

Μέθοδοι υπολογισμού ιδιοτιμών, ιδιοδιανυσμάτων

MEM 106

Γραμμική Άλγεβρα Ι

Θεόδωρος Κατσαούνης

Τμήμα Μαθηματικών και Εφαρμοσμένων Μαθηματικών
Πανεπιστήμιο Κρήτης

Υπολογιστικές Μέθοδοι

- Ο αναλυτικός υπολογισμός των ιδιοτιμών, ιδιοδιανυσμάτων ενός πίνακα μπορεί να είναι πρακτικά αδύνατος ακόμα και για 4×4 πίνακες. Υπενθυμίζουμε ότι οι ιδιοτιμές είναι οι ρίζες του χαρακτηριστικού πολυωνύμου:

$$\lambda^n + a_{n-1}\lambda^{n-1} + \dots + a_0 = 0$$

- Το πρόβλημα αυτό εμφανίζεται σε πολλές εφαρμογές μεγάλης πρακτικής σημασίας, μερικές από τις οποίες θα δούμε σε επόμενη διάλεξη.
- Υπάρχουν αρκετές μέθοδοι στη βιβλιογραφία για τον αριθμητικό υπολογισμό - προσέγγιση των ιδιοτιμών και ιδιοδιανυσμάτων. Θα παρουσιάσουμε τις δύο που χρησιμοποιούνται πιο συχνά στη πράξη :

- Μέθοδος των δυνάμεων
- Η μέθοδος QR

Μέθοδος των δυνάμεων

Μέθοδος των δυνάμεων : Ιδέα

- Παρουσιάστηκε με την σημερινή της μορφή το 1929 από τον Αυστριακό μαθηματικό R. Mises
- Μέθοδος των δυνάμεων : ουσιαστικά μια οικογένεια μεθόδων
- **Βασική ιδέα** : ο συνεχής πολλαπλασιασμός με ένα πίνακα A η κυρίαρχη(μεγαλύτερη) ιδιοτιμή του πίνακα μεγενθύνεται εκθετικά

$$Ax = \lambda x \Rightarrow A^2x = \lambda^2x \Rightarrow \dots \Rightarrow A^kx = \lambda^kx$$

- Η μέθοδος παράγει διαδοχικές προσεγγίσεις του κυρίαρχου ιδιοδιανύσματος. Κατόπιν υπολογίζουμε την αντίστοιχη κυρίαρχη ιδιοτιμή.
- **Βασικό μειονέκτημα** : πρακτικά μπορεί να μας δώσει διανύσματα με πολύ μικρές ή πολύ μεγάλες τιμές που να μην μπορούν να αναπαρασταθούν στο υπολογιστή.
- **Εύκολη διόρθωση** : κανονικοποίηση του διανύσματος σε κάθε βήμα της επανάληψης.

Μέθοδος των δυνάμεων : Αλγόριθμος

Υποθέτουμε ότι ο πίνακας $A \in \mathbb{R}^{n,n}$ διαγωνοποιείται με κυρίαρχη ιδιοτιμή λ και x το αντίστοιχο κυρίαρχο ιδιοδιάνυσμα.

Μέθοδος των δυνάμεων

Διαλέγουμε τυχαίο $x_0 \in \mathbb{R}^n$

Για $k = 1, 2, \dots$

$$x_k = Ax_{k-1} - \text{πολ/σμός πίνακα-διάνυσμα}$$

$$x_k = \frac{x_k}{\|x_k\|_\infty} - \text{κανονικοποίηση}$$

Ο αλγόριθμος ξεκινά με την επιλογή ενός τυχαίου διανύσματος στον $x_0 \in \mathbb{R}^n$. Με αυτό το τρόπο αυξάνονται οι πιθανότητες κάποια συνιστώσα του x_0 να είναι προς την διεύθυνση του κυρίαρχου ιδιοδιανύσματος x . Ακολουθεί ο πολ/σμός πίνακα με διάνυσμα και τέλος γίνεται η κανονικοποίηση του διανύσματος. Στο τερματισμό το x_k είναι μια προσέγγιση του x .

