6 Positive Definite Matrices （正定矩阵）

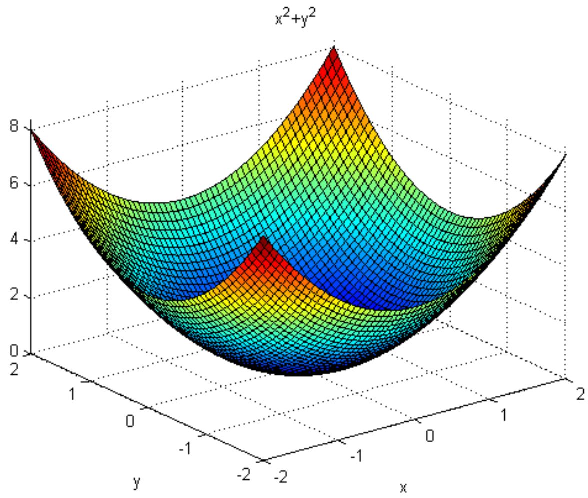
6.1

MINIMA, MAXIMA AND SADDLE POINTS

(最小值、最大值和鞍点)

Definite vs. Indefinite

Quadratic forms (二次型)



- *The signs of the eigenvalues* are often crucial.
- The new and highly important problem is to recognize a *minimum point*. This arises throughout science and engineering and every problem of optimization (优化).
- We will find a test that can be applied directly to $\boldsymbol{A}$, which will *guarantee that all those eigenvalues are positive* (*negative*, …).
- The test brings together three of the most basic ideas — ***pivots, determinants,*** and ***eigenvalues***.

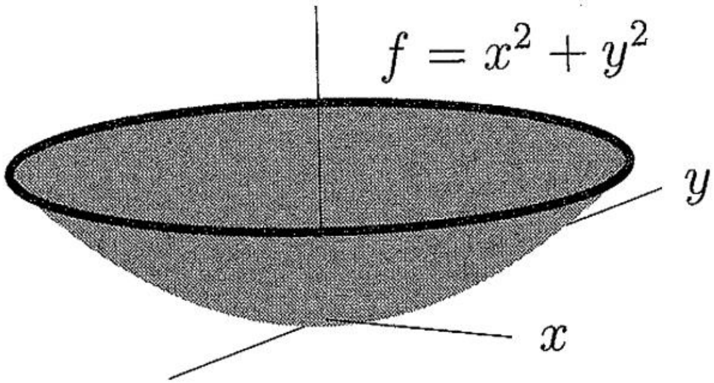
I. Definite vs Indefinite

Every ***quadratic form*** (二次型)

$$f(x,y) = ax^2 + 2bxy + cy^2$$

has a stationary point at the origin, where $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0.$

A local minimum *would* also be a global minimum. The surface $z = f(x,y)$ will then be shaped like a bowl, resting on the origin.



a bowl

When $f(x,y)$ is *strictly positive* at all other points (the bowl goes up), it is called **positive definite (正定)**, i.e.,

$$f(x,y) > 0, \forall (x,y) \neq (0,0).$$

I. Definite vs Indefinite

- **Question.** *What conditions on a, b, and c ensure that the quadratic* $f(x,y) = ax^2 + 2bxy + cy^2$ *is positive definite*?

- **(i)** If $f(x,y)$ is positive definite, then necessarily $a > 0$.

  (fix $y = 0$ and look in the x direction where $f(x,0) = ax^2$)

- **(ii)** If $f(x,y)$ is positive definite, then necessarily $c > 0$.

  (fix $x = 0$ and look in the y direction where $f(0,y) = cy^2$)

- $a > 0$ *and* $c > 0$ *do not guarantee that* $f(x,y)$ *is always positive*, a large cross term $2bxy$ can pull the graph below zero.

For instance, $f(x,y) = x^2 - 10xy + y^2$.

Here $a = 1$ and $c = 1$ are both positive, but f is not positive definite, because $f(1,1) = -8$. The conditions $a > 0$ and $c > 0$ ensure that $f(x,y)$ is positive on the x and y axes. But this function is negative on the line $x = y$, because $b = -5$ overwhelms a and c.

