Positive Definiteness

Lecture 12

Zhang Ce

Department of Electrical and Electronic Engineering
Southern University of Science and Technology

2022.12.20

Table of Contents

Table of Contents

Last Lecture, We Discuss...

Two parts in last lecture:

1. Similar Transformation
   definition; properties; geometric interpretations; tests for similarity; Schur's lemma; normal matrices; Jordan form
2. Topic: Exercise Problems for Chapter 5
   eigenvalues of A^T; $I - \alpha\alpha^T$; $I - uv^T$; rank; determinant; $A^2 = A$; $I + iH$; similarity of $A + A^T, A + A^{-1}$, $B + B^T, B + B^{-1}$

Properties for Similar Matrices

Similar matrices A and B have the following properties:

- A has the same eigenvalues as B.
- A has the same number of independent eigenvectors as B. (Difference: M^{-1})
- $\det A = \det B$, $trace(A) = trace(B)$.
- $rank(A) = rank(B)$. (who can give me a translation?)
- A and B have the same characteristic polynomial.

If A and B have the same characteristic polynomial (same eigenvalues), they are not always similar.

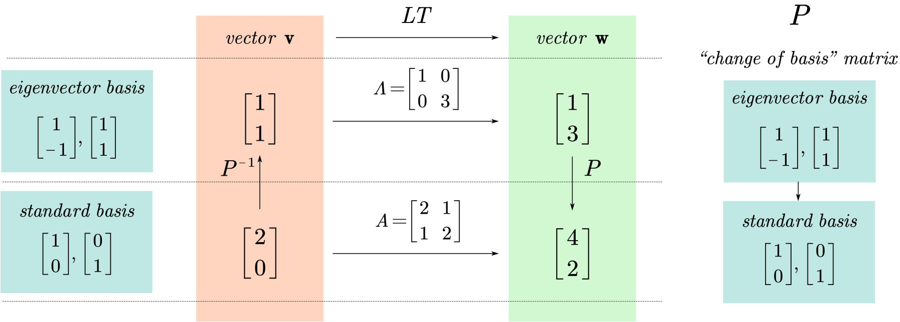
$$A = \begin{bmatrix} 5 & 3 \\ 0 & 5 \end{bmatrix}, B = \begin{bmatrix} 5 & 0 \\ 0 & 5 \end{bmatrix}$$

Those properties are necessary, not sufficient (Even if you find 2 matrices that can satisfy all of them, we can not say they are similar).

Similarity Transformation: Change of Basis

Recall: matrix diagonalization $A = P\Lambda P^{-1}$, we can say $A \backsim \Lambda$.

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



Essence: change basis and simplify the linear transformation matrix.

Normal Matrices

Theorem

For a matrix of degree n, there exists a unitary matrix U of degree n such that $U^{-1}AU = T$ is triangular. The eigenvalues of A appear along the diagonal of the similar matrix T.

For some matrices, $T = \Lambda$. For that case, the matrices are called normal.

Normal matrices contain:

- Real symmetric; Hermitian. They have all real eigenvalues.
- Real skew-symmetric; skew-Hermitian. They have all imaginary eigenvalues (or zero!).
- Orthogonal; unitary. They have eigenvalues $|\lambda| = 1$.

For normal matrices, $NN^H = N^H N$. All the 3 kinds satisfy this condition.

The Jordan Form

For non-diagonalizable matrices A and B, how to determine whether they are similar? Remind that the 5 properties for similar matrices are not sufficient.

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

For the example above, those matrices have the same eigenvalues ($\lambda = 2$ with algebraic multiplicity 4), the same number of independent eigenvectors (2 in this example). But they are not similar.

Equivalent Property for Similarity:

A and B share the same Jordan blocks.

If you want to know how to find the Jordan form for non-diagonalizable matrices, choose *MA109: Linear Algebra II* (but be careful!).

Table of Contents

Properties of Real Symmetric Matrices

Real symmetric matrix is one of the most important applications in our daily life. What properties can you get from $A = A^T$? List as much as possible.

1. $A = LDL^T$ factorization.
   All real symmetric matrices can be decomposed to $A = LDL^T$ without changing the nullspace (solution). In this factorization, we are interested in pivots.
2. $A = Q\Lambda Q^T$ factorization.
   All real symmetric matrices can be decomposed to $A = Q\Lambda Q^T$ without changing the eigenvalues. In this factorization, we are interested in eigenvalues and eigenvectors.
3. Eigenvalues of A are all real, eigenvectors of A can be chosen all orthogonal.

