ΓΡΑΜΜΙΚΗ ΑΛΓΕΒΡΑ

Ε. Γαλλόπουλος & Ε. Στεφανόπουλος

CEID

Διάλεξη 10

8 Μαΐου 2018

Έννοιες, λέξεις κλειδιά, υπενθύμιση I

Υπενθύμιση της 9ης δ.

- Εσωτερικό γινόμενο, νόρμα και ελάχιστα τετράγωνα σε διανυσματικούς χώρους συναρτήσεων.
- Gram-Schmidt σε χώρους πολυωνύμων και βέλτιστη προσέγγιση. συναρτήσεων.
- Gram-Schmidt σε χώρους πολυωνύμων και βέλτιστη προσέγγιση. συναρτήσεων.
- Ενδιαφέρουσες βαθμωτές συναρτήσεις μητρώων: Ίχνος και Ορίζουσα
- αξιωματικός ορισμός ορίζουσας
- τρόποι υπολογισμού ορίζουσας
- ιδιότητες ορίζουσας
- επίλυση συστήματος και αντιστροφή μητρώου μέσω οριζουσών.
- τύπος Sylvester
- επίλυση συστήματος και αντιστροφή μητρώου μέσω οριζουσών.

Έννοιες, λέξεις κλειδιά, υπενθύμιση II

Σήμερα θα συζητήσουμε (όροι 10ης δ.):

- Ιδιοτιμές και ιδιοδιανύσματα: Τι, γιατί, ορισμοί.
- Χαρακτηριστικό πολυώνυμο.
- Θεώρημα Cayley-Hamilton και εφαρμογές.
- Συνοδευτικό μητρώο πολυωνύμου.
- Δεξιά και αριστερά ιδιοδιανύσματα.
- Κανονικοποιημένα ιδιοδιανύσματα.
- Υπολογισμός ιδιοτιμών και ιδιοδιανυσμάτων.
- Ελλειμματικές ιδιοτιμές και μητρώα.
- Διαγωνιοποίηση και διαγωνιοποιήσιμα μητρώα.
- Φασματικό ανάπτυγμα μητρώου.
- Ιδιοτιμές και ιδιοδιανύσματα δυνάμεων μητρώου και πολυωνύμων μητρώου.
- Απόδειξη του Θ Cayley-Hamilton.
- Υπολογισμός ιδιοδιανυσμάτων και ιδιοτιμών: η μέθοδος δύναμης.

Υπό συζήτηση ενότητες

Down with Determinants!

Sheldon Axler

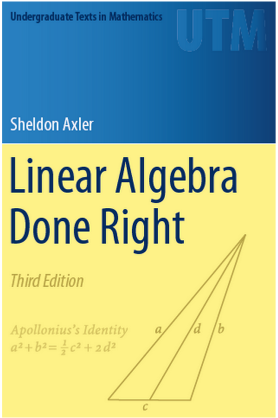
This paper was published in the *American Mathematical Monthly* 102 (1995), 139-154.

In 1996 this paper received the Lester R. Ford Award for expository writing from the Mathematical Association of America.

Abstract: This paper shows how linear algebra can be done better without determinants. The standard proof that a square matrix

complexification discussed earlier.

Now we are ready for the formal definition. The *determinant* of T, denoted $\det T$, is defined to be the product of the eigenvalues of T, counting multiplicity. This definition would not be possible with the traditional approach to eigenvalues, because that method uses determinants to prove that eigenvalues exist. With the techniques used here, we already know (by Theorem 3.11(a)) that T has $\dim V$ eigenvalues, counting multiplicity. Thus our simple definition makes sense.



This paper focuses on showing that determinants should be banished from much of the theoretical part of linear algebra. Determinants are also useless in the computational part of linear algebra. For example, Cramer's rule for solving systems of linear equations is already worthless for 10×10 systems, not to mention the much larger systems often encountered in the real world. Many computer programs efficiently calculate eigenvalues numerically—none of them uses determinants. To emphasize the point, let me quote a numerical analyst. Henry Thacher, in a review (*SIAM News*, September 1988) of the *Turbo Pascal Numerical Methods Toolbox*, writes,

> I find it hard to conceive of a situation in which the numerical value of a determinant is needed: Cramer's rule, because of its inefficiency, is completely impractical, while the magnitude of the determinant is an indication of neither the condition of the matrix nor the accuracy of the solution.

A Randomized Algorithm for Approximating the Log Determinant of a Symmetric Positive Definite Matrix

Christos Boutsidis * Petros Drineas † Prabhanjan Kambadur ‡
Eugenia-Maria Kontopoulou § Anastasios Zouzias ¶

Abstract

We introduce a novel algorithm for approximating the logarithm of the determinant of a symmetric positive definite (SPD) matrix. The algorithm is randomized and approximates the traces of a small number of matrix powers of a specially constructed matrix, using the method of Avron and Toledo [AT11]. From a theoretical perspective, we present additive and relative error bounds for our algorithm. Our additive error bound works for any SPD matrix, whereas our relative error bound works for SPD matrices whose eigenvalues lie in the interval $(\theta_1, 1)$, with $0 < \theta_1 < 1$; the latter setting was proposed in [HMS15]. From an empirical perspective, we demonstrate that a C++ implementation of our algorithm can approximate the logarithm of the determinant of large matrices very accurately in a matter of seconds.

Αλγεβρικό πρόβλημα ιδιοτιμών

Στις επόμενες διαλέξεις επικεντρωνόμαστε στο εξής πρόβλημα:

Δοθέντος $A \in \mathbb{R}^{n\times n}$ ή $A \in \mathbb{C}^{n\times n}$ να βρεθούν τα ζεύγη $\lambda \in \mathbb{C}, x \in \mathbb{C}^n$ τέτοια ώστε

$$Ax = \lambda x.$$

ή αντίστοιχα,

Δοθέντος $A \in \mathbb{R}^{n\times n}$ ή $A \in \mathbb{C}^{n\times n}$ να βρεθούν τα ζεύγη $\lambda \in \mathbb{C}, y \in \mathbb{C}^n$ τέτοια ώστε

$$y^* A = \lambda y^*.$$

όπου $y^* = \bar{y}^\top$ και με $\bar{y}$ συμβολίζεται το διάνυσμα με τιμές τις συζυγείς του y. Αν $y \in \mathbb{R}^n$ τότε $y^* = y^\top$.

Ιδιοτιμές και ιδιοδιανύσματα I

Ορισμός

Έστω ότι $A \in \mathbb{C}^{n\times n}$. Τότε κάθε μιγαδικός αριθμός λ για τον οποίο το μητρώο $(\lambda I - A)$ είναι μη αντιστρέψιμο ονομάζεται **ιδιοτιμή του A**.

- Έπεται ότι η διάσταση του μηδενόχωρου $\dim \operatorname{null}(\lambda I - A) \geq 1$. Κάθε διάνυσμά $x \in \operatorname{null}(\lambda I - A)$ ονομάζεται (δεξιό) **ιδιοδιάνυσμα του A** (για την ιδιοτιμή λ).
- Έπεται ότι η διάσταση του μηδενόχωρου $\dim \operatorname{null}(\lambda I - A)^* \geq 1$. Κάθε διάνυσμά $y \in \operatorname{null}(\lambda I - A)^*$ ονομάζεται (αριστερό) **ιδιοδιάνυσμα του A** (για την ιδιοτιμή λ), οπότε $y^* A = \lambda y^*$.

Οι ιδιοτιμές και τα ιδιοδιανύσματα ενός μητρώου αναδεικνύουν πολλά σημαντικά ζητήματα για ένα μητρώο και για τους διανυσματικούς (υπο)χώρους που συνδέονται με αυτό και χρησιμοποιούνται πολύ στις εφαρμογές.

Ιδιοτιμές και ιδιοδιανύσματα II

Ερωτήματα:

- Υπάρχουν ιδιοτιμές; Πόσες υπάρχουν;
- Υπάρχουν ιδιοδιανύσματα; Πόσα ιδιοδιανύσματα (ή καλύτερα, ποιές είναι οι διαστάσεις των αντίστοιχων μηδενοχώρων;)
- Πού και πώς εντοπίζονται (ιδιοδιανύσματα, ιδιοτιμές); Πώς υπολογίζονται;
- Τι χαρακτηριστικά έχουν και γιατί μας ενδιαφέρουν;

Το μηδέν ΜΠΟΡΕΙ να είναι ιδιοτιμή, το μηδενικό διάνυσμα ΔΕΝ θεωρείται ιδιοδιάνυσμα.

SIAM REVIEW
Vol. 48, No. 3, pp. 569–581

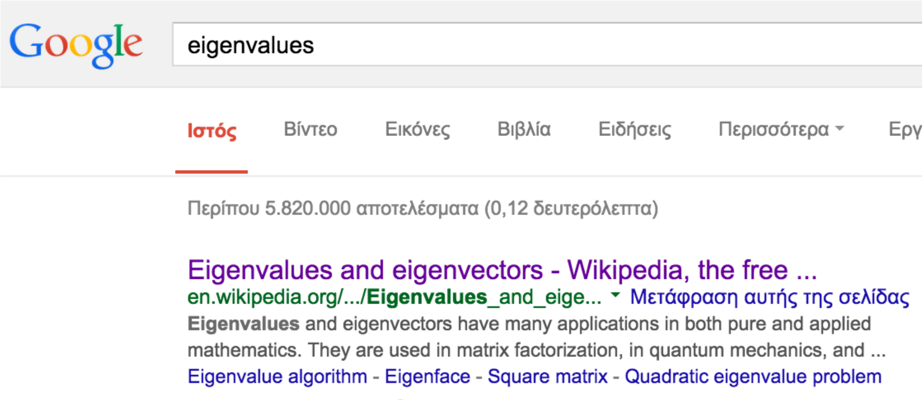
The $25,000,000,000 Eigenvector: The Linear Algebra behind Google*

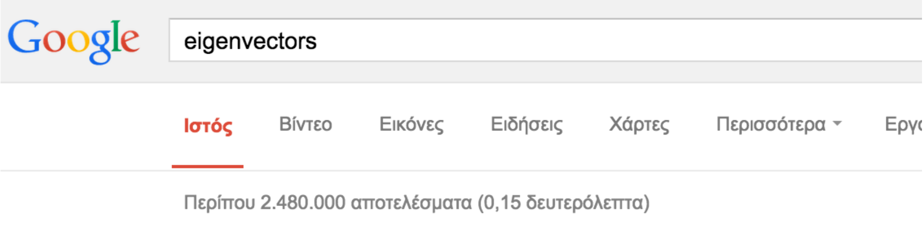
Kurt Bryan†
Tanya Leise‡

Abstract. Google's success derives in large part from its PageRank algorithm, which ranks the importance of web pages according to an eigenvector of a weighted link matrix. Analysis of the PageRank formula provides a wonderful applied topic for a linear algebra course. Instructors may assign this article as a project to more advanced students or spend one or two lectures presenting the material with assigned homework from the exercises. This material also complements the discussion of Markov chains in matrix algebra. Maple and *Mathematica* files supporting this material can be found at www.rose-hulman.edu/~bryan.