Μέθοδος των δυνάμεων : Παράδειγμα 1

Έστω $A = \begin{pmatrix} -7 & 2 \\ 8 & -1 \end{pmatrix}$, με κυρίαρχη ιδιοτιμή $\lambda = -9$ και αντίστοιχο κυρίαρχο ιδιοδιάνυσμα ένα πολλαπλάσιο του $x = (-1, 1)$. Ας εφαρμόσουμε την μέθοδο με αρχικό διάνυσμα $x_0 = (1, 0)$

$$x_1 = Ax_0 = \begin{pmatrix} -7 & 2 \\ 8 & -1 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \end{pmatrix} = \begin{pmatrix} -7 \\ 8 \end{pmatrix}$$

$$x_2 = Ax_1 = \begin{pmatrix} -7 & 2 \\ 8 & -1 \end{pmatrix} \begin{pmatrix} -7 \\ 8 \end{pmatrix} = \begin{pmatrix} 65 \\ -64 \end{pmatrix}$$

$$x_3 = Ax_2 = \begin{pmatrix} -7 & 2 \\ 8 & -1 \end{pmatrix} \begin{pmatrix} 65 \\ -64 \end{pmatrix} = \begin{pmatrix} -583 \\ 584 \end{pmatrix}$$

Κανονικοποίηση με $\|x_3\|_\infty = 584$

$$x_3 = \frac{1}{584}(-583, 584) = (-0.99828, 1) \approx (-1, 1)$$

Μέθοδος των δυνάμεων : Παράδειγμα 2

Έστω $A = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix}$, με κυρίαρχη ιδιοτιμή $\lambda = -2$ και αντίστοιχο κυρίαρχο ιδιοδιάνυσμα ένα πολλαπλάσιο του $x = (3, 1)$. Ας εφαρμόσουμε την μέθοδο με αρχικό διάνυσμα $x_0 = (1, 1)$

$$x_1 = Ax_0 = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} 1 \\ 1 \end{pmatrix} = \begin{pmatrix} -10 \\ -4 \end{pmatrix}$$

$$x_2 = Ax_1 = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} -10 \\ -4 \end{pmatrix} = \begin{pmatrix} 28 \\ 10 \end{pmatrix}$$

$$x_3 = Ax_2 = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} 28 \\ 10 \end{pmatrix} = \begin{pmatrix} -64 \\ -22 \end{pmatrix}$$

$$x_4 = Ax_3 = \begin{pmatrix} 2 & -12 \\ 1 & -5 \end{pmatrix} \begin{pmatrix} -64 \\ -22 \end{pmatrix} = \begin{pmatrix} 136 \\ 46 \end{pmatrix}$$

Κανονικοποίηση με $\|x_4\|_\infty = 136$

$$x_4 = \frac{1}{136}(136, 46) = (1, 0.33823) \approx \frac{1}{3}(3, 1) = (1, 0.333333)$$

Μέθοδος των δυνάμεων : κώδικας σε Python

Υλοποίηση της μεθόδου με Python . Ο κώδικας χρησιμοποιεί τον πίνακα του παραδείγματος 2.

```
import numpy as np
import numpy.linalg as LA
n=2; k=0;
xk=np.random.rand(n)
A=np.array([[2, -12], [1, -5]])
dxk=1.; eps=1.e-11
print(k, "   ",xk, "   ", dxk)
while (dxk > eps) :
    z=xk
    xk = np.matmul(A, xk)
    xk = xk / LA.norm(xk,np.inf)
    dxk = LA.norm(z+xk)
    k+=1
print(k, "   ",xk, "   ", dxk)
```

Ο αλγόριθμος κατασκευάζει μια ακολουθία διανυσμάτων $\{x_k\}$ προσεγγίσεις του κυρίαρχου ιδιοδιανύσματος x και τερματίζει όταν $dx_k = \|x_{k+1} - x_k\|_2 < \epsilon$, ϵ ένας μικρός αριθμός. Κανονικοποίηση : διαιρούμε όλες τις συνιστώσες του x_k με $\|x_k\|_\infty$ σε κάθε βήμα του αλγορίθμου.

