

Lecture Notes: Matrix Algebra

Part E: Quadratic Forms and Their Definiteness

Peter J. Hammond

latest version 2025 September 19th

Outline

Quadratic Forms and Their Definiteness

Quadratic Forms and Symmetric Matrices

Definiteness of a Quadratic Form: Definition

Definiteness of a Quadratic Form: Some Simple Tests

The Two Variable Case

The Block Diagonal Case

Definiteness Tests by Symmetric Pivoting

Quadratic Form Invariance

Symmetric Straightforward Pivoting

Adjusted Symmetric Pivoting

Sylvester's Criterion for (Semi-)Definite Quadratic Forms

Sylvester's Criterion: Statement and Preliminary Examples

Proving Sylvester's Criterion

Quadratic Forms: Preliminary Exercise

Exercise

Let \mathbf{A} be any $n \times n$ matrix.

For each $j \in \mathbb{N}_n$, recall that $\mathbf{e}^j = (\delta_{ij})_{i=1}^n$ denotes the j th column of the identity matrix \mathbf{I}_n , and that $(\mathbf{e}^i)^\top = (\delta_{ij})_{j=1}^n$ is the i th row of \mathbf{I}_n .

For each $i, j \in \mathbb{N}_n$, explain why:

1. $\mathbf{A}\mathbf{e}^j$ is the j th column \mathbf{a}^j of the matrix \mathbf{A} , whose i th component is $\sum_{k=1}^n a_{ik}\delta_{kj} = a_{ij}$;
2. $(\mathbf{e}^i)^\top \mathbf{A}\mathbf{e}^j = \sum_{k=1}^n \delta_{ik}a_{kj}$, which equals the single element a_{ij} of the matrix \mathbf{A} .

Definition of Quadratic Form

Definition

A **quadratic form** on the n -dimensional Euclidean space \mathbb{R}^n is a mapping

$$\mathbb{R}^n \ni \mathbf{x} \mapsto q(\mathbf{x}) = \mathbf{x}^\top \mathbf{Q} \mathbf{x} = \sum_{i=1}^n \sum_{j=1}^n x_i q_{ij} x_j \in \mathbb{R}$$

where \mathbf{Q} is a **symmetric** $n \times n$ matrix.

The quadratic form $\mathbf{x}^\top \mathbf{Q} \mathbf{x}$ is **diagonal** just in case the matrix \mathbf{Q} is diagonal, with $\mathbf{Q} = \mathbf{\Lambda} = \mathbf{diag}(\lambda_1, \lambda_2, \dots, \lambda_n)$.

In this case $\mathbf{x}^\top \mathbf{Q} \mathbf{x}$ reduces to $\mathbf{x}^\top \mathbf{\Lambda} \mathbf{x} = \sum_{i=1}^n \lambda_i (x_i)^2$.

The Hessian Matrix of a Quadratic Form in Two Variables

Exercise

Given the quadratic form $q(x, y) = (x, y) \begin{pmatrix} a & b \\ c & d \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$,

show that, even if the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is not symmetric, the Hessian matrix of its second-order partial derivatives

is the symmetric matrix $\begin{pmatrix} q''_{xx} & q''_{xy} \\ q''_{yx} & q''_{yy} \end{pmatrix} = \begin{pmatrix} 2a & b + c \\ b + c & 2d \end{pmatrix}$.

The Hessian Matrix of a Quadratic Form in n Variables

Exercise

Consider the n -variable quadratic form

$$\mathbb{R}^n \ni \mathbf{x} \mapsto q(\mathbf{x}) = \mathbf{x}^\top \mathbf{A} \mathbf{x} \in \mathbb{R}$$

Show that, even if the matrix \mathbf{A} is not symmetric, the $n \times n$ Hessian matrix \mathbf{H} whose elements h_{ij}

are the constant second-order partial derivatives $\frac{\partial^2 q}{\partial x_i \partial x_j}$

is the symmetric matrix $\mathbf{A} + \mathbf{A}^\top$.

Symmetry Loses No Generality

Requiring \mathbf{Q} in $\mathbf{x}^\top \mathbf{Q} \mathbf{x}$ to be symmetric loses no generality.

This is because, given a general non-symmetric $n \times n$ matrix \mathbf{A} , repeated transposition implies that

$$\mathbf{x}^\top \mathbf{A} \mathbf{x} = (\mathbf{x}^\top \mathbf{A} \mathbf{x})^\top = \frac{1}{2}[\mathbf{x}^\top \mathbf{A} \mathbf{x} + (\mathbf{x}^\top \mathbf{A} \mathbf{x})^\top] = \frac{1}{2} \mathbf{x}^\top (\mathbf{A} + \mathbf{A}^\top) \mathbf{x}$$

Hence $\mathbf{x}^\top \mathbf{A} \mathbf{x} = \mathbf{x}^\top \mathbf{A}^\top \mathbf{x} = \mathbf{x}^\top \mathbf{Q} \mathbf{x}$

where \mathbf{Q} is the **symmetrized** matrix $\frac{1}{2}(\mathbf{A} + \mathbf{A}^\top)$.

Note that \mathbf{Q} is indeed symmetric because

$$\mathbf{Q}^\top = \frac{1}{2}(\mathbf{A} + \mathbf{A}^\top)^\top = \frac{1}{2}[\mathbf{A}^\top + (\mathbf{A}^\top)^\top] = \frac{1}{2}(\mathbf{A}^\top + \mathbf{A}) = \mathbf{Q}$$

Outline

Quadratic Forms and Their Definiteness

Quadratic Forms and Symmetric Matrices

Definiteness of a Quadratic Form: Definition

Definiteness of a Quadratic Form: Some Simple Tests

The Two Variable Case

The Block Diagonal Case

Definiteness Tests by Symmetric Pivoting

Quadratic Form Invariance

Symmetric Straightforward Pivoting

Adjusted Symmetric Pivoting

Sylvester's Criterion for (Semi-)Definite Quadratic Forms

Sylvester's Criterion: Statement and Preliminary Examples

Proving Sylvester's Criterion

Definiteness of a Quadratic Form

When $\mathbf{x} = \mathbf{0}$, then $\mathbf{x}^\top \mathbf{Q} \mathbf{x} = 0$. Otherwise:

Definition

The quadratic form $\mathbb{R}^n \ni \mathbf{x} \mapsto \mathbf{x}^\top \mathbf{Q} \mathbf{x} \in \mathbb{R}$,

as well as its associated symmetric $n \times n$ matrix \mathbf{Q} , are both:

positive definite just in case $\mathbf{x}^\top \mathbf{Q} \mathbf{x} > 0$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$;

negative definite just in case $\mathbf{x}^\top \mathbf{Q} \mathbf{x} < 0$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$;

positive semi-definite just in case $\mathbf{x}^\top \mathbf{Q} \mathbf{x} \geq 0$ for all $\mathbf{x} \in \mathbb{R}^n$;

negative semi-definite just in case $\mathbf{x}^\top \mathbf{Q} \mathbf{x} \leq 0$ for all $\mathbf{x} \in \mathbb{R}^n$;

indefinite just in case there exist both \mathbf{x}^+ and \mathbf{x}^- in \mathbb{R}^n
such that $(\mathbf{x}^+)^\top \mathbf{Q} \mathbf{x}^+ > 0$ and $(\mathbf{x}^-)^\top \mathbf{Q} \mathbf{x}^- < 0$. □

Given the domain $\mathcal{Q}_{n \times n}$ of symmetric $n \times n$ matrices,

the sign of each $\mathbf{Q} \in \mathcal{Q}_{n \times n}$ is indicated,

using some obvious abbreviations, by the **definiteness function**

$$\mathcal{Q}_{n \times n} \ni \mathbf{Q} \mapsto \text{def}(\mathbf{Q}) \in \{\text{PD}, \text{ND}, \text{PSD}, \text{NSD}, \text{ID}\}$$

Definiteness of a Diagonal Quadratic Form

Theorem

Suppose that \mathbf{Q} is the diagonal matrix $\mathbf{diag}(\lambda_1, \dots, \lambda_n)$, so that $\mathbf{x}^\top \mathbf{Q} \mathbf{x} = \sum_{i=1}^n \lambda_i (x_i)^2$.