Key words. linear algebra, PageRank, eigenvector, stochastic matrix

Τι λέει και τι γράφει ο κόσμος για τις ιδιοτιμές και τα ιδιοδιανύσματα;





Mathematics for Quantum Mechanics
John David Jackson
LOOK INSIDE!
Eigenvalues of Matrices
Françoise Chatelin
Free Vibrations of Beams and Frames
Eigenvalues and Eigenfunctions
Igor A. Karnovsky
Olga I. Lebed
Probability and its Applications
Eigenvalues, Inequalities, and Ergodic Theory
Numerical Methods for Large Eigenvalue Problems
Revised Edition
Yousef Saad
Geometric Control Theory and Sub-Riemannian Geometry
Eigenvalues of Symmetric Five-Diagonal Matrices
L. W. Ehrlich
The Asymptotic Distribution of Eigenvalues of Partial Differential Operators
Extremum Problems for Eigenvalues of Elliptic Operators
Antoine Henrot
On the Eigenvalues of the Manakov System
LAMBERT
Eigenvalues, Embeddings and Generalised Trigonometric Functions
Jan Lang
David Edmunds
Springer
Volume IX
Surveys in Differential Geometry
Eigenvalues of Laplacians and other geometric operators
Eigenvalues in Uniform Waveguides Using Transmission Line Equivalences
COLLOQUIUM PUBLICATIONS
Random Matrices, Frobenius Eigenvalues, and Monodromy
London Mathematical Society Lecture Note Series
Cambridge University Press
Isaac Elishakoff
Eigenvalues of Inhomogeneous Structures
Unusual Closed-Form Solutions
CRC Press
Memoirs of the American Mathematical Society
Kuznetsov's Trace Formula and the Heck Eigenvalues of Maass Forms
A. Knightly
C. Li
American Mathematical Society
Calculus of Variations
and Partial Differential Equations
G. Prodi (Ed.)
Eigenvalues of Non-Linear Problems
CIME
The Symmetric Eigenvalue Problem
Beresford N. Parlett
Classics In Applied Mathematics
Multiparameter Eigenvalue Problems
Sturm-Liouville Theory
F.V. Atkinson
Angelo B. Mingarelli
Eigenvalues of Matrices
Françoise Chatelin
Wiley

Παράδειγμα

Από το $A = \begin{pmatrix} -3 & 2 \\ -12 & 7 \end{pmatrix}$ κατασκευάζουμε το παραμετροποιημένο

$$A(\lambda) = \lambda I - A = \begin{pmatrix} \lambda + 3 & -2 \\ 12 & \lambda - 7 \end{pmatrix}$$

Οι ιδιοτιμές του A είναι όλες εκείνες οι τιμές λ (ενδεχομένως μιγαδικές) για τις οποίες το $A(\lambda)$ δεν είναι αντιστρέψιμο. Στην περίπτωσή μας, μπορούμε να επαληθεύσουμε ότι τα $A(1), A(3)$ δεν είναι αντιστρέψιμα. Επίσης

$$A(1)u_1 = 0, A(3)u_2 = 0$$

και λύνοντας τα αντίστοιχα ομογενή συστήματα προκύπτει ότι (ειδικές) λύσεις τους είναι οι $u_1 = (1,2)^\top$ και $u_2 = (1,3)^\top$. Προφανώς ισχύει ότι $\mathrm{null}(A(1)) = \mathrm{span}\{u_1\}$ και $\mathrm{null}(A(3)) = \mathrm{span}\{u_2\}$. Επομένως, ως ιδιοδιάνυσμα του A για την ιδιοτιμή 1 μπορούμε να επιλέξουμε οποιοδήποτε διάνυσμα του $\mathrm{null}(A(1))$. Συνήθως, επιλέγουμε διανύσματα που έχουν κανονικοποιηθεί με κάποιον τρόπο, π.χ. τέτοια ώστε $\|u_1\|_2 = \|u_2\|_2 = 1$.

Προσοχή: Το παραπάνω παράδειγμα δεν εξηγεί με ποιόν τρόπο επιλέξαμε τα $A(1)$ και $A(3)$. Ούτε γιατί αυτά είναι τα μόνα δύο μη αντιστρέψιμα μητρώα του τύπου $\lambda I - A$. Ο σχεδιασμός συστηματικών μεθόδων αναζήτησης και υπολογισμού τους είναι ένα από τα βασικά ζητούμενα όχι μόνον αυτού του κεφαλαίου, αλλά και πολλών βιβλίων και της σημερινής έρευνας.

Ένα χαρακτηριστικό πολυώνυμο για κάθε μητρώο

Οι πληροφορίες που θέλουμε εξάγονται από ένα ειδικό πολυώνυμο που υπολογίζεται από την ορίζουσα $\det(\lambda I - A)$ ως προς τη μεταβλητή λ.

Εξετάζουμε το $\det(\lambda I - A)$:

- Αν

$$\begin{aligned} \lambda I - A &= \begin{pmatrix} \lambda - \alpha_{11} & -\alpha_{12} \\ -\alpha_{21} & \lambda - \alpha_{22} \end{pmatrix} \\ \det(\lambda I - A) &= (-\alpha_{11} + \lambda)(-\alpha_{22} + \lambda) - \alpha_{21}\alpha_{12} \\ &= \lambda^2 - (\alpha_{11} + \alpha_{22})\lambda + \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12} \end{aligned}$$

Ένα χαρακτηριστικό πολυώνυμο για κάθε μητρώο

Οι πληροφορίες που θέλουμε εξάγονται από ένα ειδικό πολυώνυμο που υπολογίζεται από την ορίζουσα $\det(\lambda I - A)$ ως προς τη μεταβλητή λ.

Εξετάζουμε το $\det(\lambda I - A)$:

- Αν

$$\begin{aligned}
\lambda I - A &= \begin{pmatrix} \lambda - \alpha_{11} & -\alpha_{12} \\ -\alpha_{21} & \lambda - \alpha_{22} \end{pmatrix} \\
\det(\lambda I - A) &= (-\alpha_{11} + \lambda)(-\alpha_{22} + \lambda) - \alpha_{21}\alpha_{12} \\
&= \lambda^2 - (\alpha_{11} + \alpha_{22})\lambda + \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}
\end{aligned}$$

- Το $\det(A - \lambda I)$ είναι πολυώνυμο 2ου βαθμού ως προς λ.

Ένα χαρακτηριστικό πολυώνυμο για κάθε μητρώο

Χαρακτηριστικά πολυώνυμα και ιδιοτιμές

Οι πληροφορίες που θέλουμε εξάγονται από ένα ειδικό πολυώνυμο που υπολογίζεται από την ορίζουσα $\det(\lambda I - A)$ ως προς τη μεταβλητή λ.

Εξετάζουμε το $\det(\lambda I - A)$:

- Αν

$$\begin{aligned}
\lambda I - A &= \begin{pmatrix} \lambda - \alpha_{11} & -\alpha_{12} \\ -\alpha_{21} & \lambda - \alpha_{22} \end{pmatrix} \\
\det(\lambda I - A) &= (-\alpha_{11} + \lambda)(-\alpha_{22} + \lambda) - \alpha_{21}\alpha_{12} \\
&= \lambda^2 - (\alpha_{11} + \alpha_{22})\lambda + \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}
\end{aligned}$$

- Το $\det(A - \lambda I)$ είναι πολυώνυμο 2ου βαθμού ως προς λ.
- Προσέξτε τους συντελεστές των 2 μεγαλύτερων δυνάμεων λ^2, λ και τη σταθερά:

$$\begin{aligned}
\gamma_2 &= 1, \\
\gamma_1 &= -(\alpha_{11} + \alpha_{22}) \\
\gamma_0 &= \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}
\end{aligned}$$

Ορισμοί με ορίζουσες

Χαρακτηριστικό πολυώνυμο και ιδιοτιμές

- *Χαρακτηριστικό πολυώνυμο* ενός μητρώου $A \in \mathbb{R}^{n\times n}$, ονομάζεται το πολυώνυμο βαθμού n με το οποίο ισούται η ορίζουσα

$$\begin{aligned} p(\lambda) &= \det(\lambda I - A) \text{ ή «ισοδύναμα»} \\ \hat{p}(\lambda) &= \det(A - \lambda I) = (-1)^n p(\lambda). \end{aligned}$$

Ορισμοί με ορίζουσες

Χαρακτηριστικό πολυώνυμο και ιδιοτιμές

- *Χαρακτηριστικό πολυώνυμο* ενός μητρώου $A \in \mathbb{R}^{n\times n}$, ονομάζεται το πολυώνυμο βαθμού n με το οποίο ισούται η ορίζουσα

$$\begin{aligned} p(\lambda) &= \det(\lambda I - A) \text{ ή «ισοδύναμα»} \\ \hat{p}(\lambda) &= \det(A - \lambda I) = (-1)^n p(\lambda). \end{aligned}$$

- Αν γράψουμε

$$p(\lambda) = \lambda^n + \gamma_{n-1}\lambda^{n-1} + \cdots + \gamma_1\lambda + \gamma_0$$

ισχύει ότι $\gamma_{n-1} = -\mathrm{trace}(A)$, $\gamma_0 = (-1)^n \det(A)$

Ορισμοί με ορίζουσες

Χαρακτηριστικό πολυώνυμο και ιδιοτιμές

- *Χαρακτηριστικό πολυώνυμο* ενός μητρώου $A \in \mathbb{R}^{n\times n}$, ονομάζεται το πολυώνυμο βαθμού n με το οποίο ισούται η ορίζουσα

$$\begin{aligned} p(\lambda) &= \det(\lambda I - A) \text{ ή «ισοδύναμα»} \\ \hat{p}(\lambda) &= \det(A - \lambda I) = (-1)^n p(\lambda). \end{aligned}$$

- Αν γράψουμε

$$p(\lambda) = \lambda^n + \gamma_{n-1}\lambda^{n-1} + \cdots + \gamma_1\lambda + \gamma_0$$

ισχύει ότι $\gamma_{n-1} = -\mathrm{trace}(A)$, $\gamma_0 = (-1)^n \det(A)$

- προσέξτε ότι αν το μέγεθος του μητρώου είναι περιττό, οι συντελεστές των p και $\hat{p}$ έχουν διαφορετικό πρόσημο.

