Η μέθοδος των δυνάμεων

Χρησιμοποιείται για τον προσδιορισμό της μέγιστης ιδιοτιμής και του αντίστοιχου ιδιοδιανύσματος ενός πίνακα. Ονομάζεται έτσι επειδή βασίζεται στην εσωτερική κατασκευή δυνάμεων του πίνακα A.

Έστω A $n \times n$ με ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_n$:
(δηλαδή διαγωνοποιείται $X^{-1}AX = \mathrm{diag}\{\lambda_1, \dots, \lambda_n\}$)

$$|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \dots \geq |\lambda_n|)$$

Δηλαδή λ_1 η μέγιστη ιδιοτιμή και έστω v_1 το αντίστοιχο ιδιοδιάνυσμα $\rightarrow (\lambda_1, v_1)$

Αλγόριθμος Μεθόδου Δυνάμεων (POWER)

B1 : Επιλέγετε $x^{(0)} \rightarrow$ προσέγγιση ιδιοδιανύσματος v_1

B2 : repeat

$$y^{(k)} = A \cdot x^{(k-1)} \longrightarrow \text{Πολλαπλασιασμός } O(n^2)$$

$$\lambda^{(k)} = \| y^{(k)} \|_\infty \quad \text{or} \quad \lambda^{(k)} = x^{(k)T} A x^{(k)}$$

$$x^{(k)} = \frac{y^{(k)}}{\| y^{(k)} \|_\infty}$$

ΘΕΩΡΗΜΑ: $\boxed{x^{(k)} \xrightarrow[k \to \infty]{} \mathrm{span}\{v_1\}, \quad \lambda^{(k)} \xrightarrow[k \to \infty]{} \lambda_1}$

Απόδειξη: Αφού ο A έχει ιδιοτιμές $\lambda_1, \lambda_2, \dots, \lambda_n$ τα ιδιοδιανύσματα $v_1, v_2, \dots, v_n$ είναι γραμμ. ανεξάρτητα, έτσι

$$x^{(0)} = a_1 v_1 + a_2 v_2 + \dots + a_n v_n, \quad a_1 \neq 0$$

Προϋπόθεση
$\lambda_1 \rightarrow$ dominant

Παρατηρούμε ότι

$$A^k x^{(0)} = A^k (a_1 v_1 + a_2 v_2 + \dots + a_n v_n) =$$

$$= a_1 \lambda_1^k v_1 + a_2 \lambda_2^k v_2 + \dots + a_n \lambda_n^k v_n =$$

$$= \lambda_1^k \left(a_1 v_1 + a_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \dots + a_n \left(\frac{\lambda_n}{\lambda_1}\right)^k v_n \right)$$

Τελικά από αλγόριθμο <u>POWER</u>

$$x^{(k)} = \frac{A^k x^{(0)} = y^{(k)}}{\max(A^k x_0) = \|y^{(k)}\|_\infty} =$$

$$= \frac{\lambda_1^k \left[a_1 v_1 + a_2 \left(\frac{\lambda_2}{\lambda_1}\right)^k v_2 + \dots + a_n \left(\frac{\lambda_n}{\lambda_1}\right)^k v_n \right]}{\max(A^k x_0)}$$

λ_1 dominant

$$\left(\frac{\lambda_i}{\lambda_1}\right)^k \xrightarrow[k \to \infty]{} 0 \quad , i = 2, 3, \dots, n$$

$$\lim x^{(k)} \to c \cdot v_1$$

$$x^{(k)} \to \text{span}\{v_1\}$$

$\lambda = $

Σχετικά με την ιδιοτιμή:

Αφού $\underline{x}^{(k-1)}$ είναι περίπου ιδιοδιάνυσμα που σχετίζ με τη λ_1

Έχουμε

$$\underline{y}^{(k)} = A\,\underline{x}^{(k-1)} \overset{\approx v_1}{\cong} \lambda_1 \cdot \underline{x}^{(k-1)} \Rightarrow$$

$$\|y^{(k)}\|_\infty \approx \lambda_1 \cdot \|x^{(k-1)}\|_\infty$$

Από την κατασκευή

$x^{(k-1)}_{p_k-1} = 1$ (μέγιστη συντεταγμένη του $\underline{x}^{(k-1)}$)

Άρα και $y^{(k)}_{p_k}$ συγκλίνει στην ιδιοτιμή λ_1.