Then the diagonal quadratic form $\mathbf{x} \mapsto \sum_{i=1}^n \lambda_i (x_i)^2 \in \mathbb{R}$ is:

positive definite if and only if $\lambda_i > 0$ for $i = 1, 2, \dots, n$;

negative definite if and only if $\lambda_i < 0$ for $i = 1, 2, \dots, n$;

positive semi-definite if and only if $\lambda_i \geq 0$ for $i = 1, 2, \dots, n$;

negative semi-definite if and only if $\lambda_i \leq 0$ for $i = 1, 2, \dots, n$;

indefinite if and only if there exist $i, j \in \{1, 2, \dots, n\}$
such that $\lambda_i > 0$ and $\lambda_j < 0$.

Proof.

The proof is left as an exercise.

The result is obvious if $n = 1$, and straightforward if $n = 2$.

Working out these two cases first suggests the proof for $n > 2$. \square

Concavity or Convexity of a Quadratic Form

Exercise

Show that, as a function $\mathbb{R}^n \ni \mathbf{x} \mapsto \mathbf{x}^\top \mathbf{Q} \mathbf{x} \in \mathbb{R}$,
the quadratic form $\mathbf{x}^\top \mathbf{Q} \mathbf{x}$ is:

strictly convex if and only if \mathbf{Q} is positive definite;

strictly concave if and only if \mathbf{Q} is negative definite;

convex if and only if \mathbf{Q} is positive semi-definite;

concave if and only if \mathbf{Q} is negative semi-definite.

Otherwise $\mathbf{x}^\top \mathbf{Q} \mathbf{x}$ is neither concave nor convex
if and only if \mathbf{Q} is indefinite.

The solution is more suited to Pablo's lectures than mine!

Outline

Quadratic Forms and Their Definiteness

Quadratic Forms and Symmetric Matrices

Definiteness of a Quadratic Form: Definition

Definiteness of a Quadratic Form: Some Simple Tests

The Two Variable Case

The Block Diagonal Case

Definiteness Tests by Symmetric Pivoting

Quadratic Form Invariance

Symmetric Straightforward Pivoting

Adjusted Symmetric Pivoting

Sylvester's Criterion for (Semi-)Definite Quadratic Forms

Sylvester's Criterion: Statement and Preliminary Examples

Proving Sylvester's Criterion

Definiteness of a Quadratic Form: Simple Tests

Even if \mathbf{Q} is not a diagonal matrix,
its diagonal elements still provide useful information.

Proposition

The quadratic form $\mathbb{R}^n \ni \mathbf{x} \mapsto \mathbf{x}^\top \mathbf{Q} \mathbf{x}$ is:

1. positive definite only if $q_{ii} > 0$ for all $i \in \mathbb{N}_n$;
2. positive semi-definite only if $q_{ii} \geq 0$ for all $i \in \mathbb{N}_n$;
3. negative definite only if $q_{ii} < 0$ for all $i \in \mathbb{N}_n$;
4. negative semi-definite only if $q_{ii} \leq 0$ for all $i \in \mathbb{N}_n$;
5. indefinite if there exist $i, j \in \mathbb{N}_n$ such that $q_{ii} > 0 > q_{jj}$.

Proof.

For each $i \in \mathbb{N}_n$, recall that $(\mathbf{e}^i)^\top \mathbf{Q} \mathbf{e}^i = q_{ii}$
where \mathbf{e}^i denotes column i of \mathbf{I}_n .

The result follows from checking the signs
of $(\mathbf{e}^i)^\top \mathbf{Q} \mathbf{e}^i = q_{ii}$ and $(\mathbf{e}^j)^\top \mathbf{Q} \mathbf{e}^j = q_{jj}$ in the 5 different cases. \square

Semi-Definiteness of a Quadratic Form

Suppose that the diagonal of the symmetric $n \times n$ matrix \mathbf{Q} has at least one zero element.

By the previous proposition, the quadratic form $\mathbb{R}^n \ni \mathbf{x} \mapsto \mathbf{x}^T \mathbf{Q} \mathbf{x}$ cannot be either positive definite or negative definite.

But could it still be
either positive semi-definite or negative semi-definite?

Semi-Definiteness of a Quadratic Form: Simple Test

Proposition

Suppose that the diagonal of the symmetric $n \times n$ matrix \mathbf{Q} has two zero elements q_{ii} and q_{jj} with $i \neq j$.

Then the quadratic form $\mathbb{R}^n \ni \mathbf{x} \mapsto \mathbf{x}^\top \mathbf{Q} \mathbf{x}$ is indefinite unless one has $q_{ij} = q_{ji} = 0$.

Proof.

Consider the particular column vector $\mathbf{x} = \alpha \mathbf{e}^i + \beta \mathbf{e}^j \in \mathbb{R}^n$ where α and β are any two real scalars.

Routine calculation shows that, because \mathbf{Q} is symmetric, and $(\mathbf{e}^i)^\top \mathbf{Q} \mathbf{e}^i = (\mathbf{e}^j)^\top \mathbf{Q} \mathbf{e}^j = q_{ii} = q_{jj} = 0$, one has

$$\mathbf{x}^\top \mathbf{Q} \mathbf{x} = \alpha\beta[(\mathbf{e}^i)^\top \mathbf{Q} \mathbf{e}^j + (\mathbf{e}^j)^\top \mathbf{Q} \mathbf{e}^i] = 2q_{ij}\alpha\beta$$

As the pair (α, β) ranges over all of $\mathbb{R}^2 \setminus \{\mathbf{0}\}$, one has $\alpha\beta \gtrless 0$ according as $\text{sgn } \alpha = \pm \text{sgn } \beta$.

So $\mathbf{x}^\top \mathbf{Q} \mathbf{x} = 2q_{ij}\alpha\beta$ is indefinite unless $q_{ij} = q_{ji} = 0$. □

Outline

Quadratic Forms and Their Definiteness

Quadratic Forms and Symmetric Matrices

Definiteness of a Quadratic Form: Definition

Definiteness of a Quadratic Form: Some Simple Tests

The Two Variable Case

The Block Diagonal Case

Definiteness Tests by Symmetric Pivoting

Quadratic Form Invariance

Symmetric Straightforward Pivoting

Adjusted Symmetric Pivoting

Sylvester's Criterion for (Semi-)Definite Quadratic Forms

Sylvester's Criterion: Statement and Preliminary Examples

Proving Sylvester's Criterion

The Two Variable Case: Completing the Square

In the 2×2 case, the typical quadratic form is

$$\mathbb{R}^2 \ni \begin{pmatrix} x \\ y \end{pmatrix} \mapsto (x \ y) \begin{pmatrix} a & h \\ h & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2hxy + by^2$$

Assuming that $a \neq 0$, one can complete the square

by writing $ax^2 + 2hxy + by^2$ as $a \left(x + \frac{h}{a}y\right)^2 + \left(b - \frac{h^2}{a}\right)y^2$,

which can be verified term by term.

First, the quadratic form $ax^2 + 2hxy + by^2$

is neither positive nor negative definite in case:

- ▶ $a = 0$, because then $ax^2 + 2hxy + by^2 = 0$ when $x \neq 0$ and $y = 0$;
- ▶ $a \neq 0$ but $ab - h^2 = 0$, because then $ax^2 + 2hxy + by^2 = 0$ when $y \neq 0$ and $x = -hy/a$.

Tests Based on Completing the Square

We are considering the quadratic form which, in case $a \neq 0$, after completing the square, becomes

$$ax^2 + 2hxy + by^2 = a \left(x + \frac{h}{a}y \right)^2 + \left(b - \frac{h^2}{a} \right) y^2$$

If $a > 0$, because $b - \frac{h^2}{a} = \frac{1}{a}(ab - h^2)$, the quadratic form is:

positive definite if and only if $ab - h^2 > 0$;

positive semi-definite if and only if $ab - h^2 \geq 0$;

indefinite if and only if $ab - h^2 < 0$.

If $a < 0$, because $b - \frac{h^2}{a} = \frac{1}{a}(ab - h^2)$, the quadratic form is:

negative definite if and only if $ab - h^2 > 0$;

negative semi-definite if and only if $ab - h^2 \geq 0$;

indefinite if and only if $ab - h^2 < 0$.

Completing the Square as Symmetric Pivoting, I

Given the 2×2 matrix $\begin{pmatrix} a & h \\ h & b \end{pmatrix}$, provided that $a \neq 0$, the downward pivoting operation involves adding $-h/a$ times row 1 to row 2.

In symbols, this downward pivoting operation is represented by the 2×2 matrix $\mathbf{E}_{11}^\downarrow = \mathbf{E}_{2+(-h/a)1} = \begin{pmatrix} 1 & 0 \\ -h/a & 1 \end{pmatrix}$.

