Lecture 12. Sylvester's Criterion
November 2022

Law of Inertia

Note that a quadratic form q in variables x_i can be written in different ways as a linear combination of squares of new variables, even if the new variables are required to be linear combinations of the x_i. For example, if $q = x_1^2 - 4x_1x_2 - x_2^2$ then

$$q = (x_1 - 2x_2)^2 - 5x_2^2 \qquad \text{and} \qquad q = 5x_1^2 - (2x_1 + x_2)^2$$

The question arises: How are these changes of variables related, and what properties do they share? To investigate this, we need a new concept.

DEFINITION 12.1 A quadratic form $q = q(\mathbf{x})$ is said to be

- **positive definite** over a subspace $U \subset V$ if $q(\mathbf{x}) > 0$ for any $\mathbf{x} \in U$, $\mathbf{x} \neq \mathbf{0}$;
- **positive semidefinite** over a subspace $U \subset V$ if $q(\mathbf{x}) \geqslant 0$ for any $\mathbf{x} \in U$;
- **negative definite** over a subspace $U \subset V$ if $q(\mathbf{x}) < 0$ for any $\mathbf{x} \in U$, $\mathbf{x} \neq \mathbf{0}$;
- **negative semidefinite** over a subspace $U \subset V$ if $q(\mathbf{x}) \leqslant 0$ for any $\mathbf{x} \in U$;
- **indefinite** over a subspace $U \subset V$ if $q(\mathbf{x}) > 0$ and $q(\mathbf{y}) < 0$ for some $\mathbf{x}, \mathbf{y} \in U$.

DEFINITION 12.2 A quadratic form $q = q(\mathbf{x}) = \mathbf{x}^T A\mathbf{x}$ and the associated matrix A are said to be **positive definite**, **positive semidefinite**, **negative definite**, **negative semidefinite**, or **indefinite** if q is positive definite, positive semidefinite, negative definite, negative semidefinite, or indefinite over the entire V, respectively.

DEFINITION 12.3 Let V be an n-dimensional vector space and q be a quadratic form on V of rank r.

- The dimension of a largest subspace on which q is positive definite is called the **positive index of inertia** of q and is denoted by n_+.
- The dimension of a largest subspace on which q is negative definite is called the **negative index of inertia** of q and is denoted by n_-.
- The triple (n_+, n_-, n_0), where $n_0 = n - r$ is a number of zeros, is called the **signature** of q.

As we can see from the definition, the indices of inertia depend only on q and not on the way it is expressed.

LEMMA 12.4 If a quadratic form q is represented as a signed sum of squares

$$q(\mathbf{x}) = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_r^2, \qquad 0 \leqslant k \leqslant r \leqslant n,$$

then $n_+ = k$ and $n_- = r - k$.

Proof. Let $V_+ \subset V$ be a largest subspace over which the quadratic form is positive definite. By definition, $\dim V_+ = n_+$.

Consider a basis $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ such that for $\mathbf{x} = (x_1, x_2, \ldots, x_n)$ we have

$$q(\mathbf{x}) = x_1^2 + \cdots + x_k^2 - x_{k+1}^2 - \cdots - x_r^2, \qquad , 0 \leqslant k \leqslant r \leqslant n,$$

and show that $n_+ = k$ and $n_- = r - k$.

Indeed, let $V_1 = \operatorname{span}\{\mathbf{e}_1, \dots, \mathbf{e}_k\}$ and $V_2 = \operatorname{span}\{\mathbf{e}_{k+1}, \dots, \mathbf{e}_n\}$. Since for any vector $\mathbf{x} \in V_1$ we have $\mathbf{x} = (x_1, \dots, x_k, 0, \dots, 0)^T$, then it is clear that

$$q(\mathbf{x}) = x_1^2 + \cdots + x_k^2 > 0 \qquad \text{if} \qquad \mathbf{x} \neq \mathbf{0}.$$

If $\mathbf{x} \in V_2$ we have $\mathbf{x} = (0, \dots, 0, x_{k+1}, \dots, x_n)^T$, then it is clear that

$$q(\mathbf{x}) = -x_{k+1}^2 - \cdots - x_r^2 \leqslant 0.$$

We have $\dim V_2 = n - k$. Had we had $\dim V_+ = n_+ > k$, then V_+ and V_2 would have a nonzero vector $\mathbf{v}$ in common because $\dim V_+ + \dim V_2 > n$. Therefore,

$$q(\mathbf{v}) > 0 \quad \text{as } \mathbf{v} \in V, \qquad q(\mathbf{v}) \leqslant 0 \quad \text{as } \mathbf{v} \in V_2.$$

This contradiction proves $n_+ \leqslant k$. On the other hand, the quadratic form is positive definite over V_1. Then $n_+ \geqslant k$ and, therefore, $n_+ = k$. By analogy, we can prove that $n_- = r - k$. $\square$

THEOREM 12.5 (Sylvester's Law of Inertia) Let A and B be symmetric $n \times n$ matrices. Then $A \overset{c}{\sim} B$ if and only if they have the same signature.