Μέθοδος των δυνάμεων : αποτελέσματα αλγορίθμου

Αποτελέσματα εκτέλεσης του παραπάνω κώδικα για το παράδειγμα 2.

k	x_k	dx_k
1	(-1 -0.35082011)	0.6591786104319128
2	(1 0.34124647)	0.00957364047847803
3	(-1 -0.33711056)	0.00413590705408573
.		
.		
.		
30	(1 0.33333333)	2.6922297724496502e-11
31	(-1 -0.33333333)	1.3461065595521404e-11
32	(1 0.33333333)	6.7305050421850865e-12

Στο υπολογιστή σας η ακολουθία θα είναι λίγο διαφορετική λόγω της τυχαιότητας στο αρχικό διάνυσμα. Βλέπουμε ότι

$$x_{32} = (1, 0.3333333) = \frac{1}{3}x. \text{ Γενικά } \frac{x_k}{\|x_k\|} \rightarrow \frac{x}{\|x\|}$$

Μέθοδος των δυνάμεων : Ιδιοτιμή;

- Θεωρία : πρώτα υπολογίζουμε την ιδιοτιμή και μετά το ιδιοδιάνυσμα.
- Μέθοδος : δουλεύει ανάποδα, πρώτα το ιδιοδιάνυσμα.
- Ερώτημα : Πως θα βρούμε τη αντίστοιχη ιδιοτιμή;
- Πηλίκο Rayleigh

$$Ax = \lambda x \Rightarrow x^T Ax = \lambda x^T x \Rightarrow \lambda = \frac{x^T Ax}{x^T x}$$

- Επειδή $x_k \rightarrow x$, περιμένουμε ότι ο λόγος $\frac{x_k^T Ax_k}{x_k^T x_k}$ θα είναι μια καλή προσέγγιση της ιδιοτιμής. Για τα παραδείγματα έχουμε
 - ▶ Παράδειγμα 1: $\frac{x_3^T Ax_3}{x_3^T x_3} = -8.994851272863448 \approx -9$
 - ▶ Παράδειγμα 2: $\frac{x_4^T Ax_4}{x_4^T x_4} = -2.057248204929167 \approx -2$
- Μπορούμε επίσης να προσθέσουμε στο τέλος του κώδικα την γραμμή
`lambda = np.dot(np.matmul(A, y), y)/np.dot(y, y)`
θα πάρουμε ως αποτέλεσμα `-2.0000000000967284` για το παράδειγμα 2.

Μέθοδος των δυνάμεων : Ιδιοτιμή;

- Θεωρία : πρώτα υπολογίζουμε την ιδιοτιμή και μετά το ιδιοδιάνυσμα.
- Μέθοδος : δουλεύει ανάποδα, πρώτα το ιδιοδιάνυσμα.
- Ερώτημα : Πως θα βρούμε τη αντίστοιχη ιδιοτιμή;
- Πηλίκο Rayleigh

$$Ax = \lambda x \Rightarrow x^T Ax = \lambda x^T x \Rightarrow \lambda = \frac{x^T Ax}{x^T x}$$

- Επειδή $x_k \rightarrow x$, περιμένουμε ότι ο λόγος $\frac{x_k^T Ax_k}{x_k^T x_k}$ θα είναι μια καλή προσέγγιση της ιδιοτιμής. Για τα παραδείγματα έχουμε
 - ▶ Παράδειγμα 1: $\frac{x_3^T Ax_3}{x_3^T x_3} = -8.994851272863448 \approx -9$
 - ▶ Παράδειγμα 2: $\frac{x_4^T Ax_4}{x_4^T x_4} = -2.057248204929167 \approx -2$
- Μπορούμε επίσης να προσθέσουμε στο τέλος του κώδικα την γραμμή
`lambda = np.dot(np.matmul(A, y), y)/np.dot(y, y)`
θα πάρουμε ως αποτέλεσμα `-2.0000000000967284` για το παράδειγμα 2.