Applied to the original matrix, the result is

$$\mathbf{E}_{11}^\downarrow \begin{pmatrix} a & h \\ h & b \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -h/a & 1 \end{pmatrix} \begin{pmatrix} a & h \\ h & b \end{pmatrix} = \begin{pmatrix} a & h \\ 0 & b - h^2/a \end{pmatrix}$$

Completing the Square as Symmetric Pivoting, II

Starting with the equation $\mathbf{E}_{11}^\downarrow \begin{pmatrix} a & h \\ h & b \end{pmatrix} = \begin{pmatrix} a & h \\ 0 & b - h^2/a \end{pmatrix}$,
suppose now we post-multiply each side by the transpose $(\mathbf{E}_{11}^\downarrow)^\top$.

This completes a **symmetric** pivoting operation whose result is

$$\begin{aligned} \mathbf{E}_{11}^\downarrow \begin{pmatrix} a & h \\ h & b \end{pmatrix} (\mathbf{E}_{11}^\downarrow)^\top &= \begin{pmatrix} a & h \\ 0 & b - h^2/a \end{pmatrix} \begin{pmatrix} 1 & -h/a \\ 0 & 1 \end{pmatrix} \\ &= \begin{pmatrix} a & 0 \\ 0 & b - h^2/a \end{pmatrix} = \mathbf{diag}(a, b - h^2/a) \end{aligned}$$

The two-part symmetric pivoting operation
converts the original quadratic form $ax^2 + 2hxy + by^2$
to the diagonal form $az^2 + (b - h^2/a)w^2$

where $\begin{pmatrix} x \\ y \end{pmatrix} = \mathbf{E}_{11}^\downarrow \begin{pmatrix} z \\ w \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ -h/a & 1 \end{pmatrix} \begin{pmatrix} z \\ w \end{pmatrix}$

and so $\begin{pmatrix} z \\ w \end{pmatrix} = (\mathbf{E}_{11}^\downarrow)^{-1} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ h/a & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix}$.

Example Where Symmetric Pivoting Is Impossible

Example

Let \mathbf{A} be the 2×2 symmetric matrix $\begin{pmatrix} 0 & h \\ h & 0 \end{pmatrix}$.

Both diagonal elements are zero.

These zeroes make symmetric pivoting impossible, so one cannot complete the square in the quadratic form

$$(x \ y) \mathbf{A} \begin{pmatrix} x \\ y \end{pmatrix} = (x \ y) \begin{pmatrix} 0 & h \\ h & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = (x \ y) \begin{pmatrix} hy \\ hx \end{pmatrix} = 2hxy$$

Fortunately the definiteness of \mathbf{A} is easy to determine directly.

If $h \neq 0$ then \mathbf{A} is indefinite.

If $h = 0$ then $\mathbf{A} = \mathbf{0}_{2 \times 2}$,
the only 2×2 symmetric matrix that is both
positive semi-definite and negative semi-definite.

Outline

Quadratic Forms and Their Definiteness

Quadratic Forms and Symmetric Matrices

Definiteness of a Quadratic Form: Definition

Definiteness of a Quadratic Form: Some Simple Tests

The Two Variable Case

The Block Diagonal Case

Definiteness Tests by Symmetric Pivoting

Quadratic Form Invariance

Symmetric Straightforward Pivoting

Adjusted Symmetric Pivoting

Sylvester's Criterion for (Semi-)Definite Quadratic Forms

Sylvester's Criterion: Statement and Preliminary Examples

Proving Sylvester's Criterion

The Block Diagonal Case: Proposition

Proposition

Suppose that $\mathbf{A} = \mathbf{diag}(\mathbf{B}, \mathbf{C})$

is a symmetric block diagonal matrix.

Then \mathbf{A} is positive definite (resp. semi-definite) if and only if both blocks \mathbf{B} and \mathbf{C} are positive definite (resp. semi-definite).

Corollary

Suppose that $\mathbf{A} = \mathbf{diag}(\mathbf{A}^{(1)}, \dots, \mathbf{A}^{(k)})$ is symmetric.

Then \mathbf{A} is positive definite (resp. semi-definite) if and only if each block $\mathbf{A}^{(i)}$ ($i \in \mathbb{N}_k$) is positive definite (resp. semi-definite).

The corollary is easily proved by induction on k .

Remark

As usual, the result for a negative (semi-)definite matrix \mathbf{A} follows from the corresponding result for the positive (semi-)definite matrix $-\mathbf{A}$.

The Block Diagonal Case: Proof

Proof.

Consider the block diagonal quadratic form

$$(\mathbf{y}^\top, \mathbf{z}^\top) \mathbf{A} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} = (\mathbf{y}^\top, \mathbf{z}^\top) \begin{pmatrix} \mathbf{B} & \mathbf{0} \\ \mathbf{0} & \mathbf{C} \end{pmatrix} \begin{pmatrix} \mathbf{y} \\ \mathbf{z} \end{pmatrix} = \mathbf{y}^\top \mathbf{B} \mathbf{y} + \mathbf{z}^\top \mathbf{C} \mathbf{z}$$

If \mathbf{A} is positive definite (resp. semi-definite),

then $(\mathbf{y}^\top, \mathbf{0}^\top) \mathbf{A} \begin{pmatrix} \mathbf{y} \\ \mathbf{0} \end{pmatrix} = \mathbf{y}^\top \mathbf{B} \mathbf{y} > 0$ (resp. ≥ 0) for all $\mathbf{y} \neq \mathbf{0}$,

and $(\mathbf{0}^\top, \mathbf{z}^\top) \mathbf{A} \begin{pmatrix} \mathbf{0} \\ \mathbf{z} \end{pmatrix} = \mathbf{z}^\top \mathbf{C} \mathbf{z} > 0$ (resp. ≥ 0) for all $\mathbf{z} \neq \mathbf{0}$.

So both \mathbf{B} and \mathbf{C} are positive definite (resp. semi-definite).

Conversely, if both \mathbf{B} and \mathbf{C} are positive definite, then so is \mathbf{A} because $\mathbf{y}^\top \mathbf{B} \mathbf{y} + \mathbf{z}^\top \mathbf{C} \mathbf{z} > 0$ unless both $\mathbf{y} = \mathbf{0}$ and $\mathbf{z} = \mathbf{0}$.

But if both \mathbf{B} and \mathbf{C} are only positive semi-definite, then $\mathbf{y}^\top \mathbf{B} \mathbf{y} + \mathbf{z}^\top \mathbf{C} \mathbf{z} \geq 0$, so \mathbf{A} is positive semi-definite. □

Outline

Quadratic Forms and Their Definiteness

Quadratic Forms and Symmetric Matrices

Definiteness of a Quadratic Form: Definition

Definiteness of a Quadratic Form: Some Simple Tests

The Two Variable Case

The Block Diagonal Case

Definiteness Tests by Symmetric Pivoting

Quadratic Form Invariance

Symmetric Straightforward Pivoting

Adjusted Symmetric Pivoting

Sylvester's Criterion for (Semi-)Definite Quadratic Forms

Sylvester's Criterion: Statement and Preliminary Examples

Proving Sylvester's Criterion

Quadratic Form Invariance: Statement of Lemma

Lemma

Suppose that \mathbf{A} and \mathbf{B} are $n \times n$ symmetric matrices,
and there exists an invertible $n \times n$ matrix \mathbf{R} such that $\mathbf{B} = \mathbf{R}\mathbf{A}\mathbf{R}^\top$.

Then the definiteness function

$$\mathcal{Q}_{n \times n} \ni \mathbf{Q} \mapsto \text{def}(\mathbf{Q}) \in \{PD, ND, PSD, NSD, ID\}$$

satisfies $\text{def}(\mathbf{B}) = \text{def}(\mathbf{A})$.

Quadratic Form Invariance: Proof of Lemma

Proof.

For any $\mathbf{x}, \mathbf{u} \in \mathbb{R}^n$ with $\mathbf{x} = \mathbf{R}^T \mathbf{u}$ and so $\mathbf{u} = (\mathbf{R}^T)^{-1} \mathbf{x}$, note that for $\mathbf{B} = \mathbf{R} \mathbf{A} \mathbf{R}^T$ one has

1. $\mathbf{x}^T \mathbf{A} \mathbf{x} = (\mathbf{R}^T \mathbf{u})^T \mathbf{A} \mathbf{R}^T \mathbf{u} = \mathbf{u}^T \mathbf{R} \mathbf{A} \mathbf{R}^T \mathbf{u} = \mathbf{u}^T \mathbf{B} \mathbf{u}$;
2. $\mathbf{x} = \mathbf{0} \iff \mathbf{u} = \mathbf{0}$ and so $\mathbf{x} \neq \mathbf{0} \iff \mathbf{u} \neq \mathbf{0}$.

