Test for Positive Definiteness (正定性的判定)

Lecture 28

Dept. of Math., SUSTech

2022.12

Test for Positive Definiteness

Positive Definiteness

Which symmetric matrices have the property that $x^T Ax > 0$ for all nonzero vectors? There are four or five different ways to answer this question. Let's first consider a 2 by 2 matrix:

$$A = \begin{bmatrix} a & b \\ b & c \end{bmatrix}.$$

A is positive definite when $a > 0$ and $ac - b^2 > 0$. From those conditions, we can obtain that both eigenvalues are positive.

$$\begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \frac{ac - b^2}{a}y^2.$$

We see that to make $x^T Ax > 0$, we need all the pivots positive.

Determinants

How about determinants? What can determinants tell about positive definiteness?

- The two parts of this book were linked by the chapter on determinants. Now we ask what part determinants play.
- It is not enough to require that the determinant of A is positive. If $a = c = -1$ and $b = 0$, then $\det A = 1$, but $A = -I$ is negative definite. The determinant test is applied not only to A itself, giving $ac - b^2 > 0$, but also to the 1 by 1 submatrix a in the upper left-hand corner.
- The natural generalization will involve all n of the upper left submatrices of A.

Test for Positive Definiteness

Here is the main theorem on positive definiteness:

Theorem

Each of the following tests is a necessary and sufficient condition for the real symmetric matrix A to be positive definite:

(I) $x^T Ax > 0$ *for all nonzero real vectors* x.

(II) *All the eigenvalues of* A *satisfy* $\lambda_i > 0$.

(III) *All the* ***upper left submatrices*** A_k *(顺序主子矩阵) have positive determinants.*

(IV) *All the pivots (without row exchanges) satisfy* $d_k > 0$.

Can you prove this theorem?

Proof.

Condition I defines a positive definite matrix. Our first step shows that each eigenvalue will be positive:

$$\text{If } Ax = \lambda x, \text{then} x^T Ax = x^T \lambda x = \lambda ||x||^2.$$

A positive definite matrix has positive eigenvalues, since $x^T Ax > 0$.

Now we go in the other direction. If all $\lambda_i > 0$, we have to prove $x^T Ax > 0$ for every vector x (not just the eigenvectors). Since symmetric matrices have a full set of orthonormal eigenvectors, any x is a combination $c_1x_1 + c_2x_2 + \cdots + c_nx_n$. Then

$$Ax = c_1Ax_1 + \cdots + c_nAx_n = c_1\lambda_1x_1 + \cdots + c_n\lambda_nx_n.$$

Because of the orthogonality $x_i^T x_j = 0$, and the normalization $x_i^T x_i = 1$,

Proof.

$$
\begin{aligned}
x^T A x &= (c_1 x_1^T + \cdots + c_n x_n^T)(c_1 \lambda_1 x_1 + \cdots + c_n \lambda_n x_n) \\
&= c_1^2 \lambda_1 + \cdots + c_n^2 \lambda_n.
\end{aligned}
$$

If every $\lambda_i > 0$, then the above equation shows that $x^T A x > 0$. Thus condition II implies condition I.

($I \Rightarrow III$) The determinant of A is the product of the eigenvalues. And if condition I holds, we already know that these eigenvalues are positive. But we also have to deal with every **upper left submatrix** A_k (顺序主子矩阵). The trick is to look at all nonzero vectors whose last $n-k$ components are zero:

$$
x^T A x = \begin{bmatrix} x_k^T & 0 \end{bmatrix} \begin{bmatrix} A_k & * \\ * & * \end{bmatrix} \begin{bmatrix} x_k \\ 0 \end{bmatrix}.
$$

Proof.

Thus A_k is positive definite. Its eigenvalues (not the same λ_i!) must be positive. Its determinant is their product, so all upper determinants are positive.

$(III \Rightarrow IV)$ According to Section 4.4, the kth pivot d_k is the ratio of $\det\, A_k$ to $\det\, A_{k-1}$. If the determinants are all positive, so are the pivots.