Μέθοδος των δυνάμεων : Ιδιοτιμή;

- Θεωρία : πρώτα υπολογίζουμε την ιδιοτιμή και μετά το ιδιοδιάνυσμα.
- Μέθοδος : δουλεύει ανάποδα, πρώτα το ιδιοδιάνυσμα.
- Ερώτημα : Πως θα βρούμε τη αντίστοιχη ιδιοτιμή;
- Πηλίκο Rayleigh

$$Ax = \lambda x \Rightarrow x^T Ax = \lambda x^T x \Rightarrow \lambda = \frac{x^T Ax}{x^T x}$$

- Επειδή $x_k \rightarrow x$, περιμένουμε ότι ο λόγος $\frac{x_k^T Ax_k}{x_k^T x_k}$ θα είναι μια καλή προσέγγιση της ιδιοτιμής. Για τα παραδείγματα έχουμε
 - ▶ Παράδειγμα 1: $\frac{x_3^T Ax_3}{x_3^T x_3} = -8.994851272863448 \approx -9$
 - ▶ Παράδειγμα 2: $\frac{x_4^T Ax_4}{x_4^T x_4} = -2.057248204929167 \approx -2$
- Μπορούμε επίσης να προσθέσουμε στο τέλος του κώδικα την γραμμή
`lambda = np.dot(np.matmul(A, y), y)/np.dot(y, y)`
θα πάρουμε ως αποτέλεσμα `-2.0000000000967284` για το παράδειγμα 2.

Μέθοδος των δυνάμεων : Ιδιοτιμή;

- Θεωρία : πρώτα υπολογίζουμε την ιδιοτιμή και μετά το ιδιοδιάνυσμα.
- Μέθοδος : δουλεύει ανάποδα, πρώτα το ιδιοδιάνυσμα.
- Ερώτημα : Πως θα βρούμε τη αντίστοιχη ιδιοτιμή;
- Πηλίκο Rayleigh

$$Ax = \lambda x \Rightarrow x^T Ax = \lambda x^T x \Rightarrow \lambda = \frac{x^T Ax}{x^T x}$$

- Επειδή $x_k \rightarrow x$, περιμένουμε ότι ο λόγος $\frac{x_k^T Ax_k}{x_k^T x_k}$ θα είναι μια καλή προσέγγιση της ιδιοτιμής. Για τα παραδείγματα έχουμε
 - ▶ Παράδειγμα 1: $\frac{x_3^T Ax_3}{x_3^T x_3} = -8.994851272863448 \approx -9$
 - ▶ Παράδειγμα 2: $\frac{x_4^T Ax_4}{x_4^T x_4} = -2.057248204929167 \approx -2$
- Μπορούμε επίσης να προσθέσουμε στο τέλος του κώδικα την γραμμή
`lambda = np.dot(np.matmul(A, y), y)/np.dot(y, y)`
θα πάρουμε ως αποτέλεσμα `-2.0000000000967284` για το παράδειγμα 2.

Μέθοδος των δυνάμεων : Ιδιοτιμή;

- Θεωρία : πρώτα υπολογίζουμε την ιδιοτιμή και μετά το ιδιοδιάνυσμα.
- Μέθοδος : δουλεύει ανάποδα, πρώτα το ιδιοδιάνυσμα.
- Ερώτημα : Πως θα βρούμε τη αντίστοιχη ιδιοτιμή;
- Πηλίκο Rayleigh

$$Ax = \lambda x \Rightarrow x^T Ax = \lambda x^T x \Rightarrow \lambda = \frac{x^T Ax}{x^T x}$$

- Επειδή $x_k \rightarrow x$, περιμένουμε ότι ο λόγος $\frac{x_k^T Ax_k}{x_k^T x_k}$ θα είναι μια καλή προσέγγιση της ιδιοτιμής. Για τα παραδείγματα έχουμε
 - ▶ Παράδειγμα 1: $\frac{x_3^T Ax_3}{x_3^T x_3} = -8.994851272863448 \approx -9$
 - ▶ Παράδειγμα 2: $\frac{x_4^T Ax_4}{x_4^T x_4} = -2.057248204929167 \approx -2$
- Μπορούμε επίσης να προσθέσουμε στο τέλος του κώδικα την γραμμή
`lambda = np.dot(np.matmul(A, y), y)/np.dot(y, y)`
θα πάρουμε ως αποτέλεσμα `-2.0000000000967284` για το παράδειγμα 2.