From these two statements one can verify each of the following four equivalences:

$$\forall \mathbf{x} \neq \mathbf{0} : \mathbf{x}^T \mathbf{A} \mathbf{x} \begin{cases} > \\ < \\ \geq \\ \leq \end{cases} 0 \iff \forall \mathbf{u} \neq \mathbf{0} : \mathbf{u}^T \mathbf{B} \mathbf{u} \begin{cases} > \\ < \\ \geq \\ \leq \end{cases} 0$$

In addition, it also follows from these four equivalences that \mathbf{A} is indefinite if and only if \mathbf{B} is indefinite. □

Quadratic Form Invariance: Counter Example

Example

Suppose that \mathbf{A} and \mathbf{B} are $n \times n$ symmetric matrices, where \mathbf{A} is either positive or negative definite.

Suppose too that there exists a singular $n \times n$ matrix \mathbf{S} such that $\mathbf{B} = \mathbf{SAS}^\top$.

Then $|\mathbf{S}| = |\mathbf{S}^\top| = 0$, so \mathbf{S}^\top is also singular.

Hence there exists $\mathbf{y} \neq \mathbf{0}$ such that $\mathbf{S}^\top \mathbf{y} = \mathbf{0}$.

Then $\mathbf{y}^\top \mathbf{B} \mathbf{y} = \mathbf{y}^\top \mathbf{SAS}^\top \mathbf{y} = 0$.

It follows that \mathbf{B} is neither positive nor negative definite.

Outline

Quadratic Forms and Their Definiteness

Quadratic Forms and Symmetric Matrices

Definiteness of a Quadratic Form: Definition

Definiteness of a Quadratic Form: Some Simple Tests

The Two Variable Case

The Block Diagonal Case

Definiteness Tests by Symmetric Pivoting

Quadratic Form Invariance

Symmetric Straightforward Pivoting

Adjusted Symmetric Pivoting

Sylvester's Criterion for (Semi-)Definite Quadratic Forms

Sylvester's Criterion: Statement and Preliminary Examples

Proving Sylvester's Criterion

Symmetric Maximal Diagonalization: Definition

Definition

A **symmetric maximal diagonalization** of an $n \times n$ matrix \mathbf{A}

takes the form $\mathbf{R}\mathbf{A}\mathbf{R}^T = \begin{pmatrix} \mathbf{D}_{r \times r} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{pmatrix}$ where:

1. the integer $r \in \mathbb{Z}$ satisfying $0 \leq r \leq \min\{m, n\}$ is the rank;
2. $\mathbf{D}_{r \times r}$ is an $r \times r$ diagonal matrix which is invertible because all its r diagonal elements are non-zero;
3. \mathbf{R} is an invertible $n \times n$ matrix that, because $|\mathbf{R}| = 1$, represents a determinant preserving row operation.

In case $0 < r < n$, the symmetric maximal diagonalization of the $n \times n$ matrix \mathbf{A} needs the full expression for the 2×2 partitioned matrix on the right-hand side.

Otherwise, in case $r = n$, this partitioned matrix reduces to $\mathbf{D}_{n \times n}$.

Straightforward Symmetric Pivoting, First Step

Start with any $n \times n$ symmetric matrix \mathbf{A} , also denoted by $\mathbf{A}^{(0)}$, which we can write in partitioned form as $\mathbf{A} = \begin{pmatrix} a_{11} & \mathbf{a}_{>1,1}^\top \\ \mathbf{a}_{>1,1} & \mathbf{A}_{>1,>1} \end{pmatrix}$, where $\mathbf{a}_{>1,1}$ is a column $n - 1$ -vector with transpose $\mathbf{a}_{>1,1}^\top$.

Provided that $a_{11} \neq 0$, we can pivot symmetrically about a_{11} by:

1. pre-multiplying \mathbf{A} by the lower triangular downward pivot matrix $\mathbf{E}_{11}^\downarrow$;
2. post-multiplying the product $\mathbf{E}_{11}^\downarrow \mathbf{A}$ by the upper triangular transpose $(\mathbf{E}_{11}^\downarrow)^\top$ of the downward pivot matrix $\mathbf{E}_{11}^\downarrow$.

The combined effect is to transform \mathbf{A} to the symmetric matrix

$$\mathbf{A}^{(1)} = \mathbf{E}_{11}^\downarrow \mathbf{A} (\mathbf{E}_{11}^\downarrow)^\top = \begin{pmatrix} a_{11} & \mathbf{0}_{1,>1} \\ \mathbf{0}_{>1,1} & \mathbf{A}_{>1,>1}^{(1)} \end{pmatrix}$$

Straightforward Symmetric Pivoting, Start of Step k

For each $k \in \mathbb{N}$ with $1 < k < n$,
step k starts with the $n \times n$ matrix $\mathbf{A}^{(k-1)}$ which,
by induction on k , takes the symmetric form

$$\mathbf{A}^{(k-1)} = \begin{pmatrix} \mathbf{D}_{<k,<k}^{(k-1)} & \mathbf{0}_{<k,k} & \mathbf{0}_{<k,>k} \\ \mathbf{0}_{k,<k}^\top & a_{kk}^{(k-1)} & (\mathbf{a}_{>k,k}^{(k-1)})^\top \\ \mathbf{0}_{>k,<k} & \mathbf{a}_{>k,k}^{(k-1)} & \mathbf{A}_{>k,>k}^{(k-1)} \end{pmatrix}$$

where:

1. $\mathbf{D}_{<k,<k}^{(k-1)}$ is a $(k-1) \times (k-1)$ diagonal matrix;
2. $\mathbf{a}_{>k,k}^{(k-1)}$ is a column $n-k$ -vector;
3. $\mathbf{A}_{>k,>k}^{(k-1)}$ is an $(n-k) \times (n-k)$ symmetric matrix.

Straightforward Symmetric Pivoting, Step k

Starting from $\mathbf{A}^{(k-1)} = \begin{pmatrix} \mathbf{D}_{<k,<k}^{(k-1)} & \mathbf{0}_{<k,k} & \mathbf{0}_{<k,>k} \\ \mathbf{0}_{k,<k}^\top & a_{kk}^{(k-1)} & (\mathbf{a}_{>k,k}^{(k-1)})^\top \\ \mathbf{0}_{>k,<k} & \mathbf{a}_{>k,k}^{(k-1)} & \mathbf{A}_{>k,>k}^{(k-1)} \end{pmatrix},$

provided that the pivot element satisfies $a_{kk}^{(k-1)} \neq 0$, we can:

- ▶ pre-multiply by the lower triangular pivot matrix $\mathbf{E}_{kk}^\downarrow$;
- ▶ post-multiply by the upper triangular transpose $(\mathbf{E}_{kk}^\downarrow)^\top$.

The result of the symmetric pivot operation is $\mathbf{A}^{(k)} = \mathbf{E}_{kk}^\downarrow \mathbf{A}^{(k-1)} (\mathbf{E}_{kk}^\downarrow)^\top$ where

$$\mathbf{A}^{(k)} = \begin{pmatrix} \mathbf{D}_{\leq k, \leq k}^{(k)} & \mathbf{0}_{\leq k, k+1} & \mathbf{0}_{\leq k, > k+1} \\ \mathbf{0}_{k+1, \leq k}^\top & a_{k+1, k+1}^{(k)} & (\mathbf{a}_{>k+1, k+1}^{(k)})^\top \\ \mathbf{0}_{>k+1, \leq k} & \mathbf{a}_{>k+1, k+1}^{(k)} & \mathbf{A}_{>k+1, >k+1}^{(k)} \end{pmatrix}$$

The $k \times k$ diagonal matrix $\mathbf{D}_{\leq k, \leq k}^{(k)}$ is the diagonal matrix $\mathbf{D}_{<k, <k}^{(k-1)}$ with one extra non-zero pivot element $a_{kk}^{(k-1)}$ on the diagonal.

Conclusion of Straightforward Symmetric Pivoting

Provided that the successive pivot elements $a_{kk}^{(k)}$, for $k \in \mathbb{N}_n$, are all non-zero, straightforward symmetric pivoting can continue all the way until k reaches n .

In this case, after all the n stages, straightforward symmetric pivoting ends with the $n \times n$ matrix $\mathbf{A}^{(n)}$, which equals the diagonal matrix $\mathbf{D}_{\leq n, \leq n}^{(n)}$.

The last diagonal element $a_{nn}^{(n-1)}$ could be zero. This does not matter because, after step $n - 1$, no more pivoting is required.