$(IV \Rightarrow I)$ We are given positive pivots, and must deduce that $x^T Ax > 0$. This is what we did in the 2×2 case, by completing the square. The pivots were the numbers outside the squares. To see how that happens for symmetric matrices of any size, we go back to elimination of a symmetric matrix:

$$A = LDL^T.$$

Example 1

Example 1 Decide for or against the positive definiteness of

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

(a) Each test is enough by itself.

(b) It is beautiful that elimination and completing the square are actually the same.

(c) Every diagonal entry a_{ii} must be positive.

Example 1

The pivots are d_i are not to be confused with the eigenvalues. For a typical positive definite matrix, they are two completely different sets of positive numbers. In our 3 by 3 example, probably the determinant test is the easiest:

$$\det(A_1) = 2, \det(A_2) = 3, \det(A_3) = 4.$$

The pivots are the ratios $d_1 = 2, d_2 = \frac{3}{2}, d_3 = \frac{4}{3}$. Ordinarily the eigenvalue test is the longest computation. For this A we know the λ's are all positive:

$$\lambda_1 = 2 - \sqrt{2}, \lambda_2 = 2, \lambda_3 = 2 + \sqrt{2}.$$

Even though it is the hardest to apply to a single matrix, eigenvalues can be the most useful test for theoretical purposes.

Positive Definite Matrices and Least Squares

So far, we connected positive definite matrices to pivots(Chapter 1), determinants (Chapter 4), and eigenvalues (Chapter 5). Now we see them in the least-squares problems of Chapter 3, coming from the rectangular matrices of Chapter 2.

Theorem

The symmetric matrix A is positive definite if and only if (V) There is a matrix R with independent columns such that $A = R^T R$.

The key is to recognize $x^T A x$ as $x^T R^T R x = (Rx)^T (Rx)$. Thus $x^T R^T R x > 0$ and $R^T R$ is positive definite. It remains to find an R for which $A = R^T R$.

Choices for R

- We almost done this twice already:

$$A = LDL^T = (L\sqrt{D})(\sqrt{D}L^T)$$

This Cholesky decomposition has the pivots split evenly between L and L^T.

- Eigenvalues:

$$A = Q\Lambda Q^T = (Q\sqrt{\Lambda})(\sqrt{\Lambda}Q^T)$$

So take $R = \sqrt{\Lambda}Q^T$.

- We also can take QR.

Semidefinite Matrices

The tests for semidefiniteness will relax $x^T Ax > 0, \lambda > 0, d > 0$, and $\det > 0$, to allow zeros to appear.

Theorem

Each of the following tests is a necessary and sufficient condition for the real symmetric matrix A to be positive semidefinite:

(I') $x^T Ax \geq 0$ *for all nonzero real vectors x(this defines positive semidefinite).*

(II') *All the eigenvalues of A satisfy $\lambda_i \geq 0$.*

(III') *No* ***principal submatrices*** *(主子矩阵) have negative determinants.*

(IV') *No pivots are negative.*

(V') *There is a matrix R, possibly with dependent columns, such that $A = R^T R$.*

Example 2

Example 2 Decide whether the following matrix is positive definite, negative definite, semidefinite, or indefinite:

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

A is positive semidefinite, by all five tests.

Ellipsoids in n Dimensions

Throughout this book, geometry has helped the matrix algebra. A linear equation produced a plane. The system $Ax = b$ gives an intersection of planes.

- Ellipse in two dimensions, and an ellipsoid in n dimensions.
- The equation to consider is $x^T Ax = 1$.
- A is identity, diagonal, general matrices.

Example 3

Example 3 $A = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$ and $x^T A x = 5u^2 + 8uv + 5v^2 = 1$.

- The axes of the ellipse point toward the eigenvectors of A. Because $A = A^T$, those eigenvectors and axes are orthogonal.
- The way to see the ellipse properly is to rewrite $x^T A x = 1$:

$$5u^2 + 8uv + 5v^2 = \left(\frac{u}{\sqrt{2}} - \frac{v}{\sqrt{2}} \right)^2 + 9 \left(\frac{u}{\sqrt{2}} + \frac{v}{\sqrt{2}} \right)^2 = 1$$

- This is different from completing the square to $5(u + \frac{4}{5}v)^2 + \frac{9}{5}v^2$, with the pivots outside.