ΚΡΙΤΗΡΙΑ ΔΙΑΚΟΠΗΣ (της μεθόδου)

$$|\lambda^{(m)} - \lambda^{(m-1)}| < TOL$$

$$\|\underline{x}^{(m)} - \underline{x}^{(m-1)}\|_\infty < TOL$$

---

ponding eigenvector.
tor g. Assume that A

$\to \infty$.

ited with $\lambda_1, \ldots, \lambda_n$
$= \alpha_1 v_1 + \alpha_2 v_2 +$

aints:

constraint, because
happen.
at the method still
nvalue. For exam-
|, and let there be

independent eigenvectors associated with λ_1. Then we have

$$A^k x_0 = \lambda_1^k \left(\sum_{i=1}^{r} \alpha_i v_i + \sum_{i=r+1}^{n} \alpha_i \left(\frac{\lambda_i}{\lambda_1} \right)^k v_i \right)$$

$$\cong \lambda_1^k \sum_{1}^{r} \alpha_i v_i$$

(because $(\lambda_i/\lambda_1)^k$ is small for large values of k). This shows that in this case the power method converges to some vector in the subspace spanned by $v_1, \ldots, v_r$.

Example 8.5.1

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

$$x_0 = (1, 1, 1)^T$$

The eigenvalues of A are 0, -0.6235, and 9.6235. The normalized eigenvector corresponding to the largest eigenvalue 9.6235 is

$$(0.3851, 0.5595, 0.7339)^T$$

Κανονικοποιημένο ιδιοδιάνυσμα που αντιστοιχεί στη μεγαλύτερη ιδιοτιμή

$$\begin{bmatrix} 0.3851 \\ 0.5595 \\ 0.7339 \end{bmatrix} \sim \begin{bmatrix} 0.5247 \\ 0.7623 \\ 1.0000 \end{bmatrix}$$

$k = 1$:

$$\hat{x}_1 = A x_0 = \begin{pmatrix} 6 \\ 9 \\ 12 \end{pmatrix}$$

$$\max(\hat{x}_1) = 12$$

$$x_1 = \frac{\hat{x}_1}{\max(\hat{x}_1)} = \begin{pmatrix} \frac{1}{2} \\ \frac{3}{4} \\ 1 \end{pmatrix} = \begin{pmatrix} 0.50 \\ 0.75 \\ 1 \end{pmatrix}$$

$k = 2$:

$$\hat{x}_2 = A x_1 = \begin{pmatrix} 5.00 \\ 7.25 \\ \underline{9.50} \end{pmatrix}, \qquad x_2 = \begin{pmatrix} 0.5263 \\ 0.7632 \\ 1.0000 \end{pmatrix}$$

$$\max(\hat{x}_2) = 9.50$$

$k = 3$:

$$\hat{x}_3 = A x_2 = \begin{pmatrix} 5.0526 \\ 7.3421 \\ \underline{9.6316} \end{pmatrix}$$

$$x_3 = \frac{\hat{x}_3}{\max(\hat{x}_3)}$$

$$= \begin{pmatrix} 0.5246 \\ 0.7623 \\ 1.000 \end{pmatrix}$$

$$\max(\hat{x}_3) = \underline{9.6316} \longrightarrow$$ συγκλίνει στην μεγαλύτερη ιδιοτιμή 9.6235

και το $\{x_k\}$ συγκλίνει στη διεύθυνση του ιδιοδιανύσματος που σχετίζεται μ' αυτή την ιδιοτιμή.

Προσδιορισμός Ιδιοτιμής

Ισχύει ότι

$$A x = \lambda \cdot x \Rightarrow$$

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

Στο κ-βήμα

$$\left.\begin{array}{l} \lambda_k = \dfrac{x^{(k)^T} A x^{(k)}}{x_k^T x_k} \\ \\ \lim x^{(k)} \longrightarrow v_1 \end{array}\right\} \Rightarrow \lim_{k \to \infty} \lambda_k = \frac{v_1^T A v_1}{v_1^T v_1} = \boxed{\lambda_1}$$

Σύγκλιση της μεθόδου των δυνάμεων

Έστω $x^{(0)} = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$, $a_1 \neq 0$

$$A^k x^{(0)} = a_1 \lambda_1^k \left(x_1 + \sum_{j=2}^{n} \frac{a_j}{a_1} \left(\frac{\lambda_j}{\lambda_1}\right)^k x_j \right)$$

Αφού $x^{(k)} \in \text{span}\{A^k x^{(0)}\}$

$$\Downarrow$$

$$\text{dist}\left(\text{span}\{x^{(k)}\}, \text{span}\{x_1\}\right) = O\left(\left|\frac{\lambda_2}{\lambda_1}\right|^k\right)$$

Αφού $|\lambda_1| > |\lambda_2| \geq \cdots \geq |\lambda_n|$

$\downarrow$

dominant eigenvalue

Κριτήριο Σύγκλισης μεθόδου δυνάμεων

λ_1 dominant eigenvalue

$x^{(0)}$ να έχει παράγοντα στη διεύθυνση του x_1

$\hookrightarrow$ αντίστοιχο dominant eigenvector

(άλλες περιπτώσεις στο ΑΕΡ.)