Proof. 1) If two matrices A and B are congruent, i.e., $A \overset{c}{\sim} B$, then they represent a quadratic form q in different bases. Since the notions of indices of inertia do not depend on the choice of bases, then A and B have the same signature.

2) Suppose A and B have the same signature. If A and B both have positive index n_+ and negative index n_-, then

$$A \overset{c}{\sim} \operatorname{diag}(\underbrace{1, \dots, 1}_{n_+}, \underbrace{-1, \dots, -1}_{n_-}, 0, \dots, 0),$$
$$B \overset{c}{\sim} \operatorname{diag}(\underbrace{1, \dots, 1}_{n_+}, \underbrace{-1, \dots, -1}_{n_-}, 0, \dots, 0)$$

The statement follows by transitivity of congruence. $\square$

Definite Quadratic Forms

Now our purpose is to develop a criterion for determining whether a quadratic form is positive definite, negative definite, or indefinite.

DEFINITION 12.6 The jth **leading principal submatrix** of an $n \times n$ matrix A is defined to be the $j \times j$ submatrix consisting of the first j rows and the first j columns of A (i.e., the upper left $j \times j$ corner of A) and is denoted by $A_{(j)}$, $j = 1, \dots, n$, where $A_{(n)} = A$. The determinants $\Delta_j := \det A_{(j)}$, $j = 1, \dots, n$, are called **leading principal minors** of A. In addition, we will set $\Delta_0 = 1$.

EXAMPLE 12.7 If $A = \begin{pmatrix} 2 & -3 & 1 \\ -3 & 1 & 4 \\ 1 & 4 & 0 \end{pmatrix}$, then $A_{(1)} = (2)$, $A_{(2)} = \begin{pmatrix} 2 & -3 \\ -3 & 1 \end{pmatrix}$, and $A_{(3)} = A$. Therefore, $\Delta_0 = 1$, $\Delta_1 = 2$, $\Delta_2 = -7$, $\Delta_3 = -57$.

LEMMA 12.8 If A is positive definite, then it is invertible and $\det A > 0$.

Proof. If A is positive definite, then $A = P^T D P$ for some invertible matrix P and a diagonal matrix D with positive diagonal entries. Since $\det D > 0$, it follows that $\det A = (\det P)^2 \det D > 0$. $\square$

LEMMA 12.9 If A is positive definite, so is each principal submatrix $A_{(j)}$ for $j = 1, 2, \ldots, n$.

Proof. Write $A = \begin{pmatrix} A_{(j)} & S \\ S^T & Q \end{pmatrix}$ in block form. Since A is positive definite, then for any $\mathbf{x} \in V$, $\mathbf{x} \neq \mathbf{0}$ there must be $\mathbf{x}^T A \mathbf{x} > 0$. If, in particular, we consider the vectors $\mathbf{x} = \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix}$, where $\mathbf{y} \neq \mathbf{0}$ is in $\mathbb{R}^r$, then we obtain

$$\mathbf{x}^T A \mathbf{x} = (\mathbf{y}^T \quad \mathbf{0}) \begin{pmatrix} A_{(j)} & S \\ S^T & Q \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix} = \mathbf{y}^T A_{(j)} \mathbf{y} > 0.$$

This shows that $A_{(j)}$ is positive definite. $\square$

COROLLARY 12.10 If A is positive definite, then $\Delta_j > 0$ for every $j = 1, 2, \ldots, n$.

THEOREM 12.11 (Sylvester's Criterion) If q is a quadratic form with a symmetric matrix A, then:

(1) q is positive definite if and only if all leading principle minors of A are positive.
(2) q is negative definite if and only if the leading principle minors alternate between negative and positive values starting with a negative value for Δ_1.

Proof (1). Implication $\Rightarrow$ follows from Corollary 12.10.

Suppose $\Delta_j > 0$ for every $j = 1, 2, \ldots, n$ and show that A is positive definite.

We proceed by induction on n. The base is trivial. Consider a vector subspace $U = \operatorname{span}\{\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_{n-1}\}$ and let $\bar{q} = q|_U$ be a restriction of q to U. Since $\Delta_j > 0$ for every $j = 1, 2, \ldots, n-1$, by inductive hypothesis, $A|_{\bar{q}}$ is positive definite and, by definition, $n_+ \geqslant n - 1$. The rank of q is equal to n, because $\Delta_n = |A| > 0$. Therefore, there are only two options for the signature of q: $(n-1, 1, 0)$ or $(n, 0, 0)$. By Lemma 11.12, the determinants of two congruent matrices are either both zero or have the same sign. Therefore, the first case is not possible and the signature of q is $(n, 0, 0)$.

To prove part (2) note that A is negative definite if and only if $B = -A$ is positive definite. Since $\det(B_{(j)}) = (-1)^j \det(A_{(j)})$ we obtain the required result. $\square$

EXAMPLE 12.12 Determine whether the following quadratic form is positive definite, negative definite, or indefinite.

$$q = 5x_1^2 - 2x_1x_2 + 3x_2^2 + 4x_1x_3 + x_3^2.$$

Solution. Positive definite by Sylvester's criterion.