- **Question.** *What conditions on a, b, and c ensure that the quadratic $f(x,y) = ax^2 + 2bxy + cy^2$ is positive definite?*
- *(continued) $b > 0$ does not guarantee that $f(x,y)$ is always positive.*

For instance, in $f(x,y) = 2x^2 + 4xy + y^2$, $2b = 4 > 0$, this does not ensure a minimum, the sign of b is not important.

f does not have a minimum at $(0,0)$ because

$$f(1,-1) = 2 - 4 + 1 = -1.$$

What really matters?

- ***<u>It is the size of b, compared to a and c, that must be controlled.</u>***

The simplest technique is to complete the square:

$$f(x,y) = ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2$$

- <u>**(iii)** If $f(x,y)$ stays positive, then necessarily $ac > b^2$.</u>
- The conditions $a > 0$ and $ac > b^2$ guarantee $c > 0$.

- **Question.** *What conditions on a, b, and c ensure that the quadratic $f(x,y) = ax^2 + 2bxy + cy^2$ is positive definite?*

The simplest technique is to complete the square:

$$f(x,y) = ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2$$

- **Theorem (Test for a minimum).**

$ax^2 + 2bxy + cy^2$ is positive definite if and only if

$$a > 0 \text{ and } ac > b^2.$$

- **Remark (Test for a maximum).**

f has a maximum whenever $-f$ has a minimum.

We reverse the signs of a, b and c.

The quadratic form $ax^2 + 2bxy + cy^2$ is **negative definite (负定)** if and only if $a < 0$ and $ac > b^2$.

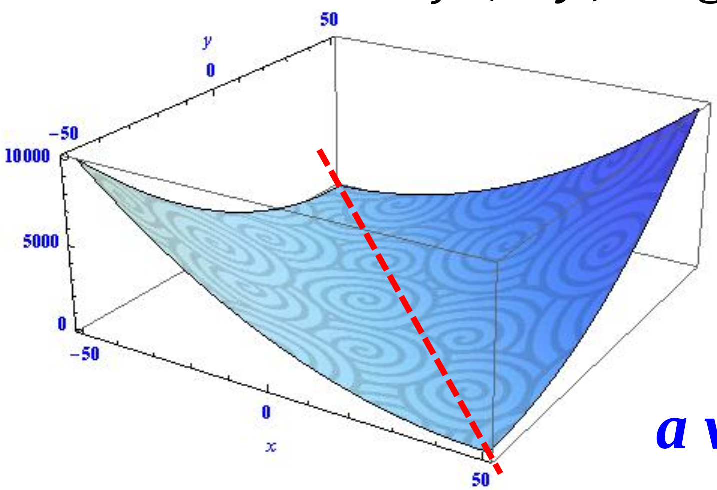
- **Singular case $ac = b^2$.**

$$f = ax^2 + 2bxy + cy^2 = a\left(x + \frac{b}{a}y\right)^2 + \left(c - \frac{b^2}{a}\right)y^2 = a\left(x + \frac{b}{a}y\right)^2$$

The quadratic form is **positive semidefinite (半正定)** when $a > 0$.

The quadratic form is **negative semidefinite (半负定)** when $a < 0$.

- **Remark.**
- The prefix *semi* allows the possibility that f can equal zero, as it will at the point $x = b, y = -a$.
- The surface $z = f(x, y)$ degenerates from a bowl into a valley.



$f(x,y) = (x+y)^2$

a valley

The valley runs along the line $x + y = 0$.

- **Saddle Point $ac < b^2$.**
- **Example. Saddle points at** $(0,0)$

$$f_1 = 2xy \quad \text{and} \quad f_2 = x^2 - y^2 \quad \text{and} \quad ac - b^2 = -1$$

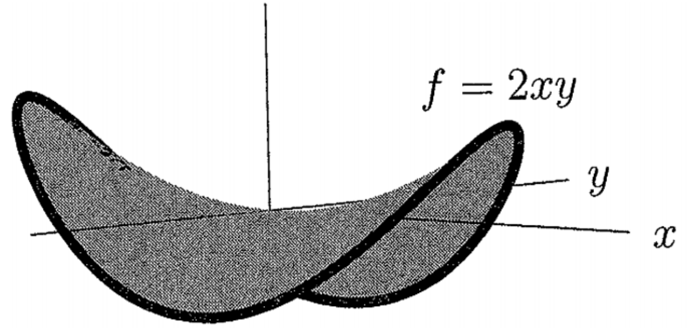
In f_1, $b = 1$ dominates $a = c = 0$.