Ορισμοί με ορίζουσες

Χαρακτηριστικό πολυώνυμο και ιδιοτιμές

- *Χαρακτηριστικό πολυώνυμο* ενός μητρώου $A \in \mathbb{R}^{n\times n}$, ονομάζεται το πολυώνυμο βαθμού n με το οποίο ισούται η ορίζουσα

$$\begin{aligned} p(\lambda) &= \det(\lambda I - A) \text{ ή «ισοδύναμα»} \\ \hat{p}(\lambda) &= \det(A - \lambda I) = (-1)^n p(\lambda). \end{aligned}$$

- Αν γράψουμε

$$p(\lambda) = \lambda^n + \gamma_{n-1}\lambda^{n-1} + \cdots + \gamma_1\lambda + \gamma_0$$

  ισχύει ότι $\gamma_{n-1} = -\mathrm{trace}(A)$, $\gamma_0 = (-1)^n \det(A)$
- προσέξτε ότι αν το μέγεθος του μητρώου είναι περιττό, οι συντελεστές των p και $\hat{p}$ έχουν διαφορετικό πρόσημο.
- Οι ρίζες του χ.π. ονομάζονται ΙΔΙΟΤΙΜΕΣ (του μητρώου).

Ορισμοί με ορίζουσες

Χαρακτηριστικό πολυώνυμο και ιδιοτιμές

- *Χαρακτηριστικό πολυώνυμο* ενός μητρώου $A \in \mathbb{R}^{n\times n}$, ονομάζεται το πολυώνυμο βαθμού n με το οποίο ισούται η ορίζουσα

$$
\begin{aligned}
p(\lambda) &= \det(\lambda I - A) \text{ ή «ισοδύναμα»} \\
\hat{p}(\lambda) &= \det(A - \lambda I) = (-1)^n p(\lambda).
\end{aligned}
$$

- Αν γράψουμε

$$p(\lambda) = \lambda^n + \gamma_{n-1}\lambda^{n-1} + \cdots + \gamma_1\lambda + \gamma_0$$

ισχύει ότι $\gamma_{n-1} = -\mathrm{trace}(A)$, $\gamma_0 = (-1)^n \det(A)$

- προσέξτε ότι αν το μέγεθος του μητρώου είναι περιττό, οι συντελεστές των p και $\hat{p}$ έχουν διαφορετικό πρόσημο.
- Οι ρίζες του χ.π. ονομάζονται ΙΔΙΟΤΙΜΕΣ (του μητρώου).
- Το πολυώνυμο έχει n ρίζες και άρα n ιδιοτιμές. Μερικές ή όλες μπορεί να είναι μεταξύ τους ίσες (πολλαπλές). Το σύνολο των ιδιοτιμών λέγεται φάσμα του A.

Μια κρίσιμη ιδιότητα του Χ.Π. - αφετηρία της θεωρίας μητρώων (Cayley'1858)

Έστω ότι $A \in \mathbb{R}^{2\times 2}$, οπότε

$$\det(\lambda I - A) \quad = \quad \lambda^2 - (\alpha_{11} + \alpha_{22})\lambda + \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}.$$

Αντικαθιστούμε το λ με το A:

$$\begin{aligned}
p(A) &= A^2 - (\alpha_{11} + \alpha_{22})A + (\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12})I \\
&= \begin{pmatrix} \alpha_{11}^2 + \alpha_{12}\alpha_{21} & \alpha_{11}\alpha_{12} + \alpha_{12}\alpha_{22} \\ \alpha_{21}\alpha_{11} + \alpha_{22}\alpha_{21} & \alpha_{21}\alpha_{12} + \alpha_{22}^2 \end{pmatrix} - (\alpha_{11} + \alpha_{22}) \begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} + \begin{pmatrix} (\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}) & 0 \\ 0 & (\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}) \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix}
\end{aligned}$$

Μια κρίσιμη ιδιότητα του Χ.Π. - αφετηρία της θεωρίας μητρώων (Cayley' 1858)

Έστω ότι $A \in \mathbb{R}^{2\times 2}$, οπότε

$$\det(\lambda I - A) = \lambda^2 - (\alpha_{11}+\alpha_{22})\lambda + \alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}.$$

Αντικαθιστούμε το λ με το A:

$$\begin{aligned} p(A) &= A^2 - (\alpha_{11}+\alpha_{22})A + (\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12})I \\ &= \begin{pmatrix} \alpha_{11}^2 + \alpha_{12}\alpha_{21} & \alpha_{11}\alpha_{12} + \alpha_{12}\alpha_{22} \\ \alpha_{21}\alpha_{11} + \alpha_{22}\alpha_{21} & \alpha_{21}\alpha_{12} + \alpha_{22}^2 \end{pmatrix} - (\alpha_{11}+\alpha_{22})\begin{pmatrix} \alpha_{11} & \alpha_{12} \\ \alpha_{21} & \alpha_{22} \end{pmatrix} + \begin{pmatrix} (\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}) & 0 \\ 0 & (\alpha_{11}\alpha_{22} - \alpha_{21}\alpha_{12}) \end{pmatrix} \\ &= \begin{pmatrix} 0 & 0 \\ 0 & 0 \end{pmatrix} \end{aligned}$$

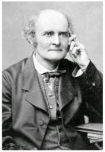
Φαίνεται ότι $p(A) = 0$. Αυτή η ιδιότητα ισχύει για **οποιοδήποτε τετραγωνικό μητρώο,** *$A \in \mathbb{C}^{n\times n}$.*

Θεώρημα Cayley-Hamilton

Έστω το χαρακτηριστικό πολυώνυμο $p(\lambda) = \det(\lambda I - A)$. Τότε

$$\boxed{p(A) = 0_{n\times n}}.$$

A Memoir on the Theory of Matrices
Author(s): Arthur Cayley
Source: *Philosophical Transactions of the Royal Society of London*, Vol. 148 (1858), pp. 17-37
Published by: The Royal Society
Stable URL: http://www.jstor.org/stable/108649
Accessed: 03/05/2010 17:45



II. *A Memoir on the Theory of Matrices.* By ARTHUR CAYLEY, *Esq., F.R.S.*

Received December 10, 1857,—Read January 14, 1858.

THE term matrix might be used in a more general sense, but in the present memoir I consider only square and rectangular matrices, and the term matrix used without qualification is to be understood as meaning a square matrix; in this restricted sense, a set of quantities arranged in the form of a square, *e. g.*

$$\begin{pmatrix} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{pmatrix}$$

is said to be a matrix. The notion of such a matrix arises naturally from an abbreviated notation for a set of linear equations, viz. the equations

$$X = ax + by + cz,$$
$$Y = a'x + b'y + c'z,$$
$$Z = a''x + b''y + c''z,$$

may be more simply represented by

$$(X, Y, Z) = \begin{pmatrix} a, & b, & c \\ a', & b', & c' \\ a'', & b'', & c'' \end{pmatrix}(x, y, z),$$

ΠΡΟΣΟΧΗ

Η $\det$ είναι συνάρτηση όλων των στοιχείων του μητρώου, π.χ. για $n = 2$, όλων των

$$\alpha_{11} - \lambda, \alpha_{12}, \alpha_{21}, \alpha_{22} - \lambda$$

Επομένως ΔΕΝ μπορούμε να πούμε ότι $p(A) = \det(A - AI)\det(0) = 0$.

Δείτε την Wikipedia για μερικές αποδείξεις του θεωρήματος (υπάρχουν πολλές !)
Θα έχουμε την ευκαιρία να τη συζητήσουμε σε επόμενη διάλεξη ....

Μια (ακόμα) ιδιομορφία των μητρώων

... επιπλέον των $AB \neq BA$, $AB = 0$ ακόμα και αν $A \neq 0, B \neq 0$, ... που έπεται από το θεώρημα Cayley-Hamilton:

Για τις δυνάμεις $A^n = -\gamma_{n-1}A^{n-1} - \gamma_{n-2}A^{n-2} - \cdots + \gamma_1 A + \gamma_0 I.$

Δηλ. για κάθε μητρώο $A \in \mathbb{R}^{n\times n}$, οι δυνάμεις μεγαλύτερες από $n-1$ μπορούν να γραφτούν ως γραμμικός συνδυασμός χαμηλοτέρων δυνάμεων!

Για το αντίστροφο (αν υπάρχει)

$$\begin{aligned} 0 &= A^n + \gamma_{n-1}A^{n-1} + \cdots + \gamma_1 A + \gamma_0 I \\ &= A^{-1}\left(A^n + \gamma_{n-1}A^{n-1} + \cdots + \gamma_1 A + \gamma_0 I\right) \end{aligned}$$

επομένως

$$A^{-1} = -\frac{1}{\gamma_0}\left(A^{n-1} + \gamma_{n-1}A^{n-2} + \cdots + \gamma_1 I\right)$$

Ιδιοτιμές χωρίς ορίζουσες (κατά S. Axler)

Για κάθε $A \in \mathbb{C}^{n\times n}$, ορίζεται ως ιδιοτιμή κάθε βαθμωτός λ για τον οποίο το μητρώο $A-\lambda I$ δεν είναι αντιστρέψιμο και ιδιοχώρος του A που αντιστοιχεί στο λ ο μηδενόχωρος του $A-\lambda I$. **Θα δείξουμε ότι υπάρχει τουλάχιστον μια ιδιοτιμή.**

- Δίνεται $A \in \mathbb{R}^{n\times n}$ και μη μηδενικό $x \in \mathbb{R}^n$. Τότε τα $n+1$ διανύσματα

$$x, Ax, A^2x, \dots., A^nx$$

  είναι οπωσδήποτε γραμμικά εξαρτημένα (γιατί;).
- Επομένως υπάρχουν συντελεστές $\gamma_0, \cdots, \gamma_n$ όχι όλοι μηδέν τέτοιοι ώστε

$$0 = \gamma_0 x + \gamma_1 Ax + \cdots \gamma_n A^n x = p(A)x$$

- Δεν γνωρίζουμε ακριβώς το βαθμό του πολυωνύμου! Μπορεί να είναι μικρότερος του n.
- Αφού $p(A)x = 0$ και $x \neq 0$, το μητρώο $p(A)$ δεν είναι αντιστρέψιμο.
- Αντικαθιστώντας το A με βαθμωτή μεταβλητή ζ, $p(\zeta)=0$, και έστω ότι $\deg p = m \leq n$ και ότι οι ρίζες του είναι $\rho_1, \dots, \rho_m \in \mathbb{C}$. Τότε

$$p(z) = \gamma(\zeta-\rho_1)(\zeta-\rho_2)(\zeta-\rho_m).$$

  όπου γ είναι ο συντελεστής του μεγιστοβάθμιου όρου.
- Έπεται ότι ένας τουλάχιστον από τους παράγοντες του

$$p(A) = \gamma(A-\rho_1 I)(A-\rho_2 I),(A-\rho_m I)$$

  δεν είναι αντιστρέψιμος. Αν είναι ο παράγοντας $A-\rho_j I$, το ρ_j θα είναι ιδιοτιμή.

