

Positive Definite Matrices (正定矩阵)

Lecture 27

Dept. of Math.

2024.5

Minima, Maxima, and Saddle Points

- 1 Introduction
- 2 Minima, Maxima, and Saddle Points
- 3 Definite versus Indefinite: Bowl versus Saddle
- 4 Higher Dimensions: Linear Algebra
- 5 Homework Assignment 27

Minimum Point, Positive Definiteness

Questions:

1. Given a matrix, how could you determine the signs of the eigenvalues?
2. What test can guarantee that all of the eigenvalues of a matrix are positive?
3. How to find a minimum point efficiently? This arises throughout science and engineering and every problem of optimization. The mathematical problem is to move the second derivative test $F'' > 0$ into n dimensions.

Examples

Here are two examples:

* 小高數的
相遇

$$F(x,y) = 7 + 2(x+y)^2 - y \sin y - x^3$$

$$f(x,y) = 2x^2 + 4xy + y^2$$

兩函數在(0,0)

一階、二階導完全相同

$f(x,y)$ 极值

何時

一令以 $f(x,y)$ 的情況

Does either $F(x,y)$ or $f(x,y)$ have a minimum at the point $x = y = 0$?

$$\frac{\partial F}{\partial x} \Big|_{(0,0)} = 4(x+y) - 3x^2 \Big|_{(0,0)} = 0$$

$$\begin{aligned} \frac{\partial F}{\partial y} \Big|_{(0,0)} &= 4(x+y) - \sin y - y \cos y \Big|_{(0,0)} \\ &= 0 \end{aligned}$$

$$\frac{\partial F}{\partial x^2} \Big|_{(0,0)} = 4 \quad \frac{\partial F}{\partial y^2} \Big|_{(0,0)} = 4 = \frac{\partial F}{\partial x \partial y} \Big|_{(0,0)}$$

$$\frac{\partial F}{\partial y^2} \Big|_{(0,0)} = 2$$

$$\frac{\partial f}{\partial x} \Big|_{(0,0)} = 4x + 4y \Big|_{(0,0)} = 0$$

$$\frac{\partial f}{\partial y} \Big|_{(0,0)} = 4x + 4y \Big|_{(0,0)} = 0$$

$$\frac{\partial f}{\partial x^2} \Big|_{(0,0)} = 4$$

$$\frac{\partial f}{\partial y^2} \Big|_{(0,0)} = 4 = \frac{\partial f}{\partial x \partial y} \Big|_{(0,0)}$$

$$\frac{\partial f}{\partial y^2} \Big|_{(0,0)} = 2$$

$$f(x, y) = 2x^2 + 4xy + y^2$$
$$= 2(x+y)^2 - y^2$$

在 $y = -x$ 直线上 各在比 $(0,0)$ 小的点

$\Rightarrow (0,0)$ 不是极小值

Remark 1

(0,0)处取值无关 (无关)

The zero-order terms $F(0,0) = 7$ and $f(0,0) = 0$ have no effect on the answer. They simply raise or lower the graphs of F and f .

The linear terms of F give a necessary condition:

To have any chance of a minimum, the first derivatives must vanish at $x=y=0$: **一阶导必须为0 才可能有极值**

$$\frac{\partial F}{\partial x} = 4(x+y) - 3x^2 = 0 \quad \text{and}$$

$$\frac{\partial F}{\partial y} = 4(x+y) - y\cos y - \sin y = 0$$

Remark 2

The linear terms of f give a necessary condition:

To have any chance of a minimum, the first derivatives must vanish at

$x = y = 0$:

$$\frac{\partial f}{\partial x} = 4(x+y) = 0 \quad \text{and}$$

$$\frac{\partial f}{\partial y} = 4x - 2y = 0. \quad \text{All zero.}$$



Thus $(x, y) = (0, 0)$ is a stationary point for both functions. The surface $z = F(x, y)$ is tangent to the horizontal plane $z = 7$, and the surface $z = f(x, y)$ is tangent to the plane $z = 0$. The question is whether the graphs go above those planes or not, as we move away from the tangency point $x = y = 0$.

用 $z=7$ 切面. 人在桥头站在水平面上方

Remark 3

The second derivatives at $(0,0)$ are decisive:

$$\frac{\partial^2 F}{\partial x^2} = 4 - 6x = 4$$

$$\frac{\partial^2 F}{\partial x \partial y} = \frac{\partial^2 F}{\partial y \partial x} = 4$$

$$\frac{\partial^2 F}{\partial y^2} = 4 + y \sin y - 2 \cos y = 2$$

$$\frac{\partial^2 f}{\partial x^2} = 4$$

$$\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x} = 4$$

$$\frac{\partial^2 f}{\partial y^2} = 2.$$

These second derivatives 4, 4, 2 contain the answer. Since they are the same for F and f , they must contain the same answer. The two functions behave in exactly the same way near the origin. F has a minimum if and only if f has a minimum. We are going to show that those functions don't!