In f_2, $a = 1$ and $c = -1$ have opposite sign.

These quadratic forms are **indefinite (不定)**, because they can take either sign. So we have a stationary point that is neither a maximum or a minimum. It is called a ***saddle point*** **(鞍点)**.

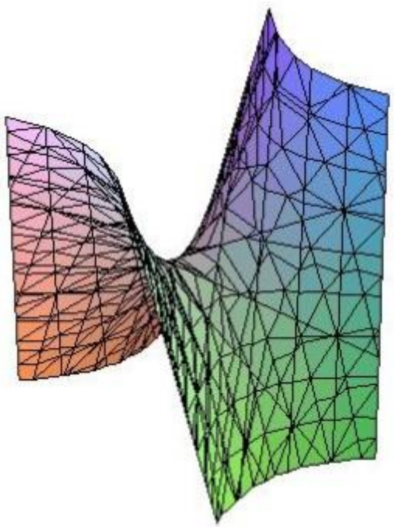
- **Remark.** *The saddles $2xy$ and $x^2 - y^2$ are practically the same, if we turn one through 45^{o} we get the other.*

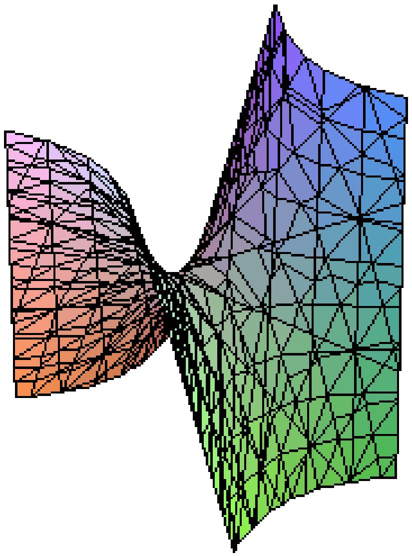
a saddle

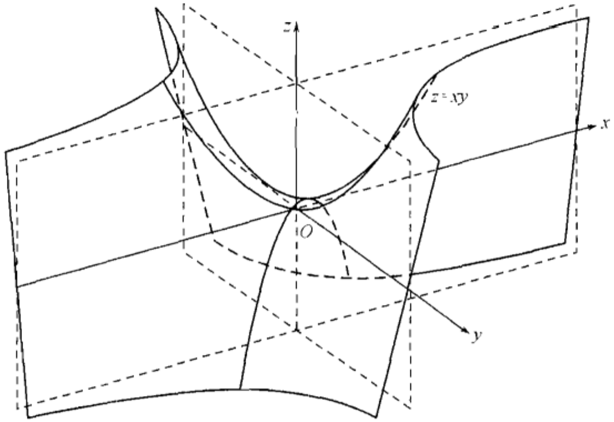


Saddle（马鞍面）

$$z = xy$$







II. Quadratic Forms (二次型) & Real Symmetric Matrices

A quadratic form $f(x, y) = ax^2 + 2bxy + cy^2$ comes directly from a symmetric 2 by 2 matrix $\boldsymbol{A}$:

$$ax^2 + 2bxy + cy^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}$$

where $\boldsymbol{x} = \begin{bmatrix} x \\ y \end{bmatrix}$.

For example, $4x^2 + 2xy - 3y^2 = \begin{bmatrix} x & y \end{bmatrix} \begin{bmatrix} 4 & 1 \\ 1 & -3 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$.

It generalizes immediately to n dimensions. (We will only discuss *real* case: $\mathbf{R}^n$)

When the variables are $x_1, \dots, x_n$, they go into a column vector $\boldsymbol{x} = (x_1, \dots, x_n)^{\mathrm{T}} \in \mathbf{R}^n$.