Παράδειγμα (Μέθοδος των δυνάμεων)

Έστω $A = \begin{bmatrix} -261 & 209 & -49 \\ -530 & 422 & -98 \\ -800 & 631 & -144 \end{bmatrix}$

$\lambda(A) = \{10, 4, 3\}$

Εάν εφαρμόσουμε τον αλγόριθμο της μεθόδου των δυνάμεων θα έχουμε:

Έστω $\underline{x}^{(0)} = [1, 0, 0]^T$

k	$\hat{\lambda}^{(k)}$
1	994.49
2	13.0606
3	10.7191
4	10.2073
5	10.0633
6	10.0198
7	10.0063
8	10.0020
9	10.0007
10	10.0002

Παρατήρηση:

Η χρησιμότητα της μεθόδου των δυνάμεων εξαρτάται από το λόγο $\left|\frac{\lambda_2}{\lambda_1}\right|$ αφού αυτός καθορίζει την ταχύτητα σύγκλισης.

Αντίστροφη Μέθοδος

Εάν A έχει ιδιοτιμές με

$$|\lambda_1| \geq |\lambda_2| \geq \dots \geq |\lambda_{n-1}| > |\lambda_n| > 0$$

τότε

A^{-1} έχει ιδιοτιμές

$$\left|\frac{1}{\lambda_n}\right| > \left|\frac{1}{\lambda_{n-1}}\right| \dots > \frac{1}{|\lambda_1|}$$

μέγιστη ιδιοτιμή του A^{-1} $\Rightarrow$ ελάχιστη του A

~~Αλγόριθμος~~ Αντίστροφη POWER

B1: Επιλέξτε $x^{(0)}$

B2: repeat

$$y^{(k)} = A^{-1} x^{(k-1)} \Rightarrow$$

$$A \cdot y^{(k)} = x^{(k-1)} \longrightarrow$$ Επίλυση συστήματος ως προς $y^{(k)}$

$$x^{(k)} = \frac{y^{(k)}}{\|y^{(k)}\|_\infty}$$

$$\lambda^{(k)} = \|y^{(k)}\|_\infty$$

Αντίστροφη μέθοδος των δυνάμεων (Inverse Iteration)

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

Αποδ: $A x_i = \lambda_i x_i \Rightarrow A^{-1} A x_i = \lambda_i A^{-1} x_i \Rightarrow$

$$\frac{1}{\lambda_i} x_i = A^{-1} x_i$$

Έστω $A \in \mathbb{R}^{n\times n}$ έχει n γραμμ. ανεξαρτ. ιδιοδιαν. $\lambda_i \in \mathbb{R}$

Έστω ποσότητα $q \in \mathbb{R}$ κοντά στην ιδιοτιμή λ_k:

$$|\lambda_k - q| < |\lambda_i - q|, \quad i = 1, 2, \dots, n, \quad i \neq k \Rightarrow$$

$\lambda_k - q$ η μικρότερη (κατ' απολ. τιμή) ιδιοτιμή του $A - qI$

Εάν εφαρμόσουμε τη μέθοδο των δυνάμεων στον $(A-qI)^{-1}$ υπολογίζουμε την $\frac{1}{(\lambda_k - q)}$ (μεγαλύτερη ιδιοτ. του $(A-qI)^{-1}$)

και από αυτήν βρίσκουμε τη λ_k.

Η μέθοδος των δυνάμεων με shift

Σε ορισμένες περιπτώσεις η σύγκλιση μπορεί να βελτιωθεί σημαντικά εάν χρησιμοποιηθεί κατάλληλο **shift**.

Εάν το σ είναι κατάλληλο shift έτσι ώστε $\lambda_1-\sigma$ μέγιστη ιδιοτιμή του $A-\sigma I$ και εφαρμόσουμε τη μέθοδο των δυνάμεων στο shifted πίνακα $A-\sigma I$ τότε ο λόγος σύγκλισης θα καθορίζεται από το πηλίκο

$$\left|\frac{\lambda_2-\sigma}{\lambda_1-\sigma}\right| \quad \text{αντί για το} \quad \left|\frac{\lambda_2}{\lambda_1}\right|$$

Εφαρμόζοντας shifting στον πίνακα οι ιδιοτιμές μετατοπίζονται κατά σ ενώ τα ιδιοδιανύσματα παραμένουν ίδια.

Με κατάλληλη επιλογή του σ σ' ορισμένες περιπτώσεις το κλάσμα $\left|\frac{\lambda_2-\sigma}{\lambda_1-\sigma}\right|$ μπορεί να είναι σημαντικά μικρότερο

από το $\left|\frac{\lambda_2}{\lambda_1}\right|$

Μια βέλτιστη επιλογή του σ υποθέτοντας ότι $\lambda_i \in \mathbb{R}$ είναι $\frac{1}{2}(\lambda_2+\lambda_n)$. Αυτή η επιλογή για ορισμένες περιπτώσεις δίνει **γρήγορη** σύγκλιση αλλά υπάρχουν και παραδείγματα που η σύγκλιση παραμένει αργή.