Μέθοδος των δυνάμεων : Ταχύτητα σύγκλισης;

Εστω A διαγωνοποιήσιμος και οι ιδιοτιμές του διατεταγμένες $|\lambda_1| > |\lambda_2| \geq \dots \geq |\lambda_n|$ και $\{x_1, x_2, \dots, x_n\}$ αντίστοιχα ιδιοδιανύσματα τα οποία αποτελούν βάση του \mathbb{R}^n . Τότε για το αρχικό τυχαίο διάνυσμα $y_0 \in \mathbb{R}^n$ της μεθόδου έχουμε $y_0 = c_1x_1 + c_2x_2 + \dots + c_nx_n$. Οπότε μετά από k βήματα της μεθόδου έχουμε

$$\begin{aligned} A^k y_0 &= \lambda_1^k c_1 x_1 + \lambda_2^k c_2 x_2 + \dots + \lambda_n^k c_n x_n \\ &= \lambda_1^k \left(c_1 x_1 + \left(\frac{\lambda_2}{\lambda_1} \right)^k c_2 x_2 + \dots + \left(\frac{\lambda_n}{\lambda_1} \right)^k c_n x_n \right) \\ &\sim \lambda_1^k c_1 x_1 \quad \text{για μεγάλο } k \end{aligned}$$

Επειδή λ_1 είναι η κυρίαρχη ιδιοτιμή ολοι οι όροι της μορφής $\left(\frac{\lambda_i}{\lambda_1} \right)^k \rightarrow 0$ για k αρκετά μεγάλο. Ο όρος λοιπόν που καθορίζει την ταχύτητα σύγκλισης είναι ο λ_2/λ_1 γιατί συγκλίνει στο 0 με τον μικρότερο ρυθμό. Στο προηγούμενο παράδειγμα έχουμε $\lambda_2/\lambda_1 = 1/2$ δηλ. σε κάθε βήμα της μεθόδου το σφάλμα (απόσταση από x) μειώνεται κατά $1/2$.

Μέθοδος των δυνάμεων : Ταχύτητα σύγκλισης;

Παρατηρείστε στη τρίτη στήλη το λόγο μεταξύ των διαδοχικών αριθμών, είναι περίπου $1/2$.

k	y	dyk
1	(-1 -0.35082011)	0.6591786104319128
2	(1 0.34124647)	0.00957364047847803
3	(-1 -0.33711056)	0.00413590705408573
.		
.		
.		
30	(1 0.33333333)	2.6922297724496502e-11
31	(-1 -0.33333333)	1.3461065595521404e-11
32	(1 0.33333333)	6.7305050421850865e-12

Μέθοδος των δυνάμεων : Υπέρ / κατά

- Πολύ απλή μέθοδο στη εφαρμογή της : απαιτεί μόνο πολλαπλασιασμό πίνακα επι διάνυσμα.
- Η ταχύτητα σύγκλισης είναι $|\lambda_2/\lambda_1|$ η οποία μπορεί να είναι πολύ αργή αν ο λόγος είναι κοντά στο 1.
- Προσεγγίζει μόνο το κυρίαρχο ιδιοδιάνυσμα και την αντίστοιχη ιδιοτιμή. Τι γίνεται αν χρειαζόμαστε και άλλα ιδιοδιανύσματα / ιδιοτιμές;
- Παραλλαγές της μεθόδου :
 - ▶ Μετατοπισμένη μέθοδος των δυναμέων : επιτάχυνση της σύγκλισης
 - ▶ Αντίστροφη μέθοδος των δυνάμεων : η μέθοδος εφαρμόζεται στον A^{-1}
 - ▶ Μέθοδος των πηλίκων Rayleigh
- Γενικά χρησιμοποιείται ως αρχικοποίηση άλλων πιο πολύπλοκων αλγορίθμων για την εύρεση ιδιοτιμών/ιδιοδιανυσμάτων