Ανακεφαλαίωση

Βασικός ορισμός

Για ένα μητρώο $A \in \mathbb{R}^{n\times n}$ και μεταβλητή λ, η ορίζουσα $p(\lambda) = \det(\lambda I - A)$ είναι πολυώνυμο βαθμού n. Λέγεται χαρακτηριστικό πολυώνυμο και αν το γράψουμε

$$p(\lambda) \quad = \quad \lambda^n + \gamma_{n-1}\lambda^{n-1} + \cdots + \gamma_1\lambda + \gamma_0$$

τότε $\gamma_{n-1} = -\mathrm{trace}(A)$ και $\gamma_0 = (-1)^n \det(A)$.

Προσοχή:

- Το πολυώνυμο έχει ακριβώς n ρίζες $\lambda_1, \cdots \lambda_n \in \mathbb{C}^n$ που αποκαλούμε **ιδιοτιμές** του A.
- Για κάθε ιδιοτιμή, λ, ο μηδενόχωρος $\mathrm{null}(\lambda I - A)$ λέγεται **ιδιόχωρος** και έχει διάσταση τουλάχιστον 1. Κάθε διάνυσμα βάσης του χώρου ονομάζεται **ιδιοδιάνυσμα**.
- Ισχύει ότι $p(A) = 0$ **(θεώρημα Cayley-Hamilton)**.
- τα ιδιοζεύγη (ιδιοτιμές και ιδιοδιανύσματα) μητρώων παίζουν πολύ σημαντικό ρόλο σε πληθώρα εφαρμογών.

Πολλαπλασιασμός με βαθμωτό

Έστω ότι $Ax = \lambda x$ και έστω βαθμωτός $\gamma \neq 0$. Τότε

$$\gamma(Ax) \quad = \quad \gamma(\lambda x).$$

<u>Δύο ερμηνείες:</u>

$(\gamma A)x = (\gamma\lambda)x$: αν πολλαπλασιάσουμε το μητρώο με βαθμωτό, το ιδιοδιάνυσμα παραμένει αμετάβλητο και η ιδιοτιμή πολλαπλασιάζεται με το βαθμωτό.

$A(\gamma x) = \lambda(\gamma x)$: αν πολλαπλασιάσουμε το ιδιοδιάνυσμα με βαθμωτό, έχουμε πάλι ιδιοδιάνυσμα και η ιδιοτιμή παραμένει αμετάβλητη.

Επομένως, όλα τα διανύσματα συγγραμικά με το ιδιοδιάνυσμα είναι και αυτά ιδιοδιανύσματα. Συνήθως κανονικοποιούμε ώστε να λαμβάνουμε ως "εκπρόσωπο της κατεύθυνσης" το διάνυσμα που έχει μήκος 1.

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου.

.

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου.
2. Υπολογισμός των ριζών του (που είναι οι ιδιοτιμές).

.

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου.
2. Υπολογισμός των ριζών του (που είναι οι ιδιοτιμές).

.

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου. Δαπανηρή και αριθμητικά προβληματική για μεγάλα προβλήματα.
2. Υπολογισμός των ριζών του (που είναι οι ιδιοτιμές).

Για κάθε ιδιοτιμή λ_j, συνήθως ενδιαφέρει ένα σύνολο από γραμμικά ανεξάρτητα ιδιοδιανύσματα που είναι βάση για το $\mathrm{null}(\lambda_j I - A)$. Αυτά είναι ειδικές λύσεις του $(\lambda_j I - A)x = 0$. Συνήθως τα διανύσματα επιλέγονται κανονικοποιημένα.

.

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου. Δαπανηρή και αριθμητικά προβληματική για μεγάλα προβλήματα.
2. Υπολογισμός των ριζών του (που είναι οι ιδιοτιμές). Αν $n \geq 5$ τότε δεν υπάρχει αναλυτικός τύπος υπολογισμού των ριζών και πρέπει να χρησιμοποιηθεί (επαναληπτικός) αλγόριθμος προσέγγισης ριζών πολυωνύμου[1].
3. Για κάθε ιδιοτιμή, επίλυση του $(A - \lambda I)x = 0$ και επιλογή των ιδιοδιανυσμάτων (ανεύρεση των ειδικών λύσεων).

.

[1] Η αδυναμία αυτή αποτελεί ένα βασικό εύρημα των Μαθηματικών του 19ου αιώνα (Abel, Rufffini, και Galois)

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου. Δαπανηρή και αριθμητικά προβληματική για μεγάλα προβλήματα.
2. Υπολογισμός των ριζών του (που είναι οι ιδιοτιμές). Αν $n \geq 5$ τότε δεν υπάρχει αναλυτικός τύπος υπολογισμού των ριζών και πρέπει να χρησιμοποιηθεί (επαναληπτικός) αλγόριθμος προσέγγισης ριζών πολυωνύμου[1].
3. Για κάθε ιδιοτιμή, επίλυση του $(A - \lambda I)x = 0$ και επιλογή των ιδιοδιανυσμάτων (ανεύρεση των ειδικών λύσεων).

.

[1] Η αδυναμία αυτή αποτελεί ένα βασικό εύρημα των Μαθηματικών του 19ου αιώνα (Abel, Rufffini, και Galois)

Σχετικά με την επίλυση του Αλγεβρικού Προβλήματος Ιδιοτιμών

1. Εύρεση του χαρακτηριστικού πολυωνύμου.
2. Υπολογισμός των ριζών του (που είναι οι ιδιοτιμές).

Τα παραπάνω είναι μόνο για μικρά προβλήματα και δεν χρησιμοποιείται: Στην πράξη (π.χ. στη MATLAB) χρησιμοποιούνται ειδικοί αλγόριθμοι (συνάρτηση `eig`) για την εύρεση των ιδιοτιμών (πχ. αλγόριθμος QR. Μάλιστα, το χ.π. υπολογίζεται μετά από κλήση στην `eig` για τον υπολογισμό των ιδιοτιμών και με υπολογισμό των συντελεστών της δυναμομορφής από την παραγοντοποιημένη μορφή. Δηλαδή, η διαδικασία που ακολουθείται έχει την αντίθετη φορά από αυτήν που χρησιμοποιούμε εδώ[1].

[1] Η αριθμητική επίλυση του ΑΠΙ είναι ένα σημαντικό επιστημονικό αντικείμενο των περιοχών της Υπολογιστικής Γραμμικής Άλγεβρας και της Αριθμητικής Ανάλυσης.

Παραδείγματα

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

$p(\lambda) = (\lambda - 1)(\lambda - 3)$, $\Lambda(A) = \{1, 3\}$,
$\lambda_1 = 1, \mu_1 = 1$, $\lambda_2 = 3, \mu_2 = 1$.

Παραδείγματα

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

$p(\lambda) = (\lambda - 1)(\lambda - 3), \Lambda(A) = \{1, 3\},$
$\lambda_1 = 1, \mu_1 = 1, \lambda_2 = 3, \mu_2 = 1.$

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix},$$

$p(\lambda) = (\lambda - 1)^2, \Lambda(A) = \{1\},$
$\lambda_1 = 1, \mu_1 = 2.$

Παραδείγματα

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

$p(\lambda) = (\lambda - 1)(\lambda - 3)$, $\Lambda(A) = \{1, 3\}$,
$\lambda_1 = 1, \mu_1 = 1$, $\lambda_2 = 3, \mu_2 = 1$.

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix},$$

$p(\lambda) = (\lambda - 1)^2$, $\Lambda(A) = \{1\}$,
$\lambda_1 = 1, \mu_1 = 2$.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix},$$

$p(\lambda) = (\lambda - 1)(\lambda - 2)^2$,
$\Lambda(A) = \{1, 2\}$, $\lambda_1 = 1, \mu_1 = 1$,
$\lambda_2 = 2, \mu_2 = 2$.

Παραδείγματα

$$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix},$$

$p(\lambda) = (\lambda - 1)(\lambda - 3)$, $\Lambda(A) = \{1, 3\}$,
$\lambda_1 = 1, \mu_1 = 1$, $\lambda_2 = 3, \mu_2 = 1$.

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix},$$

$p(\lambda) = (\lambda - 1)^2$, $\Lambda(A) = \{1\}$,
$\lambda_1 = 1, \mu_1 = 2$.

$$A = \begin{pmatrix} 4 & 2i & 0 & -2i \\ 2i & 4 & -2i & 0 \\ 0 & -2i & 4 & 2i \\ -2i & 0 & 2i & 4 \end{pmatrix}$$

$p(\lambda) = (\lambda - 4)^2(\lambda - 4 - 4i)(\lambda - 4 + 4i)$,
$\Lambda(A) = \{2, 4 + 4i, 4 - 4i\}$, $\lambda_1 = 4, \mu_1 = 2$,
$\lambda_2 = 4 + 4i, \mu_2 = 1$, $\lambda_3 = 4 - 4i, \mu_3 = 1$.

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 1 \\ 0 & 0 & 2 \end{pmatrix},$$

$p(\lambda) = (\lambda - 1)(\lambda - 2)^2$,
$\Lambda(A) = \{1, 2\}$, $\lambda_1 = 1, \mu_1 = 1$,
$\lambda_2 = 2, \mu_2 = 2$.

Παραδείγματα

Συμβολισμοί: για κάθε διακριτή ιδιοτιμή λ_j, η πολλαπλότητά της συμβολίζεται με μ_j.

$$A = \begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix}$$

$$p(\lambda) = \lambda^4 - 8\lambda^3 + 16\lambda^2 = \lambda^2(\lambda - 4)^2$$

επομένως

$$\lambda_1 = 0, \mu_1 = 2, \lambda_2 = 4, \mu_2 = 2.$$

και

$$\Lambda(A) = \{0, 4\}$$

Συνοδευτικό μητρώο

Από τα πολυώνυμα στα μητρώα

Το $p(z) = \alpha_0 + \alpha_1\lambda + \cdots + \alpha_{n-1}\lambda^{n-1} + \lambda^n$ είναι το χ.π. των

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 & -\alpha_0 \\ 1 & 0 & \cdots & 0 & -\alpha_1 \\ 0 & 1 & 0 & \cdots & -\alpha_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 & -\alpha_{n-1} \end{pmatrix}$$ καθώς και του $A^\top$.

Ορισμός και χρήση

Το A ονομάζεται συνοδευτικό μητρώο (companion matrix) του πολυωνύμου p.