Π.χ. A 20x20, με ιδιοτιμές 20, 19, .., 2, 1. Εάν $\sigma = \frac{19+1}{2} = 10$

ο λόγος $\left|\frac{\lambda_2-\sigma}{\lambda_1-\sigma}\right| = \frac{9}{10}$ παραμένει **κοντά στο 1**. $\Rightarrow$ αργή σύγκλιση

Inverse Iteration

→ Αποτελεσματική μέθοδος υπολογ. ενός ιδιοδιανύσματος όταν είναι γνωστή καλή προσέγγιση μιας ιδιοτιμής

Εφαρμόζουμε τη μέθοδο των δυνάμεων στον πίνακα $(A-\sigma I)^{-1}$, όπου το σ ονομάζεται **shift**. Αυτό οδηγεί σε σύγκλιση στην ιδιοτιμή πολύ κοντύτερα στο σ παρά στο λ_1.

Επιλογή σ ώστε
$$|\lambda_i - \sigma| \ll |\lambda_1 - \sigma| \quad \forall i \Rightarrow$$
είναι κοντύτερα στο λ_1
$$\frac{1}{|\lambda_1-\sigma|} > \frac{1}{|\lambda_i-\sigma|} \quad \forall i$$

Αλγόριθμος

for $k = 1, 2, 3, \ldots$

$$\hat{\underline{x}}^{(k)} = (A-\sigma I)^{-1}\underline{x}^{(k-1)} \equiv \text{Επίλυση του } (A-\sigma I)\hat{\underline{x}}^{(k)} = \underline{x}^{(k-1)} \text{ Gauss}$$

$O(n^3)$ αρχικά LU, μετά $O(n^2)$ flops

$$\underline{x}^{(k)} = \frac{\hat{\underline{x}}^{(k)}}{\|\hat{\underline{x}}^{(k)}\|_\infty}$$ approximate ιδιοδιάνυσμα

$$\cancel{\lambda^{(k)} = \underline{x}^{(k)T} A\, \underline{x}^{(k)}}$$

Ο αλγόριθμος υπολογίζει το $z = \frac{1}{\lambda_1 - \sigma}$, οπότε $\lambda_1 = z^{-1} + \sigma$

έως ότου επιτευχθεί σύγκλιση

Κριτήριο: $\|(A-\sigma I)\underline{x}^{(k)}\|_\infty < c u \|A\|_\infty$

Διακοπής

$$\boxed{\|A\underline{x}^{(k)} - \underset{\text{ιδιοτιμή}}{\sigma}\underline{x}^{(k)}\| < \varepsilon}$$

Θεώρημα: Η ακολουθία $\{\underline{x}^{(k)}\}$ συγκλίνει ~~στο~~ στη διεύθυνση του ιδιοδιανύσματος που αντιστοιχεί στη λ_1

Απόδ.: Οι ιδιοτιμές του $(A-\sigma I)^{-1}$ είναι $(\lambda_1-\sigma)^{-1}, (\lambda_2-\sigma)^{-1}, \ldots, (\lambda_n-\sigma)^{-1}$ και τα ιδιοδ. $\underline{v}_1, \underline{v}_2, \ldots, \underline{v}_n$ τα ίδια με του A. Και όπως και στη μέθοδο των δυνάμεων έχουμε

$$\underline{X}^{(k)} = \frac{c_1}{(\lambda_1-\sigma)^k}\underline{v_1} + \frac{c_2}{(\lambda_2-\sigma)^k}\underline{v_2} + \dots + \frac{c_n}{(\lambda_n-\sigma)^k}\underline{v_n}$$

$$= \frac{1}{(\lambda_1-\sigma)^k}\left[c_1\underline{v_1} + c_2\left(\frac{\lambda_1-\sigma}{\lambda_2-\sigma}\right)^k + \dots + c_n\left(\frac{\lambda_1-\sigma}{\lambda_n-\sigma}\right)^k\underline{u_n}\right]$$

Το λ_1 είναι κοντύτερα στο σ και συνεπώς ο πρώτος όρος είναι ο κυρίαρχος όρος στο προηγούμενο άθροισμα, επομένως η ακολουθία $\{\underline{x_k}\}$ **συγκλίνει στη διεύθυνση του $\underline{V_1}$**.

Ευστάθεια της μεθόδου

Αρχικά μπορεί να θεωρηθεί επικίνδυνη διαδικασία επειδή εάν το σ είναι κοντά στο λ_1 ο πίνακας $(A-\sigma I)$ θα είναι ill-conditioned.

Αποδεικνύεται ότι το $\hat{x}^{(k)}$ είναι λύση του συστήματος

$$(A-\sigma I+F)\,\underline{x}^{(k)} = \underline{x}^{(k-1)}$$

Επιλογή του $\underline{x_0}$

Εκτελούμε ορισμένες επαναλήψεις της μεθόδου των δυναμέων και μετά μεταπηδούμε στην inverse iteration με το **τελευταίο** διάνυσμα που υπολογίσαμε στη μέθοδο των δυναμέων σαν **αρχικό** για την **inverse iteration**. ή εναλλακτικά

ΜΕΘΟΔΟΣ Έστω **σ κατάλληλη προσέγγιση** της λ_1 και μεταθετικός πίνακας P:

$$P(A-\sigma I) = LU$$

Τότε $P\underline{x_0} = L\underline{e}_{units}$ Τότε δύο επαναλήψεις της μεθόδου των δυναμέων αρκούν.

