ελαχιστοποιείται η ποσότητα

$$\max_{i \neq 1} |\lambda_i - \mathbf{q}| \tag{29}$$

- Αποδεικνύεταιότι η ανωτέρω ποσότητα γίνεται ελάχιστη αν

$$\mathbf{q} = 1/2(\lambda_2 + \lambda_n).$$

- Όμοια εργαζόμενοι βρίσκουμε ότι η μέγιστη ταχύτητα σύγκλισης στην $\lambda_n - \mathbf{q}$ επιταχύνεται αν επιλέξουμε

$$\mathbf{q} = 1/2(\lambda_1 + \lambda_{n-1})$$

Παρατήρηση

Με τη μέθοδο αυτή μπορούμε να υπολογίσουμε τόσο την λ_1 όσο και την λ_n, ωστόσο όμως χρειαζόμαστε κάποιες εκτιμήσεις των ιδιοτιμών λ_2 και λ_n (ή των λ_1 και λ_{n-1}) πράγμα που απαιτεί επιπλέον υπολογισμούς στην πράξη και είναι ένα μειονέκτημα αυτής της μεθόδου.

Η αντίστροφη μέθοδος των δυνάμεων

Έχει το πλεονέκτημα να υπολογίζει μια οποιαδήποτε ιδιοτιμή και το αντίστοιχο ιδιοδιάνυσμα και να έχει γρήγορη ταχύτητα σύγκλισης.

Λήμμα

Οι πίνακες $\mathbf{A}$ και $\mathbf{A}^{-1}$ έχουν τα ίδια ιδιοδιανύσματα και αν λ_i είναι μια ιδιοτιμή του $\mathbf{A}$ τότε η αντίστοιχη ιδιοτιμή του $\mathbf{A}^{-1}$ είναι η $1/\lambda_i$.

Απόδειξη

Αν $\mathbf{A}\mathbf{x}^{(i)} = \lambda_i \mathbf{x}^{(i)}$ τότε πολ/ζοντας από αριστερά με $\mathbf{A}^{-1}$ έχουμε

$$\frac{1}{\lambda_i}\mathbf{x}^{(i)} = \mathbf{A}^{-1}\mathbf{x}^{(i)}.$$

Η αντίστροφη μέθοδος των δυνάμεων

Ας υποθέσουμε ότι ο $\mathbf{A} \in \mathbb{R}^{nn}$, έχει $\mathbf{n}$ γραμμικά ανεξάρτητα ιδιοδιανύσματα και όλες οι ιδιοτιμές του είναι πραγματικές. Επίσης αν γνωρίζουμε κάποια ποσότητα $\mathbf{q} \in \mathbf{R}$ η οποία βρίσκεται πλησιέστερα στην απλή ιδιοτιμή λ_k του $\mathbf{A}$ από οποιαδήποτε άλλη ιδιοτιμή του, τότε θα ισχύει

$$|\lambda_k - \mathbf{q}| < |\lambda_i - \mathbf{q}|, \qquad i = 1(1)n, \quad i \neq k \tag{30}$$

Δηλαδή η ιδιοτιμή $\lambda_k - \mathbf{q}$ είναι η μικρότερη κατά απόλυτο τιμή ιδιοτιμή του πίνακα $\mathbf{A} - \mathbf{q}\mathbf{I}$. Συνεπώς, αν αντί του $\mathbf{A}$ χρησιμοποιήσουμε τον πίνακα $(\mathbf{A} - \mathbf{q}\mathbf{I})^{-1}$ στο βασικό επαναληπτικό σχήμα της μεθόδου των δυνάμεων, τότε είναι δυνατόν να υπολογισθεί η ποσότητα $\dfrac{1}{\lambda_k - \mathbf{q}}$ και από αυτήν η λ_k.

Ο επαναληπτικός τύπος της αντίστροφης μεθόδου των δυνάμεων

Πράγματι, αν εφαρμοστεί η ε.μ.

$$(\mathbf{A} - \mathbf{q}\mathbf{I})\mathbf{y}^{(m+1)} = \mathbf{y}^{(m)}, \qquad m = 0, 1, 2, \ldots \tag{31}$$

όπου $\mathbf{y}^{(0)} \neq \mathbf{0}$ αυθαίρετο διάνυσμα είναι δυνατόν να υπολογισθεί η απόλυτα μεγαλύτερη ιδιοτιμή του $(\mathbf{A} - \mathbf{q}\mathbf{I})^{-1}$ δηλαδή η $1/(\lambda_k - \mathbf{q})$ και το αντίστοιχο ιδιοδιάνυσμα.