But like downward pivoting, if $a_{kk}^{(k)} = 0$ for some $k \leq n$, then straightforward symmetric pivoting eventually fails. Some adjustment of at least one pivot element is needed.

Outline

Quadratic Forms and Their Definiteness

Quadratic Forms and Symmetric Matrices

Definiteness of a Quadratic Form: Definition

Definiteness of a Quadratic Form: Some Simple Tests

The Two Variable Case

The Block Diagonal Case

Definiteness Tests by Symmetric Pivoting

Quadratic Form Invariance

Symmetric Straightforward Pivoting

Adjusted Symmetric Pivoting

Sylvester's Criterion for (Semi-)Definite Quadratic Forms

Sylvester's Criterion: Statement and Preliminary Examples

Proving Sylvester's Criterion

Adjusted Symmetric Pivoting: The Matrix Sequence, I

Like straightforward pivoting, straightforward symmetric pivoting works provided each pivot element $a_{kk}^{(k-1)}$ with $k < n$ is non-zero.

Like adjusted pivoting, **adjusted** symmetric pivoting introduces sign-corrected permutation matrices to allow for the possibility that at least one relevant prospective pivot $a_{kk}^{(k-1)}$ with $k < n$ is 0.

If the adjusted symmetric pivoting process can last at least r steps, it will generate, for each $k \in \mathbb{N}_r$, an $n \times n$ symmetric matrix $\tilde{\mathbf{A}}^{(k)}$

which takes the partitioned form $\tilde{\mathbf{A}}^{(k)} = \begin{pmatrix} \tilde{\mathbf{D}}_{k \times k}^{(k)} & \mathbf{0}_{k \times (n-k)} \\ \mathbf{0}_{(n-k) \times k} & \tilde{\mathbf{A}}_{>k, >k}^{(k)} \end{pmatrix}$.

Here $\tilde{\mathbf{D}}_{k \times k}^{(k)}$ is the diagonal matrix in which, for each $p \in \mathbb{N}_k$, the non-zero diagonal element $\tilde{u}_{pp}^{(k)}$ is the adjusted pivot element $\tilde{a}_{pp}^{(p-1)}$ that was used at stage p .

Adjusted Symmetric Pivoting: The Matrix Sequence, II

Suppose that the adjusted symmetric pivoting process lasts exactly r steps, where $r \in \mathbb{N}_n$.

Starting from $\tilde{\mathbf{A}}^{(0)} = \mathbf{A}$, each successive step $k \in \mathbb{N}_r$, after adjusting any non-zero pivots, makes symmetric double use of the determinant preserving downward pivoting operation $\tilde{\mathbf{E}}_{kk}^\downarrow$.

The symmetric double application of $\tilde{\mathbf{E}}_{kk}^\downarrow$ to the matrix $\tilde{\mathbf{A}}^{(k-1)}$ leads to the symmetric matrix $\tilde{\mathbf{A}}^{(k)} = \tilde{\mathbf{E}}_{kk}^\downarrow \tilde{\mathbf{A}}^{(k-1)} (\tilde{\mathbf{E}}_{kk}^\downarrow)^\top$.

By induction on k , for each $k \in \mathbb{N}_r$ one has $\tilde{\mathbf{A}}^{(k)} = \mathbf{R}^{(k)} \mathbf{A} (\mathbf{R}^{(k)})^\top$ where $\mathbf{R}^{(k)} = \prod_{q=0}^{k-1} \tilde{\mathbf{E}}_{k-q, k-q}^\downarrow = \tilde{\mathbf{E}}_{kk}^\downarrow \tilde{\mathbf{E}}_{k-1, k-1}^\downarrow \cdots \tilde{\mathbf{E}}_{22}^\downarrow \tilde{\mathbf{E}}_{11}^\downarrow$, multiplied in that specific order.

Pivoting ceases with the matrix $\tilde{\mathbf{A}}^{(r)} = \mathbf{R}^{(r)} \mathbf{A} (\mathbf{R}^{(r)})^\top$.

Note that $\mathbf{R}^{(k)}$ is invertible as the product of invertible matrices.

Because of quadratic form invariance, the definiteness of successive matrices satisfies $\text{def}(\mathbf{A}) = \text{def}(\tilde{\mathbf{A}}^{(k)}) = \text{def}(\tilde{\mathbf{A}}^{(r)})$ for all $k \in \mathbb{N}_r$.

Adjusted Symmetric Pivoting: The End, Case 1

The $(k + 1)$ th step starts from $\tilde{\mathbf{A}}^{(k)} = \begin{pmatrix} \tilde{\mathbf{D}}_{k \times k}^{(k)} & \mathbf{0}_{k \times (n-k)} \\ \mathbf{0}_{(n-k) \times k} & \tilde{\mathbf{A}}_{>k, >k}^{(k)} \end{pmatrix}$.

Case 1: If the bottom right submatrix $\tilde{\mathbf{A}}_{>k, >k}^{(k)} = \mathbf{0}_{(n-k) \times (n-k)}$, then the $(k + 1)$ th pivot step is impossible.

All the r pivoting steps that are possible have been completed.

The final matrix takes the form $\begin{pmatrix} \mathbf{D}_{r \times r} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{pmatrix}$, where $\mathbf{D}_{r \times r}$ is an invertible $r \times r$ diagonal matrix with $r < n$.

Because $r < n$, quadratic form invariance implies that the original symmetric matrix \mathbf{A} cannot be definite.

It is positive or negative semi-definite according as the (non-zero) diagonal elements of $\mathbf{D}_{r \times r}$ are all positive or all negative — that is, according as $\mathbf{D}_{r \times r}$ is positive or negative definite.

But if the diagonal of $\mathbf{D}_{r \times r}$ has both positive and negative elements, then \mathbf{A} is indefinite.

Adjusted Symmetric Pivoting: The End, Case 2

The $(k + 1)$ th step starts from $\tilde{\mathbf{A}}^{(k)} = \begin{pmatrix} \tilde{\mathbf{D}}_{k \times k}^{(k)} & \mathbf{0}_{k \times (n-k)} \\ \mathbf{0}_{(n-k) \times k} & \tilde{\mathbf{A}}_{>k, >k}^{(k)} \end{pmatrix}$.

Case 2: Suppose that in the bottom right submatrix $\tilde{\mathbf{A}}_{>k, >k}^{(k)}$, at least two of the diagonal elements $(\tilde{a}_{qq}^{(k)})_{q=k+1}^n$ are zero, even though $\tilde{\mathbf{A}}_{>k, >k}^{(k)} \neq \mathbf{0}_{(n-k) \times (n-k)}$.

Then there exist a pair $p, q \in \mathbb{N}$ with $k < p < q \leq n$ such that $\tilde{a}_{pp}^{(k)} = \tilde{a}_{qq}^{(k)} = 0$ and yet $\tilde{a}_{pq}^{(k)} = \tilde{a}_{qp}^{(k)} \neq 0$.

So the simple test for semi-definiteness implies that the symmetric matrix $\tilde{\mathbf{A}}^{(k)}$ is indefinite.

By quadratic form invariance, so is the original matrix \mathbf{A} .

Adjusted Symmetric Pivoting: How to Adjust

The $(k + 1)$ th step starts from $\tilde{\mathbf{A}}^{(k)} = \begin{pmatrix} \tilde{\mathbf{D}}_{k \times k} & \mathbf{0}_{k \times (n-k)} \\ \mathbf{0}_{(n-k) \times k} & \tilde{\mathbf{A}}_{>k, >k}^{(k)} \end{pmatrix}$.

Case 3: Suppose that $\tilde{a}_{k+1, k+1}^{(k)} = 0$ but the matrix $\tilde{\mathbf{A}}_{>k, >k}^{(k)}$ has at least one non-zero diagonal element $\tilde{a}_{qq}^{(k)}$ with $q > k + 1$. That is, there exists at least non-zero element $\tilde{a}_{qq}^{(k)}$ on the part of the diagonal below and to the right of $\tilde{a}_{k+1, k+1}^{(k)}$.

We adjust the pivot symmetrically along the diagonal by applying one sign corrected swap matrix along with its transpose:

1. first, we pre-multiply $\tilde{\mathbf{A}}^{(k)}$ by the $n \times n$ matrix $\hat{\mathbf{T}}_{n \times n}^{q, k+1}$, which first swaps rows q and $k + 1$, then corrects the sign;
2. then we post-multiply $\hat{\mathbf{T}}_{n \times n}^{q, k+1} \tilde{\mathbf{A}}^{(k)}$ by the $n \times n$ transposed matrix $(\hat{\mathbf{T}}_{n \times n}^{q, k+1})^\top$, which first swaps columns q and $k + 1$, then corrects the sign.