Παράδειγμα: Εστω $A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{bmatrix}$, $\underline{x}_0 = [1, 1, 1]^T$

$\sigma = 9$

Ιδιοτιμές: $0, \ -0.6235, \ 9.6235$

$k = 1$:

$$\hat{x}_1 = (1, 1.5, 2)^T$$

$$x_1 = \frac{\hat{x}_1}{\|\hat{x}_1\|_2} = (0.3714, 0.5571, 0.7428)^T$$

$k = 2$:

$$\hat{x}_2 = (0.6190, 0.8975, 1.1761)^T$$

$$x_2 = \frac{\hat{x}_2}{\|\hat{x}_2\|_2} = (0.3860, 0.5597, 0.7334)^T$$

$k = 3$:

$$\hat{x}_3 = (0.6176, 0.8974, \underline{1.1772})^T$$

$$x_3 = \frac{\hat{x}_3}{\|\hat{x}_3\|_2} = (0.3850, 0.5595, 0.7340)$$

$k = 4$:

$$\hat{x}_4 = (0.6176, 0.8974, \underline{1.1772})^T$$

$$x_4 = \frac{\hat{x}_4}{\|\hat{x}_4\|_2} = (0.3850, 0.5595, 0.7340)^T$$

$k = 5$:

$$\hat{x}_5 = (0.6177, 0.8974, \underline{1.1772})^T$$

$$x_5 = \frac{\hat{x}_5}{\|(\hat{x}_5)\|_2} = (0.3851, 0.5595, 0.7339)$$

$\underline{x}_5' \cdot A \, \underline{x}_5 = 9.6230$ → προσέγγιση της ιδιοτιμής

Παρατήρηση: Είτε χρησιμοποιούμε ως scaling την $\| \ \|_\infty$ είτε τη $\| \ \|_2$ προκύπτει διάνυσμα που βρίσκεται στη διεύθυνση του ιδιοδιανύσματος που αντιστοιχεί στη λ_1.

Rayleigh Quotient

Έστω $\underline{x} \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$ συμμετρικός, το rayleigh quotient είναι το βαθμωτό

real symmetric

$$r(\underline{x}) = \frac{\underline{x}^T A \underline{x}}{\underline{x}^T \underline{x}} = \frac{(x, Ax)}{(x, x)}$$

Εάν το $\underline{x}$ είναι ιδιοδιάνυσμα τότε $r(\underline{x}) = \lambda$ η αντίστοιχη ιδιοτιμή.

MOTIVATION

Δοσμένου ενός $\underline{x}$ ποιο a συμπεριφέρεται σαν ιδιοτιμή; για το $\underline{x}$ δεδομένου ότι ελαχιστοποιεί την ποσότητα

$$\| A\underline{x} - \underbrace{a\underline{x}}_{m\times 1} \|_2$$

Πρόβλημα Ελαχίστων Τετραγώνων

$$\| \underbrace{A\underline{x}}_{\text{γνωστό}} - \underbrace{a}_{\text{αγνωστο}} \underline{x} \|_2$$

πίνακας

$$a \cdot \underline{x} = A \cdot \underline{x}$$

$\Downarrow$ Κανονικές Εξισώσεις

$$\underline{x} \underbrace{a}_{\text{αγν}} = A\underline{x}$$

$$\underline{x}^T \cdot \underline{x}\, a = \underline{x}^T A \underline{x} \Rightarrow a = \frac{\underline{x}^T A \underline{x}}{\underline{x}^T \underline{x}}$$

Το πηλίκο αυτό προσεγγίζει την ιδιοτιμή

$\longrightarrow$

__Παράδειγμα__

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 3 \end{bmatrix}, \quad \underline{x} = \begin{bmatrix} 1 \\ -0.5 \end{bmatrix}$$

Το πηλίκο Rayleigh

$$\sigma = \frac{\underline{x}^T A \underline{x}}{\underline{x}^T \underline{x}} = -0.2$$

καλή προσέγγιση της ιδιοτιμής -0.2361

__Άσκηση__ Έστω $\underline{q}_j$ ιδιοδιάνυσμα του A

Να αποδειχθεί ότι

$$r(\underline{x}) - r(\underline{q}_j) = O\left(\| \underline{x} - \underline{q}_j \|^2 \right)$$