Make sure you are familiar with $A = LDL^T$ and $A = Q\Lambda Q^T$ factorization, they are important in this chapter.

All real symmetric matrices can be decomposed to $A = Q\Lambda Q^T$ without changing the eigenvalues. By further matrix multiplication, we can get:

$$A = \begin{bmatrix} | & | & & | \\ | & | & & | \\ q_1 & q_2 & \cdots & q_n \\ | & | & & | \\ | & | & & | \end{bmatrix} \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & \\ & & \ddots & \\ & & & \lambda_n \end{bmatrix} \begin{bmatrix} - & - & q_1 & - & - \\ - & - & q_2 & - & - \\ & & \vdots & & \\ - & - & q_n & - & - \end{bmatrix}$$

$$A = Q\Lambda Q^T = \lambda_1 q_1 {q_1}^T + \lambda_2 q_2 {q_2}^T + \cdots + \lambda_n q_n {q_n}^T$$

Notice that q vectors are orthonormal, that indicates the linear transformation represented by matrix A can be decomposed to a series of projection onto lines matrix, with those lines orthogonal.

We are changing the bases and the bases are just like a new coordinate system.

Eigenvalues of Real Symmetric Matrices

We know that all the eigenvalues of A are real, a natural question to ask is, how to determine how many eigenvalues are positive, while how many eigenvalues are negative?

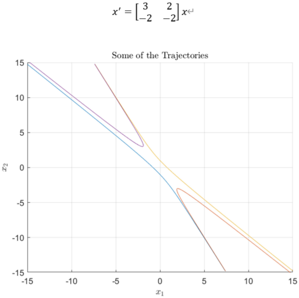
A useful property is that: number of positive pivots equals number of positive eigenvalues! (The Law of Inertia)

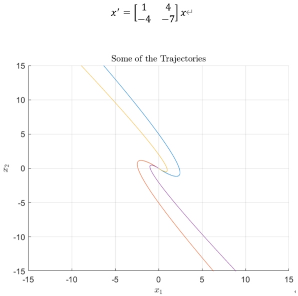
So, if you have a 80×80 matrix, and it has 32 positive pivots and 48 negative pivots, then it will also have 32 positive eigenvalues and 48 negative eigenvalues.

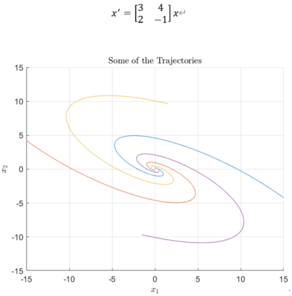
Given a 80×80 matrix, how do you find the exact value of all eigenvalues? Reminder:

- Humans are stupid when dealing with such a high-scale computation task, our design is only for computers. (think you are the programmer of MATLAB)
- Don't try to compute a determinant of 80×80 matrix, and don't try to solve a 80-degree equation!
- Gaussian elimination is efficient.

Why we are interested in signs of the eigenvalues? It has many applications in our daily life. For example, if a system is driven by a linear ordinary differential equation system, than the stability of the system is determined by the signs of eigenvalues.







If all the eigenvalues are negative, all the solution curve will go to the origin as time increases.

Definition of Positive Definiteness

We give 2 definitions here in different perspectives:

- Algebra: All the eigenvalues corresponding to this matrix are positive.
- Geometry: $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} > 0$ for all nonzero real vector x.

What is $\mathbf{x}^T A\mathbf{x}$? Do matrix multiplication in 2-dimensional:

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = a{x_1}^2 + 2bx_1x_2 + c{x_2}^2$$

Can you have a direct imagination of $\mathbf{x}^T A\mathbf{x}$? If the matrix A is positive definite, it is like a bowl. Imagine the plane $\mathbf{x}^T A\mathbf{x} = 1$, you can get an ellipse!