Adjusted Symmetric Pivoting: The Next Step

The $(k + 1)$ th step starts from $\tilde{\mathbf{A}}^{(k)} = \begin{pmatrix} \tilde{\mathbf{D}}_{k \times k}^{(k)} & \mathbf{0}_{k \times (n-k)} \\ \mathbf{0}_{(n-k) \times k} & \tilde{\mathbf{A}}_{>k, >k}^{(k)} \end{pmatrix}$.

Together, the two sign-corrected swaps $\hat{\mathbf{T}}_{n \times n}^{q, k+1}$ and $(\hat{\mathbf{T}}_{n \times n}^{q, k+1})^\top$ move the original non-zero element $\tilde{a}_{qq}^{(k)}$ in $\tilde{\mathbf{A}}^{(k)}$ up left to the $k + 1, k + 1$ position in the adjusted matrix $\hat{\mathbf{T}}_{n \times n}^{q, k+1} \tilde{\mathbf{A}}^{(k)} (\hat{\mathbf{T}}_{n \times n}^{q, k+1})^\top$.

These prior sign-corrected swaps of both rows and columns q and $k + 1$ allow us to apply the standard symmetric pivoting operation based on $\mathbf{E}_{k+1, k+1}^\downarrow$ to this new adjusted version of the matrix $\tilde{\mathbf{A}}^{(k)}$

The result of this $(k + 1)$ th adjusted symmetric pivot step is the next matrix $\tilde{\mathbf{A}}^{(k+1)} = \tilde{\mathbf{E}}_{k+1, k+1}^\downarrow \tilde{\mathbf{A}}^{(k)} (\tilde{\mathbf{E}}_{k+1, k+1}^\downarrow)^\top$ where $\tilde{\mathbf{E}}_{k+1, k+1}^\downarrow$ is the adjusted pivot matrix $\mathbf{E}_{k+1, k+1}^\downarrow \hat{\mathbf{T}}_{n \times n}^{q, k+1}$.

How Adjusted Symmetric Pivoting Ends: Case A

Given an $n \times n$ symmetric matrix \mathbf{A} , adjusted symmetric pivoting can go on through steps $k = 1, 2, \dots, r$ until it reaches a terminal symmetric matrix $\tilde{\mathbf{A}}^{(r)}$ with $r \leq n$.

There are two possible cases.

Case A: Symmetric pivoting may end after r steps with

$$\tilde{\mathbf{A}}^{(r)} = \begin{pmatrix} \mathbf{D}_{r \times r} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{0}_{(n-r) \times (n-r)} \end{pmatrix} = \mathbf{diag}(a_{11}^{(r)}, \dots, a_{rr}^{(r)}, \mathbf{0}_{n-r})$$

where $a_{kk}^{(r)}$ is the non-zero k th pivot element, for all $k \in \mathbb{N}_r$, and so $\mathbf{D}_{r \times r}$ is an invertible $r \times r$ diagonal matrix.

Then the definiteness $\text{def}(\mathbf{A})$ of the original matrix is the same as the definiteness $\text{def}(\tilde{\mathbf{A}}^{(r)})$ of the diagonal matrix, which is easy to determine.

How Adjusted Symmetric Pivoting Ends: Case B

Case B: Alternatively symmetric pivoting may end after r steps

$$\text{with } \tilde{\mathbf{A}}^{(r)} = \begin{pmatrix} \mathbf{D}_{r \times r} & \mathbf{0}_{r \times (n-r)} \\ \mathbf{0}_{(n-r) \times r} & \mathbf{S}_{(n-r) \times (n-r)} \end{pmatrix}$$

where $\mathbf{D}_{r \times r} = \mathbf{diag}(a_{11}^{(r)}, \dots, a_{rr}^{(r)})$ whose element $a_{kk}^{(r)}$ is the non-zero k th pivot element, for each $k \in \mathbb{N}_r$,

and $\mathbf{S}_{(n-r) \times (n-r)}$ is a non-zero symmetric matrix whose diagonal elements are all zero.

In this case too the definiteness $\text{def}(\mathbf{A})$ of the original matrix equals the definiteness of the non-diagonal matrix $\text{def}(\tilde{\mathbf{A}}^{(r)})$.

But in this case $\tilde{\mathbf{A}}^{(r)}$ is always indefinite.

Curtailing the Symmetric Pivoting

After k steps, symmetric pivoting reaches a partial diagonalization of the form $\tilde{\mathbf{A}}^{(k)} = \begin{pmatrix} \tilde{\mathbf{D}}_{k \times k}^{(k)} & \mathbf{0}_{k \times (n-k)} \\ \mathbf{0}_{(n-k) \times k} & \tilde{\mathbf{A}}_{>k, >k}^{(k)} \end{pmatrix}$.

The simple tests of definiteness we discussed earlier imply that, in case the diagonal matrix $\tilde{\mathbf{D}}_{k \times k}^{(k)}$ has:

1. two non-zero diagonal elements of different signs, both it and the original symmetric matrix are indefinite;
2. any zero element, neither it nor the original symmetric matrix can be either positive definite or negative definite.

These properties may allow a test of definiteness to be curtailed before the symmetric pivoting process has been completed.

Outline

Quadratic Forms and Their Definiteness

Quadratic Forms and Symmetric Matrices

Definiteness of a Quadratic Form: Definition

Definiteness of a Quadratic Form: Some Simple Tests

The Two Variable Case

The Block Diagonal Case

Definiteness Tests by Symmetric Pivoting

Quadratic Form Invariance

Symmetric Straightforward Pivoting

Adjusted Symmetric Pivoting

Sylvester's Criterion for (Semi-)Definite Quadratic Forms

Sylvester's Criterion: Statement and Preliminary Examples

Proving Sylvester's Criterion

Notation for Relevant Principal Minors

Recall the earlier definitions of the principal and leading principal minors of a determinant.

Given any $n \times n$ symmetric matrix \mathbf{A} and any non-empty subset $K \subseteq \mathbb{N}_n$ with $k = \#K$, let:

1. $\mathbf{A}_{K \times K}$ denote the $k \times k$ matrix whose elements form the symmetric submatrix $(a_{ij})_{(i,j) \in K \times K}$ made up of the rows $i \in K$ and columns $j \in K$;
2. let $\Delta_k^K = |\mathbf{A}_{K \times K}|$ denote the corresponding principal minor of order k .

In case $K = \mathbb{N}_k = \{1, 2, \dots, k\}$, let D_k denote $\Delta_k^{\mathbb{N}_k}$, which is the unique leading principal minor of order k .

Sylvester's Criterion: General Statement

Theorem (Sylvester's criterion)

Any $n \times n$ symmetric matrix \mathbf{A}
and associated quadratic form $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ are both:

positive definite $\iff D_k > 0$ for all $k = 1, \dots, n$

positive semidefinite $\iff \Delta_k^K \geq 0$ for all Δ_k^K of any order k

negative definite $\iff (-1)^k D_k > 0$ for all $k = 1, \dots, n$

negative semidefinite $\iff (-1)^k \Delta_k^K \geq 0$ for all Δ_k^K of any order k

Otherwise the quadratic form $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ and matrix \mathbf{A} are indefinite.

Note that the conditions for \mathbf{A} to be negative (semi-) definite are exactly those for $-\mathbf{A}$ to be positive (semi-) definite.

The Case of a Quadratic Form in Two Variables

The general quadratic form in 2 variables is

$$(x, y) \begin{pmatrix} a & h \\ h & b \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = ax^2 + 2hxy + by^2$$

If it is positive definite, it is positive whenever $x \neq 0$ and $y = 0$.

This implies that $ax^2 > 0$ whenever $x \neq 0$,

which holds if and only if the first leading principal minor $a > 0$.

But if $a > 0$, then completing the square implies that

$$ax^2 + 2hxy + by^2 = a(x + hy/a)^2 + (b - h^2/a)y^2$$

Given that $a > 0$, this is positive definite if and only if $b > h^2/a$,

or iff the second leading principal minor $ab - h^2 = \begin{vmatrix} a & h \\ h & b \end{vmatrix} > 0$.

For the case of 2 variables, this proves

that the real-symmetric matrix \mathbf{A} is positive definite

if and only if all the leading principal minors of \mathbf{A} are positive.

Quadratic Form in Two Variables: Exercise

Exercise

For the case of a quadratic form in two variables, prove the other cases of Sylvester's criterion.