καθώς $\underline{x} \to \underline{q}_j$

$\Rightarrow$ Το πηλίκο Rayleigh τετραγωνική εκτίμηση μιας ιδιοτιμής

__Υποδ:__ $\underline{x} = \sum_{j=1}^{m} a_j \underline{q}_j$, $\quad r(\underline{x}) = \sum_{j=1}^{m} a_j^2 \lambda_j \Big/ \sum_{j=1}^{m} a_j^2$

Έστω $\underline{x} \neq \underline{0}$ τότε το πηλίκο

$$\mu = \rho = \frac{\underline{x}^T A \underline{x}}{\underline{x}^T \underline{x}}$$

ελαχιστοποιεί την ποσότητα $\| (A - \mu I)\underline{x} \|$

olves only solution

3
5062 4.7407
1.0200

The normalized eigenvector correct to *four* digits is

$$\begin{pmatrix} 0.3851 \\ 0.5595 \\ 0.7339 \end{pmatrix}$$

The Rayleigh Quotient

THEOREM 8.5.3 Let A be a *symmetric matrix* and let x be a reasonably good approximation to an eigenvector. Then the quotient

$$R_q = \sigma = \frac{x^T A x}{x^T x}$$

is a good approximation to the eigenvalue λ for which x is the corresponding eigenvector.

⇒ $A = P^T \Sigma P$, eigenv real, orthonormal eigenvectors

Proof: Because A is symmetric, by Theorem 8.2.8, there exists a set of orthonormal eigenvectors $v_1, v_2, \ldots, v_n$. Therefore, we can write

$$x = c_1 v_1 + \cdots + c_n v_n$$ → x καλή προσέγγιση στο V_1

Assume that v_i, $i = 1, \ldots, n$ are normalized—that is, $v_i^T v_i = 1$. Then, since $Av_i = \lambda_i v_i$, $i = 1, \ldots, n$, and noting that $v_i^T v_j = 0$, $i \neq j$, we have

$$\sigma = \frac{x^T A x}{x^T x} = \frac{(c_1 v_1 + \cdots + c_n v_n)^T A (c_1 v_1 + \cdots + c_n v_n)}{(c_1 v_1 + \cdots + c_n v_n)^T (c_1 v_1 + \cdots + c_n v_n)}$$

$$= \frac{(c_1 v_1 + \cdots + c_n v_n)^T (c_1 \lambda_1 v_1 + \cdots + c_n \lambda_n v_n)}{c_1^2 + c_2^2 + \cdots + c_n^2}$$

$$= \frac{\lambda_1 c_1^2 + \lambda_2 c_2^2 + \cdots + \lambda_n c_n^2}{c_1^2 + c_2^2 + \cdots + c_n^2}$$

$$= \lambda_1 \left[\frac{1 + \left(\frac{\lambda_2}{\lambda_1}\right)\left(\frac{c_2}{c_1}\right)^2 + \cdots + \left(\frac{\lambda_n}{\lambda_1}\right)\left(\frac{c_n}{c_1}\right)^2}{1 + \left(\frac{c_2}{c_1}\right)^2 + \cdots + \left(\frac{c_n}{c_1}\right)^2} \right]$$

Because of our assumption that x is a good approximation to v_1, c_1 is larger than other c_i, $i = 2, \ldots, n$. Thus, the expression within [ ] is close to 1, which means that σ is close to λ_1. ■

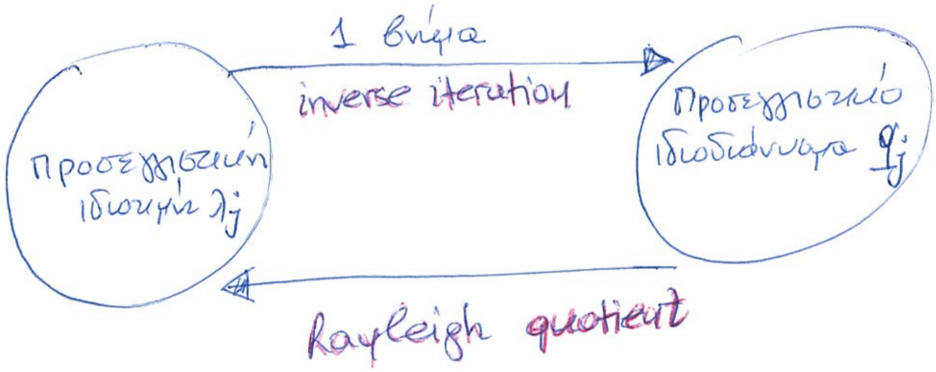
In Theorem 8.5.3 we have defined the quotient

$$R = \frac{x^T A x}{x^T x}$$

where x is an eigenvector of A and have seen that this quotient is a good approximation to an eigenvalue λ for which x is the corresponding eigenvector. In general, if x is a

Μέθοδος επανάληψης Rayleigh-Quotient (Trefethen)