Μέθοδος των δυνάμεων : Υπέρ / κατά

- Πολύ απλή μέθοδο στη εφαρμογή της : απαιτεί μόνο πολλαπλασιασμό πίνακα επι διάνυσμα.
- Η ταχύτητα σύγκλισης είναι $|\lambda_2/\lambda_1|$ η οποία μπορεί να είναι πολύ αργή αν ο λόγος είναι κοντά στο 1.
- Προσεγγίζει μόνο το κυρίαρχο ιδιοδιάνυσμα και την αντίστοιχη ιδιοτιμή. Τι γίνεται αν χρειαζόμαστε και άλλα ιδιοδιανύσματα / ιδιοτιμές;
- Παραλλαγές της μεθόδου :
 - ▶ Μετατοπισμένη μέθοδος των δυναμέων : επιτάχυνση της σύγκλισης
 - ▶ Αντίστροφη μέθοδος των δυνάμεων : η μέθοδος εφαρμόζεται στον A^{-1}
 - ▶ Μέθοδος των πηλίκων Rayleigh
- Γενικά χρησιμοποιείται ως αρχικοποίηση άλλων πιο πολύπλοκων αλγορίθμων για την εύρεση ιδιοτιμών/ιδιοδιανυσμάτων

Μέθοδος των δυνάμεων : Υπέρ / κατά

- Πολύ απλή μέθοδο στη εφαρμογή της : απαιτεί μόνο πολλαπλασιασμό πίνακα επι διάνυσμα.
- Η ταχύτητα σύγκλισης είναι $|\lambda_2/\lambda_1|$ η οποία μπορεί να είναι πολύ αργή αν ο λόγος είναι κοντά στο 1.
- Προσεγγίζει μόνο το κυρίαρχο ιδιοδιάνυσμα και την αντίστοιχη ιδιοτιμή. Τι γίνεται αν χρειαζόμαστε και άλλα ιδιοδιανύσματα / ιδιοτιμές;
- Παραλλαγές της μεθόδου :
 - ▶ Μετατοπισμένη μέθοδος των δυναμέων : επιτάχυνση της σύγκλισης
 - ▶ Αντίστροφη μέθοδος των δυνάμεων : η μέθοδος εφαρμόζεται στον A^{-1}
 - ▶ Μέθοδος των πηλίκων Rayleigh
- Γενικά χρησιμοποιείται ως αρχικοποίηση άλλων πιο πολύπλοκων αλγορίθμων για την εύρεση ιδιοτιμών/ιδιοδιανυσμάτων

Μέθοδος των δυνάμεων : Υπέρ / κατά

- Πολύ απλή μέθοδο στη εφαρμογή της : απαιτεί μόνο πολλαπλασιασμό πίνακα επι διάνυσμα.
- Η ταχύτητα σύγκλισης είναι $|\lambda_2/\lambda_1|$ η οποία μπορεί να είναι πολύ αργή αν ο λόγος είναι κοντά στο 1.
- Προσεγγίζει μόνο το κυρίαρχο ιδιοδιάνυσμα και την αντίστοιχη ιδιοτιμή. Τι γίνεται αν χρειαζόμαστε και άλλα ιδιοδιανύσματα / ιδιοτιμές;
- Παραλλαγές της μεθόδου :
 - ▶ Μετατοπισμένη μέθοδος των δυναμέων : επιτάχυνση της σύγκλισης
 - ▶ Αντίστροφη μέθοδος των δυνάμεων : η μέθοδος εφαρμόζεται στον A^{-1}
 - ▶ Μέθοδος των πηλίκων Rayleigh
- Γενικά χρησιμοποιείται ως αρχικοποίηση άλλων πιο πολύπλοκων αλγορίθμων για την εύρεση ιδιοτιμών/ιδιοδιανυσμάτων

Μέθοδος των δυνάμεων : Υπέρ / κατά

- Πολύ απλή μέθοδο στη εφαρμογή της : απαιτεί μόνο πολλαπλασιασμό πίνακα επι διάνυσμα.
- Η ταχύτητα σύγκλισης είναι $|\lambda_2/\lambda_1|$ η οποία μπορεί να είναι πολύ αργή αν ο λόγος είναι κοντά στο 1.
- Προσεγγίζει μόνο το κυρίαρχο ιδιοδιάνυσμα και την αντίστοιχη ιδιοτιμή. Τι γίνεται αν χρειαζόμαστε και άλλα ιδιοδιανύσματα / ιδιοτιμές;
- Παραλλαγές της μεθόδου :
 - ▶ Μετατοπισμένη μέθοδος των δυναμέων : επιτάχυνση της σύγκλισης
 - ▶ Αντίστροφη μέθοδος των δυνάμεων : η μέθοδος εφαρμόζεται στον A^{-1}
 - ▶ Μέθοδος των πηλίκων Rayleigh
- Γενικά χρησιμοποιείται ως αρχικοποίηση άλλων πιο πολύπλοκων αλγορίθμων για την εύρεση ιδιοτιμών/ιδιοδιανυσμάτων