(υπολ. ριζών πολυωνύμου βαθμού n) $\equiv$ (υπολ. ιδιοτιμών συνοδευτικού μητρώου)

Επαληθεύστε ότι

1. ότι κάθε ρίζα ζ του p είναι ιδιοτιμή του A
2. ... με αντίστοιχο αριστερό ιδιοδιάνυσμα το $(1, \zeta, \zeta^2, \ldots, \zeta^{n-1})$.

Παραδείγματα

$p(\lambda) = \lambda^2 + 2\lambda + 1$ και

$$A = \begin{pmatrix} 0 & -1 \\ 1 & -2 \end{pmatrix} \Rightarrow \det(\lambda I - A) = \lambda(\lambda + 2) + 1 = \lambda^2 + 2\lambda + 1$$

$p(\lambda) = \lambda^3 + 3\lambda^2 + 2\lambda + 1$ και

$$A = \begin{pmatrix} 0 & 0 & -1 \\ 1 & 0 & -2 \\ 0 & 1 & -3 \end{pmatrix} \Rightarrow \det(\lambda I - A) = \lambda(\lambda(\lambda + 3) + 2) + 1 = \lambda^3 + 3\lambda^2 + 2\lambda + 1$$

Ιδιοδιανύσματα

$A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$, Βρήκαμε ότι $\lambda_1 = 1, \mu_1 = 1$, $\lambda_2 = 3, \mu_2 = 1$.

Θέτουμε $A(1) = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}$ και επιλύουμε $A(1)x = 0$:
Η τάξη $\mathrm{rank}(A(1)) = 1$, και βρίσκουμε από τις ειδικές λύσεις ότι μητρώο μηδενοχώρου για το $A(1)$ είναι το $\begin{pmatrix} -1 \\ 1 \end{pmatrix}$ Κανονικοποιούμε και επιλέγουμε ιδιοδιάνυσμα για το λ_1: $x_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 1 \end{pmatrix}$

Θέτουμε $A(3) = \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}$ και επιλύουμε $A(3)x = 0$: Η τάξη $\mathrm{rank}(A(3)) = 1$, και βρίσκουμε από τις ειδικές λύσεις ότι μητρώο μηδενοχώρου για το $A(3)$ είναι το $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$ Κανονικοποιούμε και επιλέγουμε ιδιοδιάνυσμα για το λ_2: $x_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Παρατηρήσεις I

Αν πολλαπλασιάσουμε $A(x_1, x_2) = (\lambda_1 x_1, \lambda_2 x_2)$

$$\begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} -\frac{1}{\sqrt{2}} & 3\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 3\frac{1}{\sqrt{2}} \end{pmatrix}$$

$$= \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

Δείτε ότι αν $X = (x_1, x_2)$ και $\mathrm{fl} = \mathrm{diag}(\lambda_1, \lambda_2)$, τότε

$$AX = X\Lambda$$

Το X είναι αντιστρέψιμο γιατί οι ιδιοτιμές είναι διαφορετικές (βλ. ▸ επόμενες διαφάνειες.) Είναι επίσης ορθογώνιο (λόγω συμμετρίας του A, θα το δούμε και αυτό αργότερα).

Παρατηρήσεις II

Ισχύει επομένως ότι

$$\begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 3 \end{pmatrix}$$

Δηλαδή

$$X^{-1}AX = X^{\top}AX = \Lambda.$$

Προσοχή

Ο μετασχηματισμός $A \to X^{-1}AX$ λέγεται μετασχηματισμός ομοιότητας του A με το X. Όταν το X είναι ορθογώνιο, όπως εδώ, χαρακτηρίζεται ως ορθογώνιος μετασχηματισμός ομοιότητας. Στην περίπτωση που εξετάζουμε, η εφαρμογή του μετασχηματισμού, μετατρέπει το A σε διαγώνιο μητρώο, που περιέχει τις ιδιοτιμές. Λέμε ότι το μητρώο είναι διαγωνιοποιήσιμο. Ένα μητρώο με n γραμμικά ανεξάρτητα ιδιοδιανύσματα διαγωνιοποιείται με το μητρώο των ιδιοδιανυσμάτων του.

Ιδιοδιανύσματα Ι

Στη συνέχεια συμβολίζουμε $A(\lambda) = \lambda I - A$.

$$A = \begin{pmatrix} 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \\ 2 & 0 & 2 & 0 \\ 0 & 2 & 0 & 2 \end{pmatrix}$$, Βρήκαμε ότι $\lambda_1 = 0, \mu_1 = 2, \lambda_2 = 4, \mu_2 = 2$.

Θέτουμε $A(0) = -A$ και επιλύουμε $-Ax = 0$:

Η τάξη $\mathrm{rank}(A(0)) = 2$, και βρίσκουμε από τις ειδικές λύσεις ότι μητρώο μηδενοχώρου για το $A(0)$ είναι το $\begin{pmatrix} -1 & 0 \\ 0 & -1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$ Κανονικοποιούμε και επιλέγουμε ιδιοδιανύσματα για το $\lambda_1 = 0$: $x_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, x_2 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ -1 \\ 0 \\ 1 \end{pmatrix}$

Ιδιοδιανύσματα II

Θέτουμε $A(4) = 4I - A$ και επιλύουμε $(4I - A)x = 0$:

Η τάξη $\mathrm{rank}(A(4)) = 2$, και βρίσκουμε από τις ειδικές λύσεις ότι μητρώο μηδενοχώρου για το $A(4)$ είναι το $\begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \\ 0 & 1 \end{pmatrix}$

Κανονικοποιούμε και επιλέγουμε ιδιοδιανύσματα για το $\lambda_2 = 4$: $x_3 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}$,

$x_4 = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 \\ 1 \\ 0 \\ 1 \end{pmatrix}$

Παρατηρήστε ότι τα ιδιοδιανύσματα είναι κάθετα μεταξύ τους.

Ιδιοδιανύσματα III

Για το $A = \begin{pmatrix} 1 & 2 \\ 0 & 1 \end{pmatrix}$. βρήκαμε ότι $\Lambda(A) = \{1\}, \lambda_1 = 1, \mu_1 = 2$.

Θέτουμε $A(1) = \begin{pmatrix} -1 & -1 \\ -1 & -1 \end{pmatrix}$ και επιλύουμε $A(1)x = 0$:

Η τάξη $\mathrm{rank}(A(1)) = 1$, επομένως η διάσταση του $\mathrm{null}(A(1)) = 1$.

Παρατηρούμε ότι $\mathrm{null}(A(1)) < \mu_1$.

Από τις ειδικές λύσεις βρίσκουμε ότι το μητρώο μηδενοχώρου για το $A(1)$ είναι το $\begin{pmatrix} 1 \\ 0 \end{pmatrix}$ Κανονικοποιούμε, οπότε το ιδιοδιάνυσμα για το λ_1 είναι $x_1 = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$

Προσοχή: Όλα τα ιδιοδιανύσματα του A για το $\lambda_1 = 1$ είναι πολλαπλάσια του x_1. Δεν υπάρχουν άλλα! Εφόσον η μόνη διακριτή ιδιοτιμή είναι το $\lambda_1 = 1$, αυτό είναι το μοναδικό ιδιοδιάνυσμα του A.

Γραμμική ανεξαρτησία ιδιοδιανυσμάτων

Πρόταση

Τα ιδιοδιανύσματα που αντιστοιχούν σε διαφορετικές ιδιοτιμές είναι γραμμικά ανεξάρτητα.

Απόδειξη[2] Έστω ότι $Ax_1 = \lambda_1 x_1$, $Ax_2 = \lambda_2 x_2$ όπου $\lambda_1 \neq \lambda_2$. Τότε αν υπάρχουν μη μηδενικά γ_1, γ_2 τ.ώ.

$$0 = \gamma_1 x_1 + \gamma_2 x_2 = \gamma_1 A x_1 + \gamma_2 A x_2 = \gamma_1 \lambda_1 x_1 + \gamma_2 \lambda_2 x_2.$$

Ισχύει επίσης ότι $0 = \lambda_2(\gamma_1 x_1 + \gamma_2 x_2) = \lambda_2 \gamma_1 x_1 + \lambda_2 \gamma_2 x_2$ επομένως

$$\gamma_1 \lambda_1 x_1 + \cancel{\gamma_2 \lambda_2 x_2} = \lambda_2 \gamma_1 x_1 + \cancel{\lambda_2 \gamma_2 x_2}$$

δηλ. $0 = \gamma_1 \lambda_1 x_1 - \lambda_2 \gamma_1 x_1$ επομένως $\lambda_1 = \lambda_2$, άρα άτοπο.

Πόρισμα

Αν ένα μητρώο $A \in \mathbb{C}^{n \times n}$ έχει n διαφορετικές ιδιοτιμές, τότε θα έχει n γραμμικά ανεξάρτητα ιδιοδιανύσματα. Κατά συνέπεια, τότε τα ιδιοδιανύσματα παρέχουν μία βάση (ιδιοβάση) με πολλές ευχάριστες ιδιότητες.

[2] Δείτε την πλήρη απόδειξη στο βιβλίο του Strang:

Διαγωνιοποίηση μητρώου I

Έστω ότι γνωρίζουμε n ιδιοζεύγη (λ_j, x_j) του $A \in \mathbb{C}^{n\times n}$. Τότε

$$Ax_1 = \lambda_1 x_1, Ax_2 = \lambda_2 x_2, \ldots, Ax_n = \lambda_n x_n$$

Συλλέγουμε και διατυπώνουμε με μητρώα:

$$A[x_1, x_2, \ldots, x_n] = [x_1 \ldots, x_n] \begin{pmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{pmatrix},$$

Επομένως

$$AX = X\Lambda,$$

όπου $\Lambda = \mathrm{diag}([\lambda_1, \ldots, \lambda_n])$ και $X = [x_1 \ldots, x_n]$.

Διαγωνιοποίηση μητρώου II

<u>ΠΡΟΣΕΞΤΕ</u> Αν το X είναι αντιστρέψιμο,

$$X^{-1}AX = \Lambda,$$

δηλ. χρησιμοποιώντας τα μητρώα X (με στήλες τα δεξιά ιδιοδιανύσματα) και X^{-1} (με στήλες του X^{-*} τα αριστερά ιδιοδιανύσματα) διαγωνιοποιήσαμε το A.

Κάθε μητρώο για το οποίο υπάρχουν n γραμμικά ανεξάρτητα ιδιοδιανύσματα χαρακτηρίζεται ως διαγωνιοποιήσιμο, ειδάλλως λέγεται μη διαγωνιοποιήσιμο.