Η ιδέα της προσέγγισης μίας ιδιοτιμής με τη μέθοδο Rayleigh μπορεί να συνδυαστεί με την inverse iteration για να υπολογίζουμε διαδοχικές προσεγγίσεις μίας ιδιοτιμής και το αντίστοιχο ιδιοδιάνυσμα με ένα επαναληπτικό σχήμα. που είναι γνωστό σαν επανάληψη του πηλίκου Rayleigh (Rayleigh quotient iteration)



Αλγόριθμος

$\underline{v}^{(0)}$ = κάποιο διάνυσμα με $\|\underline{v}^{(0)}\| = 1$

$\lambda^{(0)} = (\underline{v}^{(0)})^T A\, \underline{v}^{(0)}$ → αντίστοιχο Rayleigh quotient

for $k = 1, 2, \dots$

Επίλυση του συστήματος

$(A - \lambda^{(k-1)} I)\,\underline{w} = \underline{v}^{(k-1)}$ → inverse iteration

$\underline{v}^{(k)} = \dfrac{\underline{w}}{\|\underline{w}\|}$ → κανονικοποίηση

$\lambda^{(k)} = (\underline{v}^{(k)})^T A\, (\underline{v}^{(k)})$ → πηλίκο Rayleigh

Παρατήρηση: Η σύγκλιση του αλγορίθμου είναι εντυπωσιακή. Κάθε επανάληψη τριπλασιάζει τον αριθμό των ψηφίων ακριβείας.

Αποδεικνύεται ότι $\|\underline{v}^{(k+1)} - (\pm \underline{q}_j)\| = O\left(\|\underline{v}^{(k)} - (\pm \underline{q}_j)\|^3\right)$

$$|\lambda^{(k+1)} - \lambda_j| = O\left(|\lambda^{(k)} - \lambda_j|^3\right), \quad k \to \infty$$

όπου $\underline{q}_j$ το ιδιοδιάνυσμα που αντιστοιχεί στη λ_j.

__Πολυπλοκότητα:__

Για κάθε $\lambda^{(k)}$ επιλύεται ένα καινούργιο σύστημα οπότε αυτό την καθιστά ιδιαίτερα δαπανηρή. Της τάξης $O(n^3)$ flops ανά επανάληψη. Εάν ο πίνακας είναι τριδιαγώνιος τότε τα flops μειώνονται σε $O(n)$.

__Παράδειγμα:__

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

Ιδιοτιμές

0, −0.6235, 9.6235

↓ Αντιστ. Ιδιοδιάνυσμα

$[0.3851 \quad 0.5595 \quad 0.7339]^T$

Έστω $\underline{x}_0 = \begin{bmatrix} 0.5246 \\ 0.7622 \\ 1.000 \end{bmatrix}$ → Προσδιορίζεται μετά από 3 επαναλήψεις της μεθόδου των δυνάμεων

$k = 0$:

$$\sigma_0 = \frac{x_0^T A x_0}{x_0^T x_0} = 9.6235$$

$$x_1 = \begin{pmatrix} 0.5247 \\ 0.7623 \\ 1.000 \end{pmatrix}$$

$k = 1$:

$$\sigma_1 = \frac{x_1^T A x_1}{x_1^T x_1} = 9.6235$$

$$x_2 = \begin{pmatrix} 1.000 \\ 1.4529 \\ 1.9059 \end{pmatrix}$$

Παράδειγμα: Έστω ο συμμετρικός πίνακας

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

και έστω $\underline{v}^{(0)} = (1,1,1)^T/\sqrt{3}$ αρχική προσέγγιση ιδιοδιανύσματος
Όταν εφαρμόσουμε την επανάληψη Rayleigh quotient προκύπτουν οι ακόλουθες τιμές για τα $\lambda^{(k)}$ στις πρώτες 3 επαναλήψεις

$$\lambda^{(0)} = 5, \quad \lambda^{(1)} = 5.2131\ldots, \quad \lambda^{(2)} = 5.2143197431\!84\ldots$$

Η θεωρητική τιμή της ιδιοτιμής που αντιστοιχεί στο ιδιοδιάνυσμα που είναι κοντύτερα στο $\underline{v}^{(0)}$ είναι $\lambda = 5.2143197433\!77$
Μετά από μόνο τρεις επαναλήψεις προκύπτει αποτέλεσμα που έχει 10 ψηφία ακριβείας. Τρεις επιπλέον επαναλήψεις αυξάνουν αυτό το αποτέλεσμα σε περίπου 270 ψηφία, εάν υπήρχε δυνατότητα υπολογισμού τόσο μεγάλης ακρίβειας.

Υπολογισμός της δεύτερης μεγαλύτερης ιδιοτιμής και αντίστοιχου ιδιοδιανύσματος: Μέθοδος Deflation

Αφού έχει υπολογισθεί η μέγιστη ιδιοτιμή λ_1 και το αντίστοιχο ιδιοδιάνυσμα $\underline{v}_1$ η αμέσως επόμενη ιδιοτιμή λ_2 μπορεί να υπολογισθεί χρησιμοποιώντας deflation. Η βασική ιδέα έγκειται στην αντικατάσταση του αρχικού πίνακα με έναν άλλο πίνακα μικρότερης διάστασης με τη βοήθεια μιας επιλεγμένης ιδιοτιμής και του αντίστοιχου ιδιοδιανύσματος.
Ο καινούργιος πίνακας έχει τις ίδιες ιδιοτιμές και ιδιοδιανύσματα με τον αρχικό εκτός από την ιδιοτιμή που χρησιμοποιήθηκε για την αναγωγή.