A website for fun: `https://www.wolframalpha.com/`

Try with example:

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2{x_1}^2 + 4x_1x_2 + 5{x_2}^2$$

Tests for Positive Definiteness

Tests for Positive Definiteness:

1. $f(\mathbf{x}) = \mathbf{x}^T A\mathbf{x} > 0$ for all nonzero real vector x.
2. All the eigenvalues of A satisfy $\lambda_i > 0$.
3. All the leading submatrices A_k have positive determinants.
4. All the pivots (without row exchanges) satisfy $d_k > 0$.

Test 2 and 4 are easy to understand, how about test 3? In my view, determinant combines pivots and eigenvalues together.

Although we have 4 tests, actually test 2, 3, 4 are showing the same thing.

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2{x_1}^2 + 4x_1x_2 + 5{x_2}^2$$

Please use the four tests to show that matrix A is positive definite. (You may need to verify the first test by completing the squares.)

Tests for Positive Definiteness

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2{x_1}^2 + 4x_1x_2 + 5{x_2}^2$$

Let's start with test 2:

$$\det\left(A - \lambda I\right) = \begin{vmatrix} 2-\lambda & 2 \\ 2 & 5-\lambda \end{vmatrix} = (\lambda - 1)(\lambda - 6) = 0$$

Eigenvalues are $\lambda_1 = 1, \lambda_2 = 6$, all greater than zero.

Test 3:

$$|2| = 2 > 0, \begin{vmatrix} 2 & 2 \\ 2 & 5 \end{vmatrix} = 6 > 0$$

Test 4:

$$\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \to \begin{bmatrix} \color{red}{2} & 2 \\ 0 & \color{red}{3} \end{bmatrix}$$

Tests for Positive Definiteness

$$\mathbf{x}^T A\mathbf{x} = \begin{bmatrix} x_1 & x_2 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = 2{x_1}^2 + 4x_1x_2 + 5{x_2}^2$$

Finally, check with test 1. Typically, we finish this by completing the squares.

$$2{x_1}^2 + 4x_1x_2 + 5{x_2}^2 = 2{x_1}^2 + 4x_1x_2 + 2{x_2}^2 + 3{x_2}^2 = 2\left(x_1 + x_2\right)^2 + 3{x_2}^2 > 0$$

This is a simple example, but in most of the cases, we cannot complete the squares by observation. There is actually a linear algebra way to do this! Recall your knowledge of LDL^T factorization.

$$\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Can observe that LDL^T can guide you how to complete the squares. Why?

Linear Algebra Way for Completing Squares

$$2{x_1}^2 + 4x_1x_2 + 5{x_2}^2 = 2\left(x_1 + x_2\right)^2 + 3{x_2}^2 > 0$$

$$\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 0 & 3 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}$$

Change $A = LDL^T$ in:

$$x^TAx = x^TLDL^Tx = \left(L^Tx\right)^T D \left(L^Tx\right)$$

Adopt change of variable $y = L^Tx$:

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{cases} y_1 = x_1 + x_2 \\ y_2 = x_2 \end{cases}$$

$$x^TAx = y^TDy$$

The pivots are the coefficients of the squares. In this example, 2 and 3.

Linear Algebra Way for Completing Squares

You also have $A = Q\Lambda Q^T$ factorization!

Change $A = Q\Lambda Q^T$ in:

$$x^T A x = x^T Q \Lambda Q^T x = \left(Q^T x\right)^T D \left(Q^T x\right)$$

Let's have a try!

$$\begin{bmatrix} 2 & 2 \\ 2 & 5 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \\ \frac{-1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 6 \end{bmatrix} \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix}$$

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}, \begin{cases} y_1 = \frac{2}{\sqrt{5}} x_1 - \frac{1}{\sqrt{5}} x_2 \\ y_2 = \frac{1}{\sqrt{5}} x_1 + \frac{2}{\sqrt{5}} x_2 \end{cases}$$

$$2{x_1}^2 + 4x_1x_2 + 5{x_2}^2 = \left(\frac{2}{\sqrt{5}} x_1 - \frac{1}{\sqrt{5}} x_2\right)^2 + 6\left(\frac{1}{\sqrt{5}} x_1 + \frac{2}{\sqrt{5}} x_2\right)^2 > 0$$

Note that not all real symmetric matrices can have $A = LU$ factorization, but they must have $A = Q\Lambda Q^T$ factorization.

The Principle Axes Theorem

Here is the chapter 6 meaning of $A = Q\Lambda Q^T$ factorization. I will use the knowledge in senior high school to show you.