Η μέθοδος QR

Η μέθοδος QR : Ιδέα

- Παρουσιάστηκε το 1961 από τον J. Francis
- Επαναληπτικός αλγόριθμος ο οποίος μπορεί να υπολογίσει όλες τις ιδιοτιμές και ιδιοδιανύσματα του πίνακα
- **Βασικά βήματα της μεθόδου**
 - 1ο Μετασχηματισμός του αρχικού πίνακα A , μέσω μετασχηματισμών ομοιότητας, σε σχεδόν πάνω τριγωνικό (μορφή Hessenberg), ή τριδιαγώνιο αν ο A είναι συμμετρικός.
 - 2ο Υπολογισμός της παραγοντοποίησης QR του A και αντιστροφή των παραγόντων
 - 3ο Ο πίνακας A ανάγεται σε ένα πάνω τριγωνικό μετά από μια σειρά ορθομοναδιαίων μετασχηματισμών
- **Βασικό μειονέκτημα** : Ακόμα και αν ο A είναι μεγάλος αραιός πίνακας (το πλήθος των μη-μηδενικών στοιχείων είναι πολύ-πολύ μικρό), ήδη από τα πρώτα βήματα του αλγορίθμου ο πίνακας γεμίζει, με αποτέλεσμα να αυξάνεται θεαματικά η απαίτηση σε υπολογιστική μνήμη.
- Το υπολογιστικό κόστος της μεθόδου είναι $O(n^3)$ - πράξεις ανά βήμα του αλγορίθμου.

Η μέθοδος QR : αλγόριθμος

Η παρακάτω είναι η πιο βασική μορφή του αλγορίθμου. Για μεγάλα προβλήματα εύρεσης ιδιοτιμών και ιδιοδιανυσμάτων χρησιμοποιούνται διάφορες παραλλαγές που επιταχύνουν την σύγκλιση, μειώνουν τις απαιτήσεις σε μνήμη και υπολογιστικό κόστος.

Βασικός QR αλγόριθμος

$$A_1 = A$$

Για $k = 1, 2, \dots$

$$A_k = Q_k R_k, \text{ } QR\text{-παραγοντοποίηση του } A_k$$

$$A_{k+1} = R_k Q_k, \text{ αντιστροφή των παραγόντων}$$

Έχουμε δει ήδη ότι με την μέθοδο ορθοκανονικοποίησης Gram-Schmidt μπορούμε να γράψουμε $A = QR$ όπου Q είναι ορθογώνιος πίνακας ($Q^T Q = I$) και ο R είναι ανω τριγωνικός. Στη πράξη όμως δεν ακολουθείται αυτή η διαδικασία. Για την ανάλυση $A = QR$ συνήθως χρησιμοποιούνται οι μετασχηματισμοί Householder και οι στροφές Givens.

Η μέθοδος QR : ο κώδικας Python

Ο παραπάνω αλγόριθμος είναι απλός και είναι πολύ εύκολο να υλοποιηθεί με ένα κώδικα Python με την χρήση της βιβλιοθήκης numpy.

```
import numpy as np
import numpy.linalg as LA
n=3
D=np.diag([3,2,1])
P=np.random.rand(n,n)
Pinv=LA.inv(P)
A=np.matmul(np.matmul(P, D), Pinv)
for k in range(30):
    Q, R = LA.qr(A, mode='complete')
    print("_____")
    A = np.matmul(R,Q)
    print("A(" ,k, ")\n" , A)
```

Οι πρώτες γραμμές του κώδικα φτιάχνουν έναν 3×3 πίνακα με τυχαία στοιχεία ο οποίος έχει ιδιοτιμές $\lambda = 3, 2, 1$ (μπορείτε να βάλετε ότι τιμές θέλετε). Στη συνέχεια γίνεται η ανάλυση QR και φτιάχνεται ο νέος πίνακας σαν RQ .