Householder deflation

Θα κατασκευάσουμε deflation ενός πίνακα A χρησιμοποιώντας μετασχηματισμούς ομοιότητας με πίνακες Householder, οι οποίοι εξασφαλίζουν ευστάθεια. Η μέθοδος χρησιμοποιεί το ακόλουθο θεώρημα

Θεώρημα: Έστω $(\lambda_1, \underline{v_1})$ ιδιοτιμή και το αντίστοιχο ιδιοδιάνυσμα του A και H πίνακας Householder έτσι ώστε $H\underline{v_1} = \text{span}\{\underline{e_1}\}$

Τότε

$$A_1 = HAH = \begin{bmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & & & & \\ \vdots & & A_2 & & \\ 0 & & & & \end{bmatrix} = \left[\begin{array}{c|c} \lambda_1 & \underline{b}^T \\ \hline 0 & A_2 \end{array}\right]$$

$(H^T = H)$

όπου A_2 $(n-1)\times(n-1)$ και οι ιδιοτιμές του A_2 είναι ίδιες μ' αυτές του A_1 εκτός από τη λ_1; ιδιαίτερα εάν $|\lambda_1| > |\lambda_2| \geq |\lambda_3| \geq \cdots \geq |\lambda_n|$ τότε κυρίαρχη ιδιοτιμή στον A_2 είναι η λ_2, η οποία είναι η $2^{η}$ κυρίαρχ στον A.

Απόδειξη: Από τη σχέση $A\underline{v_1} = \lambda_1 \underline{v_1}$ έχουμε

$$HAHH\underline{v_1} = \lambda_1 H\underline{v_1} \quad (\text{αφού } H^2 = I)$$

Έτσι $HAH(k\underline{e_1}) = \lambda_1 k \underline{e_1}$ αφού $H\underline{v_1} = k\underline{e_1}$, $k \in \mathbb{R}$

Τελικά $HAH\underline{e_1} = \lambda_1 \underline{e_1} \Rightarrow$ Η πρώτη στήλη του HAH είναι λ_1 φορές η πρώτη στήλη του ταυτοτικού πίνακα.

Έτσι ο HAH πρέπει να έχει τη μορφή:

$$A_1 = HAH = \begin{bmatrix} \lambda_1 & * & * & \cdots & * \\ 0 & & & & \\ \vdots & & A_2 & & \\ 0 & & & & \end{bmatrix}$$

Επειδή $\underbrace{\det(HAH - \lambda I)}^{\text{Ιδιοτιμές του HAH}} = (\lambda_1 - \lambda)\det(A_2 - \lambda I) \Rightarrow$ οι ιδιοτιμές του A_2 είναι οι ίδιες με του A_1 εκτός της λ_1. Επιπλέον εάν

$|\lambda_1| > |\lambda_2| > |\lambda_3| \geq |\lambda_4| \geq \dots \geq |\lambda_n|$

η κυρίαρχη ιδιοτιμή του A_2 είναι η λ_2, η οποία είναι η δεύτερη κυρίαρχη ιδιοτιμή του A.

Αλγόριθμος Householder Deflation

B1: Υπολογίστε την κυρίαρχη ιδιοτιμή λ_1 και το αντίστοιχο ιδιοδιάνυσμα $\underline{v_1}$ χρησιμοποιώντας μέθοδο δυνάμεων και inverse iteration.

B2: Προσδιόρισε πίνακα Householder H :

$$H\,\underline{v_1} = \begin{bmatrix} * \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

B3: Υπολόγισε τον πίνακα HAH

B4: Διέγραψε την 1η γραμμή και στήλη του HAH και βρες την κυρίαρχη ιδιοτιμή του $\overset{\text{εναπομένοντα}}{(n-1)\times(n-1)}$ πίνακα.

Παράδειγμα:

$$A = \begin{pmatrix} 0.2190 & 0.6793 & 0.5194 \\ 0.0470 & 0.9347 & 0.8310 \\ 0.6789 & 0.3835 & 0.0346 \end{pmatrix}$$

Οι ιδιοτιμές του A είναι 0.0018, −0.3083, 1.4947.

1. $\lambda_1 = 1.4947,\ v_1 = \begin{pmatrix} -0.5552 \\ -0.7039 \\ -0.4430 \end{pmatrix}$

2. $u = v_1 - \|v_1\|_2 e_1 = \begin{pmatrix} -1.5552 \\ -0.7039 \\ -0.4430 \end{pmatrix}$