For any real symmetric matrix $\boldsymbol{A}$, the product $\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}$ is a pure quadratic form $f(x_1, \dots, x_n)$:

$$f(x_1, x_2, \cdots, x_n) = \sum_{i=1}^{n} a_{ii} x_i^2 + \sum_{1 \le i < j \le n} 2\, a_{ij} x_i x_j$$

$$= a_{11} x_1^2 + \boxed{2a_{12} x_1 x_2} + 2a_{13} x_1 x_3 + \cdots + 2a_{1n} x_1 x_n$$

$$a_{22} x_2^2 + 2a_{23} x_2 x_3 + \cdots + 2a_{2n} x_2 x_n$$

$$+ \cdots\cdots\cdots$$

$$+ a_{nn} x_n^2$$

Let $a_{ij} = a_{ji} \, (1 \le i < j \le n)$, then

$$f(x_1, x_2, \cdots, x_n) = a_{11} x_1^2 + \boxed{a_{12} x_1 x_2} + \cdots + a_{1n} x_1 x_n$$

$$+ \boxed{a_{21} x_2 x_1} + a_{22} x_2^2 + \cdots + a_{2n} x_2 x_n$$

$$+ \cdots\cdots\cdots$$

$$+ a_{n1} x_n x_1 + a_{n2} x_n x_2 + \cdots + a_{nn} x_n^2$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j .$$

For any real symmetric matrix $\boldsymbol{A} \in \mathbf{R}^{n\times n}$, *the product* $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ *is a pure quadratic form* $f(x_1, \dots, x_n)$:

$$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = [x_1 \quad x_2 \quad \cdots \quad x_n]\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}x_i x_j$$

(with $a_{ij} = a_{ji}$, i.e., $\boldsymbol{A} = \boldsymbol{A}^{\mathrm{T}}$; $\boldsymbol{x} = (x_1, \dots, x_n)^{\mathrm{T}} \in \mathbf{R}^n$)

The matrix $\boldsymbol{A}$ is called the **matrix of the quadratic form**.

For example, $f(x_1, x_2, x_3) = 2x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_2x_3 + 2x_3^2$:

$$f = [x_1 \quad x_2 \quad x_3]\begin{bmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}.$$

Example 1 Let

$$f(x_1, x_2, x_3, x_4) = 2x_1^2 + x_1x_2 + 2x_1x_3 + 4x_2x_4 + x_3^2 + 5x_4^2$$

The corresponding matrix for this quadratic form is

$$\boldsymbol{A} = \begin{bmatrix} 2 & \frac{1}{2} & 1 & 0 \\ \frac{1}{2} & 0 & 0 & 2 \\ 1 & 0 & 1 & 0 \\ 0 & 2 & 0 & 5 \end{bmatrix}$$

For

$$f(x_1, x_2, x_3) = 2x_1x_2 + 4x_1x_3 - 10x_2x_3 \qquad \begin{bmatrix} 0 & 1 & 2 \\ 1 & 0 & -5 \\ 2 & -5 & 0 \end{bmatrix}$$

$$f(x_1, x_2, x_3, x_4) = x_1^2 + 2x_2^2 - x_3^2 + 5x_4^2 \qquad \begin{bmatrix} 1 & & & \\ & 2 & & \\ & & -1 & \\ & & & 5 \end{bmatrix}$$

Definition 1 For a *real* quadratic form $f(x_1, x_2, \cdots, x_n) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ in n variables, if $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0$ for *any* $\boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}} \neq \mathbf{0}$ ($\boldsymbol{x} \in \mathbf{R}^n$), then $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ is called a **positive definite quadratic form**, and the corresponding real symmetric matrix $\boldsymbol{A}$ is called a **positive definite matrix**.

如果 n 元实二次型 $f(x_1, x_2, \cdots, x_n) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$,
$\forall \boldsymbol{x} = (x_1, x_2, \cdots, x_n)^{\mathrm{T}} \neq \mathbf{0}$ ($\boldsymbol{x} \in \mathbf{R}^n$), 恒有 $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0$,
就称$\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ 为**正定二次型**; 称实对称矩阵$\boldsymbol{A}$为**正定矩阵**.