Example 3

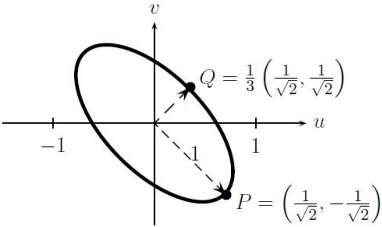

Figure 6.2: The ellipse $x^{\mathrm{T}}Ax = 5u^2 + 8uv + 5v^2 = 1$ and its principal axes.

- The major axis of the ellipse corresponds to the smallest eigenvalue of A.
- The first square equals 1 at $(1/\sqrt{2}, -1/\sqrt{2})$ at the end of the major axis. The minor axis is one-third as long, since we need $\left(\frac{1}{3}\right)^2$ to cancel the 9.

Ellipsoids in n Dimensions

The equation to consider is $x^T Ax = 1$. Any ellipsoid $x^T Ax = 1$ can be simplified in the same way. The key step is to diagonalize $A = Q\Lambda Q^T$, We straightened the picture by rotating the axes. Algebraically, the change to $y = Q^T x$ produces a sum of squares:

$$x^T Ax = (x^T Q)\Lambda(Q^T x) = y^T \Lambda y = \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 = 1$$

The major axis has $y_1 = \frac{1}{\sqrt{\lambda_1}}$ along the eigenvector with the smallest eigenvalue. The other axes are along the other eigenvectors. Their lengths are $1/\sqrt{\lambda_2}, \cdots, 1/\sqrt{\lambda_n}$. Notice that the λ's must be positive–the matrix must be positive definite–or these square roots are in trouble.

Simplifying an ellipsoid in n dimensions

The change from x to $y = Q^T x$ rotates the axes of the space, to match the axes of the ellipsoid. In the y variables we can see that it is an ellipsoid, because the equation becomes so manageable:

Theorem

Suppose $A = Q\Lambda Q^T$ with $\lambda_i > 0$. Rotating $y = Q^T x$ simplifies $x^T Ax = 1$:

$$x^T Q\Lambda Q^T x = 1, \quad y^T \Lambda y = 1, \quad \text{and} \quad \lambda_1 y_1^2 + \cdots + \lambda_n y_n^2 = 1.$$

This is the equation of an ellipsoid. Its axes have lengths $1/\sqrt{\lambda_1}, \cdots, 1/\sqrt{\lambda_n}$ from the center. In the original x-space they point along the eigenvectors of A.

The Law of Inertia

What are the elementary operations and their invariants for $x^T Ax$? The basic operation on a quadratic form is to change variables. A new vector y is related to x by some nonsingular matrix, $x = Cy$. The quadratic form becomes $y^T C^T ACy$. This shows the fundamental operation on A:

$$A \to C^T AC$$

for some nonsingular C. The symmetry of A is preserved, since $C^T AC$ remains symmetric. The real question is, what other properties are shared by A and $C^T AC$? The answer is given by Sylvester's Law of Inertia:

Congruence Transformation

Theorem

*(**Sylvester's Law of Inertia**) C^TAC has the same number of positive eigenvalues, negative eigenvalues, and zero eigenvalues as A.*

The signs of the eigenvalues are preserved by a congruence transformation. Can you prove the above theorem?

Remarks:

1. The number of positive eigenvalues of the real symmetric matrix A, p, is called the positive index of inertia (正惯性指数) of x^TAx.
2. The number of negative eigenvalues of the real symmetric matrix A, q, is called the negative index of inertia (负惯性指数) of x^TAx.
3. $p-q$ is the signature (符号差) of x^TAx. $p+q=r$, where r is the rank of A.

Examples

- Example 4 Suppose $A = I$. Then $C^T AC = C^T C$ is positive definite. Both I and $C^T C$ have n positive eigenvalues, confirming the law of inertia.
- Example 5 If $A = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$, then $C^T AC$ has a negative determinant:

$$\det C^T AC = -(\det C)^2 < 0.$$

Then $C^T AC$ must have one positive and one negative eigenvalue, like A.