Use the example in the slide above.

$$2{x_1}^2 + 4x_1x_2 + 5{x_2}^2 = \left(\frac{2}{\sqrt{5}}x_1 - \frac{1}{\sqrt{5}}x_2\right)^2 + 6\left(\frac{1}{\sqrt{5}}x_1 + \frac{2}{\sqrt{5}}x_2\right)^2 > 0$$

Change of variable and let the quadratic form equals 1:

$$\begin{cases} y_1 = \frac{2}{\sqrt{5}}x_1 - \frac{1}{\sqrt{5}}x_2 \\ y_2 = \frac{1}{\sqrt{5}}x_1 + \frac{2}{\sqrt{5}}x_2 \end{cases}$$

$$2{x_1}^2 + 4x_1x_2 + 5{x_2}^2 = {y_1}^2 + 6{y_2}^2 = 1$$

That is the equation of an ellipse! Imagine in your mind: the shape at height 1 is a ellipse. And the change of variable is like rotating the axes, which lands on the axes on the ellipse.

Congruence Transformation

Definition

For an invertible matrix C, the linear transformation $A \mapsto C^T AC$ is called a congruence transformation.

Essence: change of variable x to y and the quadratic form remains unchanged.

The symmetry, rank, signs of eigenvalues are all hold before and after the transformation. That is because, the quadratic form behind never change.

Recall another transformation we meet in 5.6: similarity transformation $B = M^{-1}AM$. Which variables are not changed?

Transform to Standard Quadratic Form

We have introduced 2 approaches: $A = LDL^T$ factorization and $A = Q\Lambda Q^T$ factorization. Here we use another method: Elementary Transformation.

$$f(x_1, x_2, x_3) = {x_1}^2 + 2{x_2}^2 + 5{x_3}^2 + 2x_1x_2 + 2x_1x_3 + 6x_2x_3$$

$$\begin{bmatrix} A \\ I \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 2 & 3 \\ 1 & 3 & 5 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 2 & 4 \\ 1 & -1 & -1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & -1 & 1 \\ 0 & 1 & -2 \\ 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} D \\ C \end{bmatrix}$$

Change of variable: $x = Cy$, then it will be standard quadratic form.

The key: do symmetric row and column operations for A, and only do column operations for I.

My suggestion: $A = LDL^T$ factorization is yyds!

Table of Contents

Introduction

We have introduced so many different decomposition for square matrices in Chapter 5, but what about non-square matrices?
All decomposition for non-square matrices:

1. $PA = LU$, P is permutation matrix, L is lower triangular matrix with 1s on the diagonal, and U is the "upper something", echelon form.
2. $A = QR$, Q is matrix with orthonormal columns, R is an upper triangular matrices.

Here we have the third one: singular value decomposition $A = U\Sigma V^T$.

Try to guess the size and properties of these 3 matrices.

Actually, U is an $m \times m$ orthogonal matrix, V is an $n \times n$ orthogonal matrix, and Σ is a $m \times n$ "diagonal" matrix.

Singular Values

How to make our eigenvalues also make sense for non-square matrices? The first thing you need to think is: how to transform $m \times n$ matrices into square matrices? (Hint: Recall Chapter 3)

Two important (square) matrices for non-square matrix A: AA^T, A^TA.

Properties of those 2 matrices:

1. They are all real symmetric and positive semidefinite.
2. They share the same rank.

$$rank\left(A^TA\right) = rank\left(A\right) = rank\left(A^T\right) = rank\left(AA^T\right)$$

3. They share the same r positive eigenvalues, and the rest are all zeros.

The singular values are the square root of the positive eigenvalues of these 2 matrices.

Singular Values

Example

Find the singular values of the following matrix:

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

$$AA^T = \begin{bmatrix} 0 & -1 & 1 \\ 2 & 0 & 0 \end{bmatrix} \begin{bmatrix} 0 & 2 \\ -1 & 0 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 4 \end{bmatrix}$$

$$A^TA = \begin{bmatrix} 0 & 2 \\ -1 & 0 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 0 & -1 & 1 \\ 2 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 4 & 0 & 0 \\ 0 & 1 & -1 \\ 0 & -1 & 1 \end{bmatrix}$$

The eigenvalues are 2 and 4, so the singular values are $\sqrt{2}, 2$.