Ταχύτητα σύγκλισης

Εξαρτάται από την ποσότητα

$$\max_{i \neq k} \left| \frac{\lambda_k - q}{\lambda_i - q} \right| \tag{32}$$

αφού

$$y^{(m)} = (A - qI)^{-1} y^{(m-1)} = (A - qI)^{-m} y^{(0)}$$

$$= \frac{\alpha_1}{(\lambda_1 - q)^m} x^{(1)} + \frac{\alpha_2}{(\lambda_2 - q)^m} x^{(2)} + \ldots + \frac{\alpha_n}{(\lambda_n - q)^m} x^{(n)}$$

$$= \frac{1}{(\lambda_k - q)^m} \left[\alpha_1 \left(\frac{\lambda_k - q}{\lambda_1 - q} \right)^m x^{(1)} + \ldots + \alpha_k x^{(k)} \right.$$

$$\left. + \ldots + \alpha_n \left(\frac{\lambda_k - q}{\lambda_n - q} \right)^m x^{(n)} \right]$$

Παρατηρήσεις

- Η επιλογή της $\mathbf{q}$ καθορίζει και την ταχύτητα σύγκλισης της μεθόδου.
- Όσο πλησιέστερα η $\mathbf{q}$ είναι στην ιδιοτιμή λ_k τόσο ταχύτερη θα είναι και η σύγκλιση της μεθόδου. Επειδή η $\mathbf{q}$ μπορεί να εκλεγεί αυθαίρετα, μπορούμε να βρούμε μια προσέγγιση σε οποιαδήποτε ιδιοτιμή του $\mathbf{A}$. Ο προσδιορισμός των $\mathbf{y}^{(m)}$ γίνεται από την επίλυση των συστημάτων
$$(\mathbf{A} - \mathbf{q}\mathbf{I})\mathbf{y}^{(m)} = \mathbf{y}^{(m-1)}, \qquad \mathbf{m} = \mathbf{1}, \mathbf{2}, \ldots \tag{33}$$
- Στην πράξη τα διανύσματα κανονικοποιούνται, με άλλα λόγια, εφαρμόζεται η παραλλαγή της μεθόδου των δυνάμεων. Τα γραμμικά συστήματα που προκύπτουν έχουν τον ίδιο πίνακα και διαφορετικά δεύτερα μέλη.
- Χρήση μιας άμεσης μεθόδου για την επίλυση τους.
- Σχηματισμός της LU διάσπασης του $\mathbf{A} - \mathbf{q}\mathbf{I}$ μόνο μία φορά. Αν λοιπόν χρησιμοποιήσουμε κανονικοποιημένα διανύσματα και την LU μέθοδο τότε το βασικό επαναληπτικό σχήμα θα είναι το παρακάτω:
$$\mathbf{L}\mathbf{z} = \mathbf{z}^{(m)}$$
$$\mathbf{U}\mathbf{y}^{(m+1)} = \mathbf{z} \tag{34}$$
όπου
$$\mathbf{L}\mathbf{U} = \mathbf{A} - \mathbf{q}\mathbf{I}$$

Σημειώματα

Σημείωμα Αναφοράς

Copyright Εθνικόν και Καποδιστριακόν Πανεπιστήμιον Αθηνών 2015, Νικόλαος Μισυρλής, "Αριθμητική Ανάλυση. Ενότητα 4 - Αριθμητικός Υπολογισμός Ιδιοτιμών και Ιδιοδιανυσμάτων" Έκδοση:1.01 . Αθήνα 2015. Διαθέσιμο από τη δικτυακή διεύθυνση:`http://opencourses.uoa.gr/courses/DI12/` .

Σημείωμα Αδειοδότησης

Διατήρηση Σημειωμάτων

Οποιαδήποτε αναπαραγωγή ή διασκευή του υλικού θα πρέπει να συμπεριλαμβάνει:

- το Σημείωμα Αναφοράς
- το Σημείωμα Αδειοδότησης
- τη δήλωση Διατήρησης Σημειωμάτων
- το Σημείωμα Χρήσης Έργων Τρίτων (εφόσον υπάρχει)

μαζί με τους συνοδευόμενους υπερσυνδέσμους.

Σημείωμα Χρήσης Έργων τρίτων

Το Έργο αυτό κάνει χρήση του ακόλουθου έργου:
" Εισαγωγή στην Αριθμητική Ανάλυση : Μια αλγοριθμική προσέγγιση, αυτο-έκδοση, Αθήνα, 2009", Νικόλαος Μισυρλής.