The Case of a Diagonal Quadratic Form

The general diagonal quadratic form in n variables is $\mathbf{x}^\top \Lambda \mathbf{x}$ where \mathbf{x} is an n -vector and Λ is an $n \times n$ diagonal matrix $\mathbf{diag}(\lambda_1, \dots, \lambda_n)$.

Then the quadratic form $\mathbf{x}^\top \Lambda \mathbf{x} = \sum_{i=1}^n \lambda_i (x_i)^2$ and matrix Λ are:

1. positive definite if and only if $\lambda_i > 0$ for $i = 1, 2, \dots, n$.

This is true **if and only if** the k -fold product $\prod_{i=1}^k \lambda_i$ is **positive** for each $k = 1, 2, \dots, n$.

But $\prod_{i=1}^k \lambda_i = |\mathbf{diag}(\lambda_1, \dots, \lambda_k)|$ is the leading principal minor D_k of order k for Λ .

2. positive semi-definite if and only if $\lambda_i \geq 0$ for $i = 1, 2, \dots, n$.

This is true **if and only if** the product $\prod_{i \in K} \lambda_i$ is **nonnegative** for every nonempty $K \subseteq \mathbb{N}_n = \{1, 2, \dots, n\}$.

But each product $\prod_{i \in K} \lambda_i$ equals the determinant $|\Lambda_{K \times K}|$ of the diagonal submatrix $\Lambda_{K \times K}$, which is the particular principal minor Δ_K^K of order $k = \#K$.

Toward the General Case

The formal proof of Sylvester's criterion for a general $n \times n$ symmetric matrix \mathbf{A} to be positive or negative definite will rely on:

1. showing that unadjusted symmetric pivoting, while it works, preserves each leading principal minor of \mathbf{A} ;
2. using unadjusted symmetric pivoting to reduce the general case to the case when \mathbf{A} is diagonal.

A similar argument allowing for adjusted symmetric pivoting will treat the case when \mathbf{A} is positive or negative semi-definite.

For large n ($n > 3?$), the best way to compute those minors, however, may well be to use symmetric pivoting ...

Outline

Quadratic Forms and Their Definiteness

Quadratic Forms and Symmetric Matrices

Definiteness of a Quadratic Form: Definition

Definiteness of a Quadratic Form: Some Simple Tests

The Two Variable Case

The Block Diagonal Case

Definiteness Tests by Symmetric Pivoting

Quadratic Form Invariance

Symmetric Straightforward Pivoting

Adjusted Symmetric Pivoting

Sylvester's Criterion for (Semi-)Definite Quadratic Forms

Sylvester's Criterion: Statement and Preliminary Examples

Proving Sylvester's Criterion

Symmetric Pivoting Preserves Leading Principal Minors

Given any $n \times n$ symmetric matrix \mathbf{A} and any $k \in \mathbb{N}_n$, let $\mathbf{A}_{\leq k, \leq k}$ denote the $k \times k$ matrix whose determinant $|\mathbf{A}_{\leq k, \leq k}|$ is the k th order leading principal minor.

Whenever $p < q \leq k$, the elementary row operation $\mathbf{A} \mapsto \mathbf{E}_{q+\alpha p} \mathbf{A}$ of adding α times row p to row q of \mathbf{A} preserves not only $|\mathbf{A}|$, but also each leading principal minor $|\mathbf{A}_{\leq k, \leq k}|$ when $\mathbf{E}_{q+\alpha p}$ is restricted to the $k \times k$ matrix $\mathbf{A}_{\leq k, \leq k}$.

The same property of **leading principal minor preservation** applies to each elementary column operation $\mathbf{A} \mapsto \mathbf{A} \mathbf{E}_{q+\alpha p}^T$.

From this, it follows that leading principal minor preservation also applies to the symmetric pivoting operation $\mathbf{A} \mapsto \mathbf{E}_{pp}^\downarrow \mathbf{A} (\mathbf{E}_{pp}^\downarrow)^T$ when it is restricted to $\mathbf{A}_{\leq k, \leq k}$, where $k > p$.

Proof by Induction: Key Ideas

We will prove Sylvester's criterion for a general $n \times n$ symmetric matrix \mathbf{A} .

Actually, we prove a superficially stronger necessary condition for \mathbf{A} to be positive definite:

all its principal minors, whether leading or not, must be positive.

The proof of this modified form of Sylvester's criterion will be by induction on n .

The result is trivial when $n = 1$ and $\mathbf{A} = (a_{11})$, whose only minor is $\det(a_{11}) = a_{11}$.

The induction hypothesis will be that Sylvester's modified criterion is valid for any $m \times m$ symmetric matrix \mathbf{A} .

The induction step will be to prove that if Sylvester's modified criterion is valid for every $(n - 1) \times (n - 1)$ symmetric matrix, then it is valid for every $n \times n$ symmetric matrix.

Proof by Induction in Four Parts

To repeat, the induction step will be to prove that if Sylvester's modified criterion is valid for every $(n - 1) \times (n - 1)$ symmetric matrix, then it is valid for every $n \times n$ symmetric matrix.

This induction step has to be proved four times for Sylvester's:

1. modified necessary condition for a positive definite matrix;
2. sufficient condition for a positive definite matrix;
3. necessary condition for a positive semi-definite matrix;
4. sufficient condition for a positive semi-definite matrix.

Each of the four proofs will occupy two slides.

Recall that the criterion

for a negative definite or semi-definite symmetric matrix \mathbf{A}

is equivalent to the same criterion

for the positive definite or semi-positive symmetric matrix $-\mathbf{A}$.

1. Proving Necessity for a Positive Definite Matrix, I

- (a) Suppose the $n \times n$ symmetric matrix \mathbf{A} is positive definite.
- (b) We have already argued that $a_{11} > 0$, as a diagonal element.
- (c) So the downward pivoting matrix $\mathbf{E}_{11}^\downarrow$ is well defined and invertible.
- (d) Because $\mathbf{E}_{11}^\downarrow$ is invertible and \mathbf{A} is positive definite, so is the block diagonal matrix $\mathbf{E}_{11}^\downarrow \mathbf{A} (\mathbf{E}_{11}^\downarrow)^\top = \mathbf{diag}(a_{11}, \mathbf{B})$ where \mathbf{B} is the $(n-1) \times (n-1)$ symmetric submatrix that results from one round of symmetric pivoting.
- (e) It follows from (d) that the block \mathbf{B} is positive definite.
- (f) Because \mathbf{B} is positive definite, the induction hypothesis implies that each principal minor Δ_k^K of $|\mathbf{B}|$ is positive.

1. Proving Necessity for a Positive Definite Matrix, II

(g) From (f) it follows that every principal minor of $\mathbf{diag}(a_{11}, \mathbf{B})$ which does not include the diagonal element a_{11} must be positive.

(h) But apart from a_{11} by itself, all the other principal minors of $\mathbf{diag}(a_{11}, \mathbf{B})$ which do include the element a_{11} take the form $a_{11}\Delta_k^K$ where Δ_k^K is a principal minor of $|\mathbf{B}|$.

(i) Because $a_{11} > 0$, it follows from (f), (g) and (h) that every principal minor of $\mathbf{diag}(a_{11}, \mathbf{B})$ must be positive.

(j) But the matrix $\mathbf{E}_{11}^\downarrow$ is determinant preserving, so $\mathbf{E}_{11}^\downarrow \mathbf{A} (\mathbf{E}_{11}^\downarrow)^\top = \mathbf{diag}(a_{11}, \mathbf{B})$ has the same principal minors as \mathbf{A} , implying that all the principal minors of \mathbf{A} are also positive. \square

2. Proving Sufficiency for a Positive Definite Matrix, I

(a) Suppose that every leading principal minor of the $n \times n$ symmetric matrix \mathbf{A} is positive.

(b) Note that (a) implies in particular that the first leading principal minor satisfies $a_{11} > 0$.

(c) So the downward pivoting matrix $\mathbf{E}_{11}^\downarrow$ is well defined and determinant preserving.

(d) But (c) implies that \mathbf{A} has the same leading principal minors as the block diagonal matrix $\mathbf{E}_{11}^\downarrow \mathbf{A} (\mathbf{E}_{11}^\downarrow)^\top = \mathbf{diag}(a_{11}, \mathbf{B})$ where \mathbf{B} is the $(n-1) \times (n-1)$ symmetric submatrix that results from one round of symmetric pivoting.