惯性定理的中文表述

Theorem

设 $f(x_1,x_2,\cdots,x_n)$ 是秩为 r 的 n 元二次型, 则一定存在可逆线性替换 $X=CY$, 把 $f(x_1,x_2,\cdots,x_n)$ 变为

$$g(y_1,y_2,\cdots,y_n)=y_1^2+y_2^2+\cdots+y_p^2-y_{p+1}^2-y_{p+2}^2-\cdots-y_r^2.$$

这个标准形称为实二次型 $f(x_1,x_2,\cdots,x_n)$ 的规范形. $f(x_1,x_2,\cdots,x_n)$ 的规范形由 $f(x_1,x_2,\cdots,x_n)$ 唯一决定.

Remark: 该定理中的“惯性”是指在变换下保持不变的东西.

Example 6

The following is an important application:

Theorem

For any symmetric matrix A, the signs of the pivots agree with the signs of the eigenvalues. The eigenvalue matrix Λ and the pivot matrix D have the same number of positive entries, negative entries, and zero entries.

- This is both beautiful and practical. It is beautiful because it brings together(for symmetric matrices) two parts of this book that were previously separate: pivots and eigenvalues.
- It is also practical, because the pivots can **locate** the eigenvalues.

One Application of the Law of Inertia

- This was almost the first practical method of computing eigenvalues. It was dominant about 1960, after one important improvement–to make A tridiagonal first.
- The the pivots are computed in $2n$ steps instead of $\frac{1}{6}n^3$. Elimination becomes fast, and the search for eigenvalues becomes simple. The current favorite is the QR method in Chapter 7.

The Generalized Eigenvalue Problem

Sometimes $Ax = \lambda x$ is replaced by $Ax = \lambda Mx$. There are two matrices rather than one. An example is the motion of two unequal masses in a line of springs:

$$m_1 \frac{d^2v}{dt^2} + 2v - w = 0$$

$$m_2 \frac{d^2w}{dt^2} - v + 2w = 0$$

This can be reduced to an eigenvalue problem:

$$Ax = \lambda Mx.$$

As long as M is positive definite, the generalized eigenvalue problem $Ax = \lambda Mx$ behaves exactly like $Ax = \lambda x$.

Equivalent problem and "M-orthogonality"

In the following discussion, M is assumed to be positive definite. As a consequence, M can be split into $R^T R$.

- Equivalent problem:

$$C^T A C y = \lambda y.$$

- The properties of $C^T A C$ lead directly to the properties of $Ax = \lambda Mx$, when $A = A^T$ and M is positive definite.
- A and M are being simultaneously diagonalized.
- As long as M is positive definite, the generalized eigenvalue problem $Ax = \lambda Mx$ behaves exactly like $Ax = \lambda x$.

一些习题

1. 设二次型 $f(x_1,x_2,x_3)=x_1^2-x_2^2+2ax_1x_3+4x_2x_1$ 的负惯性指数是 1, 则 a 的取值范围是 ____.
2. 二次型 $f(x_1,x_2,x_3)=(x_1+x_2)^2+(x_2+x_3)^2-(x_3-x_1)^2$ 的正惯性指数和负惯性指数分别为 ____.
3. 设 A 是三阶实对称矩阵，E 为三阶单位矩阵，若 $A^2+A=2E$, 且 $|A|=4$, 则二次型 x^TAx 的规范形为 ____.
4. 设二次型 $f(x_1,x_2,x_3)$ 在正交变换为 $x=Py$ 下的标准形为 $2y_1^2+y_2^2-y_3^2$, 其中 $P=(e_1,e_2,e_3)$, 若 $Q=(e_1,-e_3,e_2)$, 则 $f(x_1,x_2,x_3)$ 在正交变换 $x=Qy$ 下的标准形为 ____.
5. 设二次型 $f(x_1,x_2,x_3)=x_1^2+x_2^2+x_3^2+4x_1x_2+4x_2x_3+4x_1x_3$, 则 $f(x_1,x_2,x_3)=2$ 在空间直角坐标系下表示的二次曲面为 ____.

Homework Assignment 28

6.2: 2, 6, 7, 9, 14, 15, 24, 30, 36, 38.