For example, $f(x, y, z) = x^2 + 4y^2 + 16z^2$

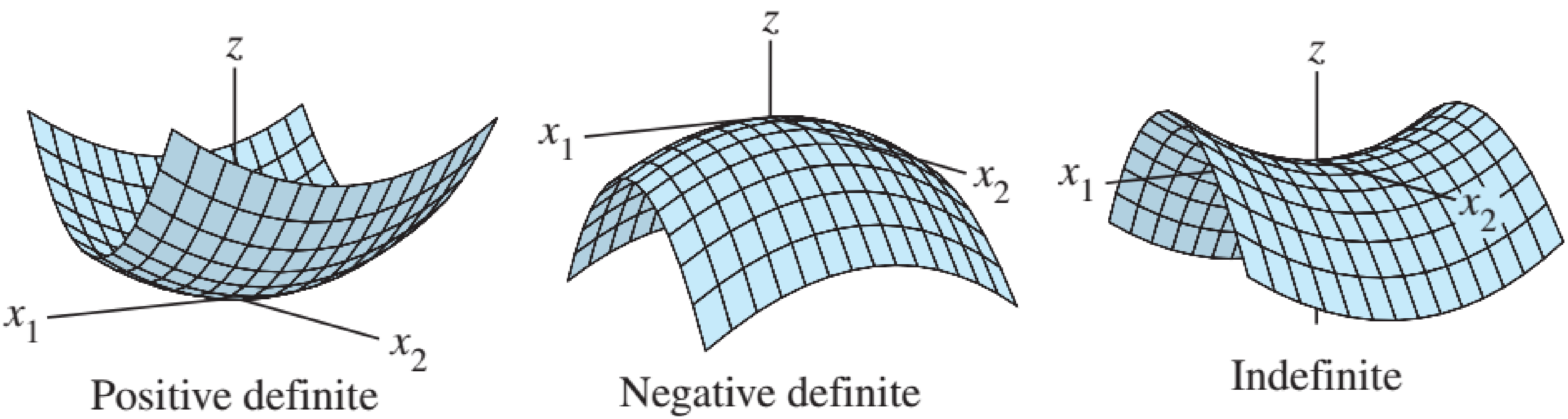
is a positive definite quadratic form.

$$\begin{bmatrix} 1 & & \\ & 4 & \\ & & 16 \end{bmatrix}$$

*Check the **eigenvalue**, **determinant**, **pivot**, …?*

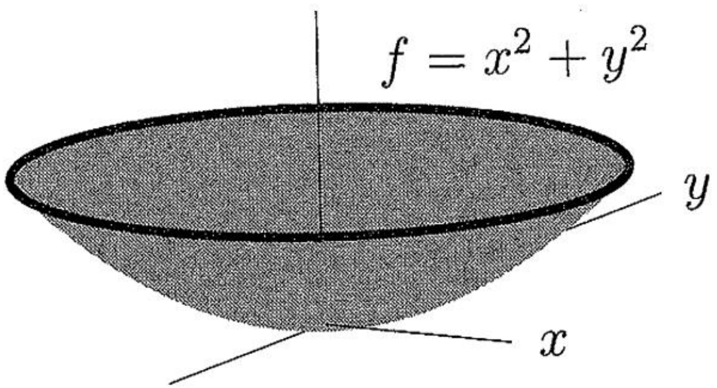
When $\boldsymbol{A}$ is an $n \times n$ matrix, the quadratic form $f(\boldsymbol{x}) = \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}$ is a real-valued function with domain $\mathbf{R}^n$.

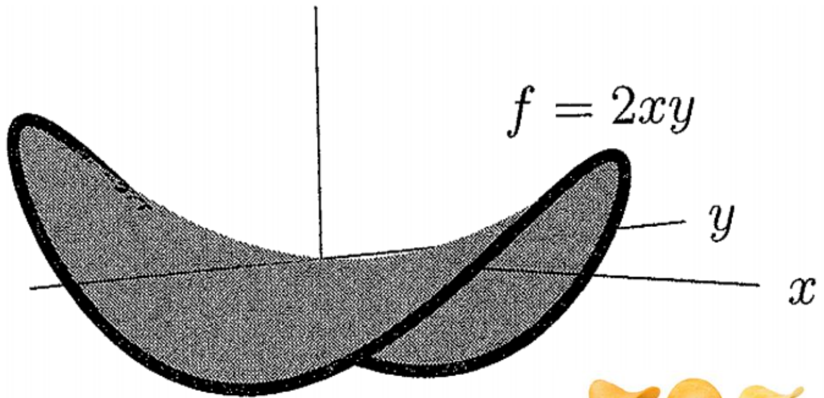
We distinguish several important classes of quadratic forms by the type of their values.



正定　　负定　　不定

Graphs of quadratic forms





a bowl

a saddle

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

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

Positive definite (正定)

Indefinite (不定)

Which symmetric matrices have the property that $\boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} > 0$ *for all nonzero vectors* $\boldsymbol{x}$?

Key words:

Definite vs Indefinite

Quadratic Forms and Real Symmetric Matrices

Homework

See Blackboard