(e) Evidently, the leading principal minors of $|\mathbf{diag}(a_{11}, \mathbf{B})|$ take the form $a_{11}, a_{11}D_1, \dots, a_{11}D_{n-1}$ where each D_k denotes the k th leading principal minor of $|\mathbf{B}|$.

2. Proving Sufficiency for a Positive Definite Matrix, II

(f) By the induction hypothesis, because (e) implies that all the leading principal minors of $|\mathbf{B}|$ are positive, the $(n - 1) \times (n - 1)$ symmetric matrix \mathbf{B} is positive definite.

(g) Then, because (b) implies that $a_{11} > 0$, it follows from (f) that $\mathbf{diag}(a_{11}, \mathbf{B})$ is positive definite.

(h) Finally, because $\mathbf{E}_{11}^\downarrow$ is invertible and (g) implies that $\mathbf{diag}(a_{11}, \mathbf{B}) = \mathbf{E}_{11}^\downarrow \mathbf{A} (\mathbf{E}_{11}^\downarrow)^\top$ is positive definite, it follows from quadratic form invariance that \mathbf{A} is also positive definite.

3. Proving Necessity for a Positive Semi-Definite Matrix, I

(a) Suppose the $n \times n$ symmetric matrix \mathbf{A} is positive semi-definite.

(b) In case all the diagonal elements of \mathbf{A} are zero, we must have $\mathbf{A} = \mathbf{0}_{n \times n}$, otherwise \mathbf{A} would be indefinite.

(c) In the trivial case when $\mathbf{A} = \mathbf{0}_{n \times n}$, all minors of $|\mathbf{A}|$ are zero.

(d) Otherwise there exists a diagonal element $a_{pp} \neq 0$, which is positive because \mathbf{A} is positive semi-definite.

(e) Let $\hat{\mathbf{T}}_{1p}$ denote the sign adjusted swap of rows 1 and p .

Use it to define an adjusted symmetric pivot operation that gives the symmetric matrix $\mathbf{E}_{11}^\downarrow \hat{\mathbf{T}}_{1p} \mathbf{A} (\mathbf{E}_{11}^\downarrow \hat{\mathbf{T}}_{1p})^\top = \mathbf{diag}(a_{pp}, \tilde{\mathbf{B}})$, where $\tilde{\mathbf{B}}$ is an $(n-1) \times (n-1)$ symmetric matrix.

(f) Because $\mathbf{E}_{11}^\downarrow \hat{\mathbf{T}}_{1p}$ is invertible, it follows from quadratic form invariance that positive semi-definiteness of \mathbf{A} implies the same for $\mathbf{diag}(a_{pp}, \tilde{\mathbf{B}})$, and so also for $\tilde{\mathbf{B}}$.

3. Proving Necessity for a Positive Semi-Definite Matrix, II

(g) Because $\tilde{\mathbf{B}}$ is positive semi-definite, the induction hypothesis implies that, for each $k \in \mathbb{N}_{n-1}$ and each $K \subseteq \mathbb{N}_{n-1}$ with $\#K = k$, the principal minor Δ_k^K of $|\tilde{\mathbf{B}}|$ is non-negative.

(h) Now each principal minor of $|\mathbf{diag}(a_{pp}, \tilde{\mathbf{B}})|$ that is not a principal minor of $|\tilde{\mathbf{B}}|$

must take the form $a_{pp} \Delta_k^K$ for some principal minor Δ_k^K of $|\tilde{\mathbf{B}}|$.

(i) But then $a_{pp} > 0$ by (d) and $\Delta_k^K \geq 0$ by (g), so (h) implies that every principal minor of $|\mathbf{diag}(a_{pp}, \tilde{\mathbf{B}})|$ is non-negative.

(j) Now $\mathbf{diag}(a_{11}, \tilde{\mathbf{B}}) = \mathbf{E}_{11}^\downarrow \hat{\mathbf{T}}_{1p} \mathbf{A} (\mathbf{E}_{11}^\downarrow \hat{\mathbf{T}}_{1p})^\top$

where $\hat{\mathbf{T}}_{1p}$ is a sign-preserving swap of two rows

and the downward pivot matrix $\mathbf{E}_{11}^\downarrow$ is determinant preserving.

It follows that there is an obvious bijection

between each of the $2^n - 1$ principal minors of $|\mathbf{diag}(a_{pp}, \tilde{\mathbf{B}})|$ and a unique corresponding principal minor of \mathbf{A} .

(k) From (i) and (j), each principal minor of \mathbf{A} is non-negative. \square

4. Proving Sufficiency for a Positive Semi-Definite Matrix, I

- (a) Suppose that every principal minor of the $n \times n$ symmetric matrix \mathbf{A} is non-negative.
- (b) In case all the diagonal elements of \mathbf{A} are zero, we must have $\mathbf{A} = \mathbf{0}_{n \times n}$, otherwise at least one principal minor of the symmetric \mathbf{A} would be negative.
- (c) In the trivial case when $\mathbf{A} = \mathbf{0}_{n \times n}$, the matrix \mathbf{A} is evidently positive semi-definite.
- (d) Otherwise there exists a non-zero diagonal element a_{pp} , which is positive because every principal minor of \mathbf{A} is ≥ 0 .
- (e) Let $\hat{\mathbf{T}}_{1p}$ denote the sign adjusted swap of rows 1 and p . Use it to define an adjusted symmetric pivot operation that gives the $n \times n$ symmetric matrix $\mathbf{E}_{11}^\downarrow \hat{\mathbf{T}}_{1p} \mathbf{A} (\mathbf{E}_{11}^\downarrow \hat{\mathbf{T}}_{1p})^\top = \mathbf{diag}(a_{pp}, \tilde{\mathbf{B}})$, where $a_{pp} > 0$ and $\tilde{\mathbf{B}}$ is an $(n-1) \times (n-1)$ symmetric matrix.

4. Proving Sufficiency for a Semi-Definite Matrix, II

(f) Because $\hat{\mathbf{T}}_{1\rho}$ is a sign-preserving swap of two rows whereas $\mathbf{E}_{11}^\downarrow$ is determinant preserving, there exists an obvious bijection between each of the principal minors of $|\mathbf{E}_{11}^\downarrow \hat{\mathbf{T}}_{1\rho} \mathbf{A} (\mathbf{E}_{11}^\downarrow \hat{\mathbf{T}}_{1\rho})^\top| = |\mathbf{diag}(a_{pp}, \tilde{\mathbf{B}})|$ and a unique corresponding principal minor of \mathbf{A} .

(g) Together (a) and (f) imply that each principal minor of $|\mathbf{diag}(a_{pp}, \tilde{\mathbf{B}})|$ is non-negative. So therefore is each principal minor of $|\tilde{\mathbf{B}}|$.

(h) By the induction hypothesis, (g) implies that the $(n-1) \times (n-1)$ matrix $\tilde{\mathbf{B}}$ is positive semi-definite.

(i) Because $a_{pp} > 0$, (h) implies that the $n \times n$ matrix $\mathbf{diag}(a_{pp}, \tilde{\mathbf{B}})$ is positive semi-definite.

(j) But $\mathbf{diag}(a_{pp}, \tilde{\mathbf{B}}) = \mathbf{E}_{11}^\downarrow \hat{\mathbf{T}}_{1\rho} \mathbf{A} (\mathbf{E}_{11}^\downarrow \hat{\mathbf{T}}_{1\rho})^\top$ where $\mathbf{E}_{11}^\downarrow \hat{\mathbf{T}}_{1\rho}$ is invertible.

(k) By quadratic form invariance, together (i) and (j) imply that \mathbf{A} is positive semi-definite. □

Envoi

Though Sylvester's Criterion has been proved, remember it is here only because it is in various textbooks, including ours.

To establish the definiteness of a symmetric matrix, especially if it is larger than 3×3 , one can and should use symmetric pivoting first.

Key reference for idea of symmetric pivoting:

Paul Binding (1991) "Simple Tests for Classifying Critical Points of Quadratics with Linear Constraints"

American Mathematical Monthly 98 (10): 949–954.

As its main topic, this paper also considers conditions for a quadratic form $\mathbf{x}^\top \mathbf{A} \mathbf{x}$ to be positive (semi-)definite subject to a constraint $\mathbf{K} \mathbf{x} = \mathbf{0}$ — in the sense that $\mathbf{x}^\top \mathbf{A} \mathbf{x} > (\geq) 0$ for all $\mathbf{x} \in \mathbb{R}^n \setminus \{\mathbf{0}_n\}$ that satisfy $\mathbf{K} \mathbf{x} = \mathbf{0}$.

The relevant tests involve "bordered Hessians".

We can move on at last!