Ιδιοδιανύσματα: Δεξιά και Αριστερά

Αν $A \in \mathbb{C}^{n \times n}$ και $\lambda \in \Lambda(A)$ τότε

$$\det(\lambda I - A) = 0$$

Αν συμβολίσουμε με r την τάξη του $\lambda I - A$,

$$r = \mathrm{rank}(\lambda I - A) \leq n - 1.$$

Αφού το μητρώο είναι τετραγωνικό, οι διαστάσεις του μηδενόχωρου και του αριστερού μηδενόχωρου του $A - \lambda I$ είναι ίσες επομένως

$$1 \leq \dim(\mathrm{null}(\lambda I - A)) = \dim(\mathrm{null}(\lambda I - A)^*) = n - r \leq n - 1$$

Ιδιοδιανύσματα είναι τα μέλη των μηδενοχώρων αυτών:

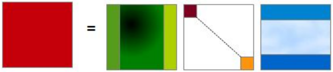
δεξιά ιδιοδιανύσματα για την ιδιοτιμή λ είναι τα διανύσματα $x \in \mathrm{null}(\lambda I - A)$ δηλ. τ.ώ. $Ax = \lambda x$.

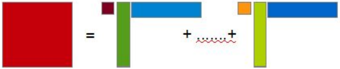
αριστερά ιδιοδιανύσματα για την ιδιοτιμή λ είναι τα διανύσματα $y \in \mathrm{null}(\lambda I - A)^*$, δηλ. τ.ώ. $y^* A = \lambda y^*$.

Ανάπτυγμα μητρώου συναρτήσει ιδιοτιμών και ιδιοδιανυσμάτων

Αν το μητρώο A διαθέτει n γ.α. ιδιοδιοδιανύσματα, τότε $A = X\Lambda X^{-1}$ όπου Λ το διαγώνιο μητρώο των ιδιοτιμών και X το μητρώο των ιδιοδιανυσμάτων.
Θυμηθείτε τη γραφή γινομένου μητρώων ως άθροισμα μητρώων πρώτης τάξης (στήλη του 1ου επί γραμμή του 2ου).
Θέτουμε για ευκολία $Y = (X^{-1})^*$, δηλ. Y είναι το συζυγές αντίστροφο του X, τότε μπορούμε να γράψουμε το A με βάση το φασματικό ανάπτυγμα.

$$\begin{aligned} A &= X\Lambda Y^* \\ &= \lambda_1 x_1 y_1^* + \lambda_2 x_2 y_2^* + \cdots + \lambda_n x_n y_n^* \end{aligned}$$





Ιδιόχωροι και γεωμετρική πολλαπλότητα ιδιοτιμής

Για κάθε διαφορετική ιδιοτιμή λ_j:

- ο μηδενόχωρος $\mathrm{null}(\lambda_j I - A)$ λέγεται ιδιόχωρος της ιδιοτιμής λ_j.
- Η διάσταση του ιδιόχωρου λέγεται γεωμετρική πολλαπλότητα της ιδιοτιμής λ_j.
- Για κάθε ιδιοτιμή $1 \leq$ γεωμ. πολλ/τα$(\lambda_j) \leq$ αλγεβρ. πολλ/τα.(λ_j)
- Κάθε ιδιοτιμή για την οποία

  $$1 < \text{αλγεβρ. πολλ/τα.} = \text{αλγεβρ. πολλ/τα.}$$

  αποκαλείται ημιαπλή.
- Αν κάποια ιδιοτιμή έχει γεωμετρική πολλαπλότητα μικρότερη της αλγεβρικής, δηλ.

  $$1 \leq \text{γεωμ. πολλ/τα}(\lambda_j) < \text{αλγεβρ. πολλ/τα.}(\lambda_j)$$

  χαρακτηρίζεται ως ελλειμματική[3]. Αν υπάρχει έστω και μια ελλειμματική, το μητρώο λέγεται ελλειμματικό.
- Τα ελλειμματικά μητρώα είναι μη διαγωνιοποιήσιμα.

[3] defective

αλγ. πολλ.	γεω. πολλ.	ΑΠ-ΓΠ	ιδιοτιμή	μητρώο
1	1	0	απλή	διαγωνιοποιήσιμο, αν
>1	>1	0	ημιαπλή	όλες οι ιδιοτιμές όπως εδώ
>1	>1	>0	ελλειμματική	ελλειμματικό (μη διαγωνιοποιήσιμο)

Παραδείγματα

Για τα παρακάτω τριγωνικά μητρώα (ιδιοτιμές στη διαγώνιο)

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 0 & 2 & 2 \\ 0 & 0 & 3 \end{pmatrix}, B = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix}, C = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}, D = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix},$$

A	*διαγωνιοποιήσιμο*	*διακριτές ιδιοτιμές*	$X = \begin{pmatrix} 1 & \frac{2\sqrt{5}}{5} & \frac{7\sqrt{69}}{69} \\ 0 & \frac{\sqrt{5}}{5} & \frac{4\sqrt{69}}{69} \\ 0 & 0 & \frac{2\sqrt{69}}{69} \end{pmatrix}$
B	*μη διαγωνιοποιήσιμο*	$B = \begin{pmatrix} I_2 & B_{1:2,2} \\ 0 & 1 \end{pmatrix}$	$X = \begin{pmatrix} 1 & 0 & -\frac{\sqrt{2}}{2} \\ 0 & 1 & -\frac{\sqrt{2}}{2} \\ 0 & 0 & 0 \end{pmatrix}, \det(X) = 0$
C	*μη διαγωνιοπ.*	*μορφή Jordan*	
D	*μη διαγωνιοπ.*	$D = \begin{pmatrix} D_{1:2,1:2} & D_{1:2,2} \\ 0 & 1 \end{pmatrix}$	
$A_{1:2,1:2}$	*διαγωνιοποιήσιμο*	*διακριτές ιδιοτιμές*	

Περί διαγωνιοποιήσιμων μητρώων

Κάθε μητρώο είναι διαγωνιοποιήσιμο ή μη διαγωνιοποιήσιμο (ελλειμματικό).

Ερώτημα Πότε είναι ένα μητρώο διαγωνιοποιήσιμο;

- Αν και μόνον αν υπάρχουν n γ.α. ιδιοδιανύσματα.
- Για παράδειγμα, όταν υπάρχουν n διαφορετικές ιδιοτιμές.

Προσοχή: Αν υπάρχουν επαναλαμβανόμενες ιδιοτιμές δεν εξασφαλίζεται ούτε η διαγωνιοποιησιμότητα ούτε η μη διαγωνιοποιησιμότητα. Εξαρτάται από το μητρώο, επομένως απαιτείται περισσότερη διερεύνηση.

- Προσέξτε ότι αυτή η κατηγοριοποίηση αφορά την ύπαρξη αντιστρέψιμου X ώστε $X^{-1}AX$ να είναι διαγώνιο. Η αντιστρεψιμότητα ή μη του A αφορά μόνον τοA (προφανώς!)
- Έστω $A(\alpha,\beta) = \begin{pmatrix} \alpha & 1 \\ 0 & \beta \end{pmatrix}$. Τότε
  - Αν $\alpha = 1, \beta = 2$, τότε το A είναι αντιστρέψιμο και διαγωνοποιήσιμο.
  - Αν $\alpha = 1, \beta = 1$, τότε το A είναι αντιστρέψιμο και μη διαγωνοποιήσιμο.
  - Αν $\alpha = 1, \beta = 0$, τότε το A είναι μη αντιστρέψιμο και διαγωνοποιήσιμο.
  - Αν $\alpha = 0, \beta = 0$, τότε το A είναι μη αντιστρέψιμο και μη διαγωνοποιήσιμο.
- (Ευτυχώς) στο (τεράστιο) σύνολο των μητρώων, τα μη αντιστρέψιμα μητρώα, όπως και τα μη διαγωνιοποιήσιμα, είναι "λίγα": *Ένα μητρώο με τυχαίες τιμές είναι συνήθως αντιστρέψιμο και διαγωνιοποιήσιμο!*

Πολλαπλές ιδιοτιμές

- Σε κάθε απλή ιδιοτιμή (αλγ. πολλ/τας ίσης με 1) αντιστοιχούν ένα δεξιό και ένα αριστερό ιδιοδιάνυσμα
- Τι γίνεται όταν μια ιδιοτιμή, π.χ. λ_j, έχει αλγ. πολλ/τα, έστω $\mu_j > 1$;
- Το κρίσιμο ερώτημα είναι κατά πόσον υπάρχουν μ_j γραμμικά ανεξάρτητα ιδιοδιανύσματα[4].
- Αν για κάθε διακριτή ιδιοτιμή, η διάσταση του αντίστοιχου μηδενόχωρου είναι ίση με την αλγεβρική πολλαπλότητά της, τότε υπάρχουν συνολικά n γραμμικά ανεξάρτητα ιδιοδιανύσματα και το μητρώο είναι διαγωνιοποιήσιμο.
- Κατά συνέπεια, τα ιδιοδιανύσματα αποτελούν βάση για όλον το διανυσματικό χώρο ($\mathbb{R}^n$ ή $\mathbb{C}^n$).
- Η βάση αυτή (ιδιοβάση) μπορεί να είναι προτιμότερη της τυπικής βάσης $\{e_1, \ldots, e_n\}$ καθώς ως προς κάθε στοιχείο της βάσης αυτής, ο πολλαπλασιασμός με το A και άλλες πράξεις γίνονται πολύ εύκολα.

[4] Στη συνέχεια θα εννοούμε τα δεξιά ιδιοδιανύσματα, αλλά το ίδιο ισχύει για τα αριστερά.

Ιδιοτιμές και ιδιοδιανύσματα δυνάμεων και πολυωνύμων μητρώου

- Για κάθε A, αν $Ax = \lambda x$ τότε $A^k x = \lambda^k x$. Αυτό φαίνεται εύκολα, π.χ. $A(Ax) = A(\lambda x) = \lambda Ax = \lambda^2 x$ κ.ο.κ..
- ομοίως $\gamma A^j x + \delta A^k x = \gamma \lambda^j x + \delta \lambda^k x$
- άρα αν οι ιδιοτιμές του A είναι $\{\lambda_1, \ldots, \lambda_n\}$ τότε οι ιδιοτιμές του πολυωνύμου $q(A) = \gamma_s A^s + \cdots + \gamma_0 I$, είναι

$$\begin{array}{rcl} q(\lambda_1) & = & \gamma_s \lambda_1^s + \cdots + + \gamma_1 \lambda_1 + \gamma_0 I \\ \vdots & \vdots & \vdots \\ q(\lambda_n) & = & \gamma_s \lambda_n^s + \cdots + \gamma_1 \lambda_n + \gamma_0 I \end{array}$$

δηλ. $\{q(\lambda_1), \ldots., q(\lambda_n)\}$.