Η μέθοδος QR : αποτελέσματα του κώδικα

Παρακάτω βλέπουμε τον αρχικό πίνακα $A(0)$, τον $A(19)$ μετά από 20 επαναλήψεις και $A(29)$ μετά από 30 επαναλήψεις. Παρατηρούμε τα διαγώνια στοιχεία συγκλίνουν στις ιδιοτιμές ενώ τα στοιχεία στο κάτω τρίγωνο πλησιάζουν το 0. Τα ιδιοδιανύσματα του πίνακα είναι οι στήλες του Q .

$A(0) =$

[2.49304485 -0.55373175 0.31876241]

[-0.98832819 3.98418722 -2.29657189]

[-0.80627541 1.93708017 -0.47723208]

$A(19) =$

[3.00015324e+00 5.10578304e-01 -4.14554908e+00]

[-3.00181545e-04 1.99984725e+00 1.38989889e+00]

[2.72578283e-10 -3.54248942e-07 9.99999507e-01]

$A(29) =$

[3.00000266e+00 5.10874749e-01 -4.14595864e+00]

[-5.20403537e-06 1.99999734e+00 1.38867653e+00]

[4.61543967e-15 -3.46131554e-10 1.00000000e+00]

Η μέθοδος QR : γιατί δουλεύει; - ιδιότητες

- Οι A_{k+1} , A_k είναι όμοιοι μεταξύ τους και όμοιοι με τον A

$$A_k = Q_k R_k \Rightarrow Q_k^T A_k = R_k \Rightarrow A_{k+1} = R_k Q_k = Q_k^T A_k Q_k \Rightarrow$$

$$A_{k+1} = Q_k^T Q_{k-1}^T A_{k-1} Q_{k-1} Q_k = \dots = Q_k^T Q_{k-1}^T \dots Q_1^T A Q_1 \dots Q_{k-1} Q_k$$

- Η ακολουθία A_k συγκλίνει σε ένα πάνω τριγωνικό πίνακα αν ο πίνακας έχει πραγματικές ιδιοτιμές. Οι ιδιοτιμές είναι τα διαγώνια στοιχεία του πίνακα.
- Αν ο A έχει και μιγαδικές ιδιοτιμές τότε η ακολουθία A_k συγκλίνει σε ένα άνω ημι-τριγωνικό πίνακα με τα στοιχεία της πρώτης κύριας υποδιαγωνίου να είναι μη-μηδενικά, (μορφή Schur).
- Μπορεί ναδειχθεί ότι για τα στοιχεία $a_{ij}^{(k)}$ κάτω από την διαγώνιο του πίνακα A_k ισχύει

$$|a_{ij}^{(k)}| \sim \left| \frac{\lambda_i}{\lambda_j} \right|^k \Rightarrow \frac{|a_{ij}^{(k+1)}|}{|a_{ij}^{(k)}|} \sim \left| \frac{\lambda_i}{\lambda_j} \right| \quad i > j$$

Η μέθοδος QR : γιατί δουλεύει; - ιδιότητες

Η τελευταία σχέση $\frac{|a_{ij}^{(k+1)}|}{|a_{ij}^{(k)}|} \sim \left| \frac{\lambda_i}{\lambda_j} \right|$ επαληθεύεται και απο τα αριθμητικά αποτελέσματα

A(30) / A(29)

```
[ 0.99999947  1.00000047 -0.99999992]
[ 0.66666454  1.0000008  -1.00000941]
[-0.3333328  -0.50000198  1.          ]
```

- Ο αλγόριθμος είναι πολύ ευσταθής. Η σύγκλιση μπορεί να είναι αργή αν οι ιδιοτιμές είναι κοντά η μία στη άλλη.
- Το υπολογιστικό κόστος της μεθόδου είναι αρκετά υψηλό γιατί σε κάθε βήμα του αλγορίθμου υπολογίζουμε την ανάλυση QR ενός πίνακα.
- Στη πράξη χρησιμοποιούνται παραλλαγές του αλγορίθμου με τεχνικές που βελτιώνουν σημαντικά το κόστος της μεθόδου, εκμεταλλεύομενοι τις ιδιότητες του πίνακα.