Συμπεράσματα

- Επομένως αν γνωρίζουμε τις ιδιοτιμές του A γνωρίζουμε και τις ιδιοτιμές οποιουδήποτε πολυωνύμου στο A.
- Τα ιδιοδιανύσματα του $q(A)$ είναι ίδια με τα ιδιοδιανύσματα του A.
- Τα αποτελέσματα επεκτείνονται και σε γενικότερες συναρτήσεις μητρώου, π.χ. Αν $X^{-1}AX = \Lambda$ τότε $X^{-1}\exp(A)X = \exp(\Lambda)$

Ανακεφαλαίωση

Βασικός ορισμός

Για ένα μητρώο $A \in \mathbb{R}^{n\times n}$ και μεταβλητή λ, η ορίζουσα $p(\lambda) = \det(\lambda I - A)$ είναι πολυώνυμο βαθμού n. Λέγεται χαρακτηριστικό πολυώνυμο και αν το γράψουμε

$$p(\lambda) \quad = \quad \lambda^n + \gamma_{n-1}\lambda^{n-1} + \cdots + \gamma_1\lambda + \gamma_0$$

τότε $\gamma_{n-1} = -\mathrm{trace}(A)$ και $\gamma_0 = (-1)^n \det(A)$.

Προσοχή:

- Το πολυώνυμο έχει ακριβώς n ρίζες $\lambda_1, \cdots \lambda_n \in \mathbb{C}^n$ που αποκαλούμε **ιδιοτιμές** του A.
- Για κάθε ιδιοτιμή, λ, ο μηδενόχωρος $\mathrm{null}(\lambda I - A)$ λέγεται **ιδιόχωρος** και έχει διάσταση τουλάχιστον 1. Κάθε διάνυσμα βάσης του χώρου ονομάζεται **ιδιοδιάνυσμα**.
- Ισχύει ότι $p(A) = 0$ **(θεώρημα Cayley-Hamilton)**.
- τα ιδιοζεύγη (ιδιοτιμές και ιδιοδιανύσματα) μητρώων παίζουν πολύ σημαντικό ρόλο σε πληθώρα εφαρμογών.

Ιδιοδιανύσματα και ιδιόχωροι

- Γενικά, δεξιός (αριστερός) ιδιόχωρος του A λέγεται ο υπόχωρος $\mathrm{null}(A-\lambda I)$ (αντίστοιχα, ο αριστερός μηδενόχωρος $\mathrm{null}(A^*-\lambda I)$).
- ΠΡΟΣΕΞΤΕ: Επειδή το μητρώο είναι τετραγωνικό, για κάθε ιδιοτιμή λ, ο δεξιός και αριστερός μηδενόχωρος έχουν ίδια διάσταση, $n-r(\lambda)$, όπου $r(\lambda)$ είναι η τάξη του $A-\lambda I$.
- Η πολλαπλότητα της κάθε ρίζας του $p(\lambda)$ ονομάζεται αλγεβρική πολλαπλότητα της ιδιοτιμής λ.
- $\forall \gamma \neq 0,\ Ax=\lambda x \Leftrightarrow (\gamma A)x=(\gamma\lambda)x$
- Η διάσταση $(n-r(\lambda))$ των μηδενόχωρων λέγεται γεωμετρική πολλαπλότητα της ιδιοτιμής.
- Ο (αριστερός) μηδενόχωρος για κάθε ιδιοτιμή καλείται (αριστερος) ιδιόχωρος της ιδιοτιμής.

Ομοιότητα μητρώων

(Strang, 6.6)

Ορισμοί

Έστω $A \in \mathbb{C}^{n\times n}$. Δίνεται αντιστρέψιμο $X \in \mathbb{C}^{n\times n}$ και $B = X^{-1}AX$. Τότε $\Lambda(B) = \Lambda(A)$. Τα μητρώα A και B αποκαλούνται όμοια και έχουν ακριβώς τις ίδιες ιδιοτιμές. Η απεικόνιση

$$A \Rightarrow X^{-1}AX$$

λέγεται μετασχηματισμός ομοιότητας.

- Ένα μητρώο είναι διαγωνιοποιήσιμο αν είναι όμοιο με ένα διαγώνιο μητρώο (των ιδιοτιμών του!)
- Διαγωνιοποιήσιμα είναι όσα μητρώα διαθέτουν ακριβώς n γραμμικά ανεξάρτητα ιδιοδιανύσματα.
- Αν τα ιδιοδιανύσματα είναι πραγματικά και κάθετα μεταξύ τους, τότε $P^\top P = I$ άρα το μητρώο P είναι ορθογώνιο και ισχύει ότι $P^\top AP = \Lambda$.

Απεικόνιση στο μιγαδικό επίπεδο

Live παραδείγματα (MATLAB) στην τάξη.

1. `A=rand(n);` μητρώο με τυχαίες τιμές στο $(0,1)$ `A=randn(n);` μητρώο με τιμές από κανονική κατανομή $N(0,1)$.
2. `A=sprand(n,p);` αραιό μητρώο μη μηδενικά στοιχεία σε τυχαίες θέσεις και τυχαίες τιμές στο $(0,1)$ `A=sprandn(n,p);` αραιό μητρώο μη μηδενικά στοιχεία σε τυχαίες θέσεις και τυχαίες τιμές στο $N(0,1)$
3. `A=gallery('cauchy',n);` μητρώο με τιμές $\frac{1}{i+j}$.
4. `A=gallery('grcar',n);` μητρώο του Grcar
5. `A=gallery('rando',n);` Μητρώο με τιμές 0 ή 1 σε τυχαίες θέσεις
6. `A=gallery('poisson',n);` Μητρώο Poisson (μπλοκ τριδιαγώνιο)
7. `A=randn(n)+ sqrt(-1)*randn(n);`
8. `A=rand(n)+ sqrt(-1)*rand(n);`
9. `A=rand(n)+ sqrt(-1)*rand(n);A=A+A';` Ερμιτιανό
10. `A=rand(n)+ sqrt(-1)*rand(n);A=A+A.';` μιγαδικό συμμετρικό

και στη συνέχεια `e=eig(full(A));plot(real(e),imag(e),'o')`.

Το φάσμα κείται στο μιγαδικό επίπεδο. Όταν το μητρώο είναι πραγματικό, η εικόνα είναι συμμετρική ως προς τον άξονα των πραγματικών. Όταν το μητρώο είναι συμμετρικό πραγματικό, οι ιδιοτιμές είναι πραγματικές.

Φασματική ακτίνα μητρώου

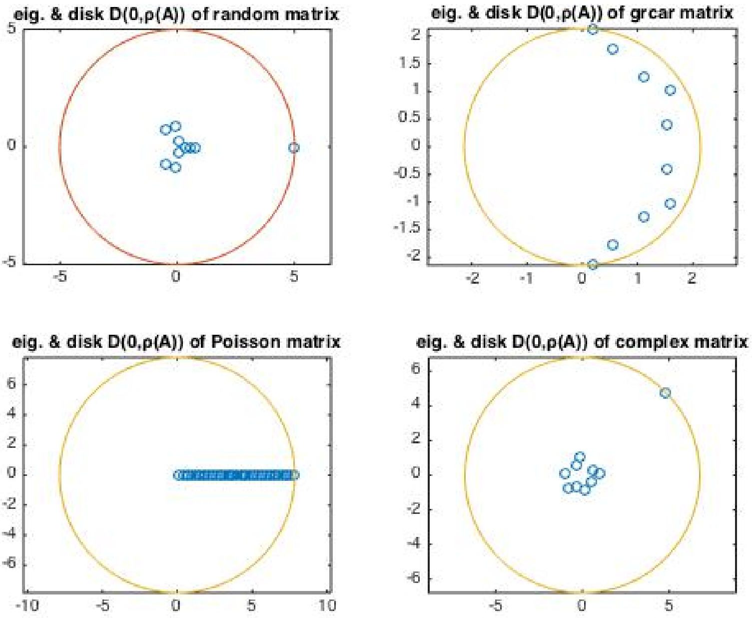
Ορισμός

Ονομάζουμε φασματική ακτίνα μητρώου την τιμή

$$\rho(A) := \max\{|\lambda| : \lambda \text{ ιδιοτιμή του } A\}$$

Επομένως, όλες οι ιδιοτιμές του A (το φάσμα) κείνται στο δίσκο με κέντρο το 0 και ακτίνα $\rho(A)$:

$$\Lambda(A) \subseteq \mathcal{D}(0, \rho(A))$$


eig. & disk D(0,ρ(A)) of random matrix
5
0
-5
-5
0
5
eig. & disk D(0,ρ(A)) of grcar matrix
2
1.5
1
0.5
0
-0.5
-1
-1.5
-2
-2
-1
0
1
2
eig. & disk D(0,ρ(A)) of Poisson matrix
6
4
2
0
-2
-4
-6
-10
-5
0
5
10
eig. & disk D(0,ρ(A)) of complex matrix
6
4
2
0
-2
-4
-6
-5
0
5

Διαχείριση δυνάμεων μητρώου μέσω διαγωνιοποίησης

Έστω $A \in \mathbb{R}^{n\times n}$ και ότι γνωρίζουμε X, Λ όπως πριν, δηλ. $X^{-1}AX = \Lambda$.

$$\begin{aligned} A &= X\Lambda X^{-1} \\ A^k &= (X\Lambda X)^k = (X\Lambda X^{-1})(X\Lambda X^{-1})\cdots(X\Lambda X^{-1}) \\ &= X\Lambda(X^{-1}X)\Lambda(X^{-1}X)\Lambda\cdots(X^{-1}X)\Lambda X^{-1} \\ &= X\Lambda^k X^{-1}. \end{aligned}$$

Παρατήρηση: Ο υπολογισμός του A^k ανάγεται στον εύκολο υπολογισμό του Λ^k και σε δύο επιμέρους πολλαπλασιασμούς μητρώων, π.χ. $X\Lambda^k \to (X\Lambda^k)X^{-1}$.

Πέραν του υπολογισμού του A^k, η διάσπαση του μητρώου σε $A = X\Lambda X^{-1}$ διευκολύνει στην διερεύνηση του παρακάτω ζητήματος.

Πώς συμπεριφέρεται το A^k για μεγάλες τιμές του k;

- Υπάρχει κάτι το ιδιαίτερο;
- Υπάρχει όριο, δηλ. κάποιο B τ.ω. $\lim_{k\to\infty} A^k = B$ υπό την έννοια ότι $\forall \varepsilon > 0, \exists \hat{k}$ τ.ώ. $\|A^k - B\| < \varepsilon, \forall k > \hat{k}$;
- Ποιό είναι αυτό;

Αποδεικνύεται ότι: Αν $A \in \mathbb{C}^{n\times n}$ τότε το όριο $\lim_{k\to\infty} A^k$ υπάρχει αν και μόνον αν

- $\rho(A) < 1$, είτε
- $\rho(A) = 1$ και α) το $\lambda = 1$ είναι η μόνη ιδιοτιμή στον μοναδιαίο κύκλο και β) ως ιδιοτιμή είναι ημιαπλή.

Παράδειγμα I

$$A = \begin{pmatrix} \frac{2}{3} & -\frac{1}{3} & 0 \\ -\frac{1}{3} & \frac{2}{3} & 0 \\ -\frac{1}{3} & \frac{1}{3} & \frac{1}{3} \end{pmatrix}, \Lambda = \begin{pmatrix} \frac{1}{3} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & \frac{1}{3} \end{pmatrix}, X = \begin{pmatrix} 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} \\ 1 & \frac{\sqrt{3}}{3} & 0 \end{pmatrix}$$

$$A^k = \begin{pmatrix} 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} \\ 1 & \frac{\sqrt{3}}{3} & 0 \end{pmatrix} \begin{pmatrix} 3^{-k} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3^{-k} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 1 \\ -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \end{pmatrix}$$

Εκτελώντας τον πολλαπλασιασμό για οποιοδήποτε k επιθυμούμε, προκύπτει το ζητούμενο, αφού στρογγυλέψουμε

$$A^4 = \begin{pmatrix} 0.5062 & -0.4938 & 0 \\ -0.4938 & 0.5062 & 0 \\ -0.4938 & 0.4938 & 0.0123 \end{pmatrix}, A^{10} = \begin{pmatrix} 0.5000 & -0.5000 & 0 \\ -0.5000 & 0.5000 & 0 \\ -0.5000 & 0.5000 & 0.0000 \end{pmatrix}$$

Οι τιμές έχουν υποστεί στρογγύλευση.

Παράδειγμα II

Καθώς $k \to \infty$,

$$A^k = \begin{pmatrix} 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} \\ 1 & \frac{\sqrt{3}}{3} & 0 \end{pmatrix} \begin{pmatrix} 3^{-k} & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 3^{-k} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 1 \\ -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \end{pmatrix}$$

οι όροι στο διαγώνιο μητρώο Λ^k που αντιστοιχούν σε ιδιοτιμές $|\lambda| < 1$ τείνουν στο 0, επομένως

$$\begin{aligned} \lim_{k\to\infty} A^k &= \begin{pmatrix} 0 & -\frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} \\ 0 & \frac{\sqrt{3}}{3} & \frac{\sqrt{2}}{2} \\ 1 & \frac{\sqrt{3}}{3} & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 1 \\ -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & 0 \\ \frac{\sqrt{2}}{2} & \frac{\sqrt{2}}{2} & 0 \end{pmatrix} \\ &= \begin{pmatrix} -\frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \\ \frac{\sqrt{3}}{3} \end{pmatrix} \begin{pmatrix} -\frac{\sqrt{3}}{2} & \frac{\sqrt{3}}{2} & 0 \end{pmatrix} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \\ -\frac{1}{2} & \frac{1}{2} & 0 \end{pmatrix} \end{aligned}$$

Απόδειξη θεωρήματος Cayley-Hamilton

Για διαγωνιοποιήσιμα μητρώα

Το χ.π. είναι το $p(\lambda) = \det(\lambda I - A)$. Τότε σύμφωνα με τα προηγούμενα, οι ιδιοτιμές του $p(A)$ θα είναι

$$\{p(\lambda_1), \cdots, p(\lambda_n)\}$$

Όμως καθώς τα λ_j εξ ορισμού είναι ρίζες του χ.π. τότε $p(\lambda_j) = 0$ επομένως $p(\Lambda) = \mathrm{diag}[p(\lambda_1), \ldots, p(\lambda_n)] = 0$, επομένως

$$p(A) = Xp(\Lambda)X^{-1} = X0X = 0$$

Συναρτήσεις μητρώων

Τα παραπάνω χρησιμοποιούνται για να ορίσουμε τυπικά και γενικές συναρτήσεις μητρώων.

- η συνάρτηση ενός διαγώνιου μητρώου, έστω $D = \mathrm{diag}[\delta_1, \cdots, \delta_n]$, είναι το διαγώνιο μητρώο που έχει στη διαγώνιο τις τιμές που παίρνει η συνάρτηση στα στοιχεία της διαγωνίου:

$$f(D) = \mathrm{diag}[f(\delta_1), \ldots, f(\delta_n)]$$

Δοθείσης συνάρτησης f, για κάθε διαγωνιοποιήσιμο μητρώο A για το οποίο η συνάρτηση f υπάρχει σε κάθε ιδιοτιμή του μητρώου (δηλ. ορίζονται οι τιμές $f(\lambda_j)$ για κάθε ιδιοτιμή λ_j του A), η συνάρτηση μητρώου $f(A)$ ορίζεται ως εξής:

$$f(A) \quad = \quad X f(\Lambda) X^{-1} = X \begin{pmatrix} f(\lambda_1) & & \\ & \ddots & \\ & & f(\lambda_n) \end{pmatrix} X^{-1}$$

όπου Λ, X είναι τα μητρώα των ιδιοτιμών και ιδιοδιανυσμάτων του A.

Συμπέρασμα: Αν γνωρίζουμε τις ιδιοτιμές, το X (ιδιοδιανύσματα) και το X^{-1}, η συνάρτηση μητρώου ανάγεται στον υπολογισμό της συνάρτησης για κάθε ιδιοτιμή και δύο πολλαπλασιασμούς μητρώων. (Διαβάστε προαιρετικά τις σελ. 390-392 του Strang)

Παραδείγματα (Strang, σελ. 382)

Ελέγξτε πώς συμπεριφέρονται οι δυνάμεις των παρακάτω μητρώων:

$$A = \begin{pmatrix} 3 & 2 \\ 1 & 4 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 2 & 0 \\ 0 & 5 \end{pmatrix}$$

,

$$B = \begin{pmatrix} 3 & 2 \\ -5 & -3 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \mathrm{i} & 0 \\ 0 & \mathrm{i} \end{pmatrix}$$

$$C = \begin{pmatrix} 5 & 7 \\ -3 & -4 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \frac{1}{2} + \frac{\sqrt{3}\mathrm{i}}{2} & 0 \\ 0 & \frac{1}{2} - \frac{\sqrt{3}\mathrm{i}}{2} \end{pmatrix}$$

$$D = \begin{pmatrix} 5 & 6.9 \\ -3 & -4 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} \frac{1}{2} + \frac{3\sqrt{5}\mathrm{i}}{10} & 0 \\ 0 & \frac{1}{2} - \frac{3\sqrt{5}\mathrm{i}}{10} \end{pmatrix}$$

$$E = \begin{pmatrix} 0.999 & 1000 \\ 0 & 0.998 \end{pmatrix}, \quad \Lambda = \begin{pmatrix} 0.999 & 0 \\ 0 & 0.998 \end{pmatrix}$$

Μέθοδος δύναμης

απλή μέθοδος για ΓΙΓΑΝΤΙΑΙΑ ΠΡΟΒΛΗΜΑΤΑ

Αν η μέγιστη σε απόλυτη τιμή ιδιοτιμή, έστω $\lambda_{\max}$, ενός $A \in \mathbb{R}^{n\times n}$ είναι μοναδική (άρα πραγματική!), με αντίστοιχο ιδιοδιάνυσμα q_1, και έστω $x \in \mathbb{R}^n$ τυχαίο, τ.ώ. $x^\top q_1 \neq 0$, τότε η ακολουθία των διανυσματων $\{Ax, A^2x, \ldots, A^kx, \ldots\}$ τείνει σε διάνυσμα του ιδιόχωρου του $\lambda_{\max}$.

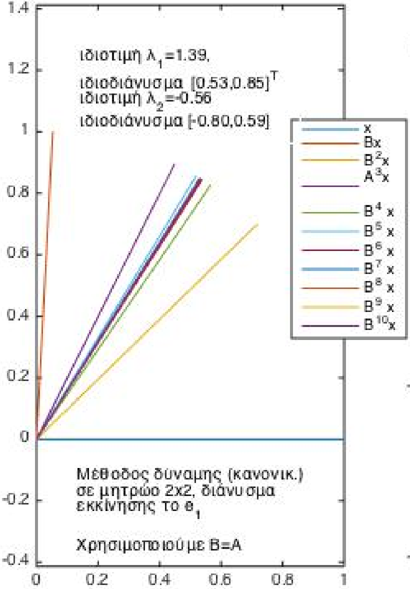
Σε περίπτωση που η μέγιστη σε απόλυτη τιμή είναι μοναδική, τότε τείνει σε διάνυσμα συγγραμμικό του q_1. (προσεγγίζεται το μέγιστο ιδιοδιάνυσμα!)

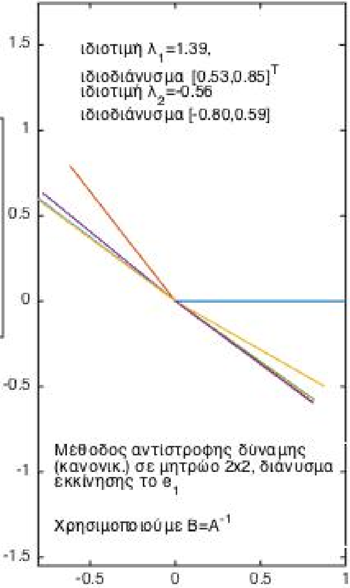
Επίσης για μεγάλα k, το παρακάτω πηλίκο,

$$\left(\frac{x^\top A^k x}{x^\top x}\right)^{1/k} \approx \lambda_{\max} \quad \text{προσεγγίζεται η μέγιστη ιδιοτιμή!}$$

Ορισμός Για κάθε A, x, το $\frac{x^\top Ax}{x^\top x}$ ονομάζεται πηλίκο Rayleigh.

Οπτικοποίηση μεθόδου δύναμης





Μέθοδος δύναμης και η Google

- Φημολογείται ότι η μέθοδος χρησιμοποιήθηκε από τους Brin & Page για τη βαθμολόγηση ιστοσελίδων.
- Θυμηθείτε τον τίτλο της εργασίας που δείξαμε στη διαφάνεια 11 της προηγούμενης διάλεξης (9):

  "The \$25,000,000,000 Eigenvector: The Linear Algebra behind Google"

- Το διάνυσμα PageRank είναι το ιδιοδιάνυσμα που αντιστοιχεί στη μέγιστη ιδιοτιμή του μητρώου της Google για το Web.
- Η μέγιστη ιδιοτιμή είναι 1 επειδή το μητρώο είναι στοχαστικό κατά στήλες (κάθε στήλη έχει άθροισμα 1 οπότε $e^\top A = e^\top$.)