Remark 4

For $F(x, y)$:

- (a) The higher-degree terms in F have no effect on the question of a local minimum, but they can prevent it from being a global minimum.
- (b) In our example the term $-x^3$ must sooner or later pull F toward $-\infty$.

For $f(x, y)$ with no higher terms, all the action is at $(0, 0)$:

- (a) Every quadratic form $f(x, y) = ax^2 + 2bxy + cy^2$ has a stationary point at the origin, where $\frac{\partial f}{\partial x} = \frac{\partial f}{\partial y} = 0$.
- (b) A local minimum would also be a global minimum.
- (c) The surface $z = f(x, y)$ will then be shaped like a bowl, resting on the origin.



Positive Definiteness

$$\left. \frac{\partial F}{\partial x} \right|_{(\alpha, \beta)} = \left. \frac{\partial F}{\partial y} \right|_{(\alpha, \beta)} = 0$$

- (a) If the stationary point of F is at $x = \alpha, y = \beta$, the only change would be to use the second derivatives at (α, β) :

$$f(x, y) = \frac{x^2}{2} \frac{\partial^2 F}{\partial x^2}(\alpha, \beta) + xy \frac{\partial^2 F}{\partial x \partial y}(\alpha, \beta) + \frac{y^2}{2} \frac{\partial^2 F}{\partial y^2}(\alpha, \beta)$$

$f(0,0)$ 与 $F(\alpha, \beta)$ 相同
 $\frac{\partial F}{\partial x}(\alpha, \beta) = \frac{\partial F}{\partial y}(\alpha, \beta) = 0$

- (b) This $f(x, y)$ behaves near $(0, 0)$ in the same way that $F(x, y)$ behaves near (α, β) .

- (c) The third derivatives are drawn into the problem when the second derivatives fail to give a definite decision. That happens when the quadratic part is singular.

$$f(x, y) > 0 \quad (x, y) \neq (0, 0)$$

- (d) For a true minimum, f is allowed to vanish only at $x = 0, y = 0$. When $f(x, y)$ is strictly positive at all other points (the bowl goes up), it is called positive definite.

二次型 $f(x, y) = ax^2 + 2bxy + cy^2$ (对称矩阵 - 分母二)
 $= [x, y] \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \mathbf{x}^T A \mathbf{x}$

Figure 6.1

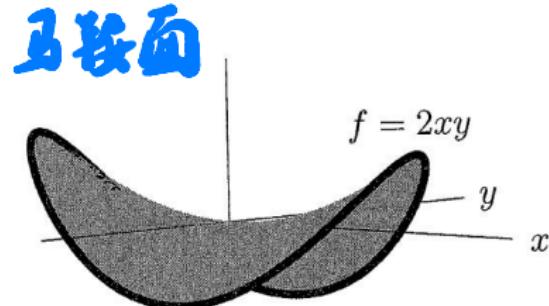
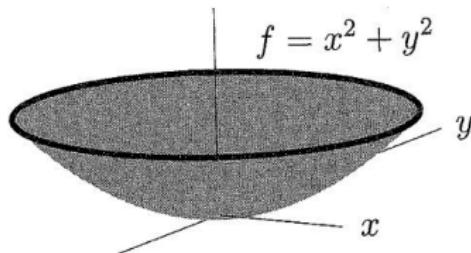


Figure 6.1: A bowl and a saddle: Definite $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and indefinite $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

The problem comes down to this: For a function of two variables x and y , what is the correct replacement for the condition $\frac{\partial^2 F}{\partial x^2} > 0$? With only one variable, the sign of the second derivative decides between a minimum or a maximum. Now we have three derivatives: $F_{xx}, F_{xy} = F_{yx}, F_{yy}$. These three numbers must determine whether or not F has a minimum.

Examples

Example 1 $f(x,y) = x^2 - 10xy + y^2$. 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 $b = -10$ overwhelms a and c .

Example 2 In our original f the coefficient $2b = 4$ was positive. Does this ensure a minimum? Again the answer is no; the sign of b is of no importance! Even though its second derivatives are positive, $2x^2 + 4xy + y^2$ is not positive definite. Neither F nor f has a minimum at $(0,0)$ because $f(1,-1) = 2 - 4 + 1 = -1$. It is the size of b , compared to a and c , that must be controlled.

General requirements

1. It is the size of b , compared to a and c , that must be controlled. We now want a necessary and sufficient condition for positive definiteness.
2. The simplest technique is to complete the square:

$$f(x, y) = ax^2 + 2bxy + cy^2 = a \left(x + \frac{b}{a}y \right)^2 + \left(c - \frac{b^2}{a} \right) y^2 \quad (1)$$

The first term on the right is never negative, when the square is multiplied by $a > 0$. But this square term must then be positive. That term has coefficient $\frac{ac - b^2}{a}$.

3. The last requirement for positive definiteness is that this coefficient must be positive:

If $ax^2 + 2bxy + cy^2$ stays positive, then necessarily $ac > b^2$.

Test for a minimum

The conditions $a > 0$ and $ac > b^2$ are just right. They guarantee $c > 0$.

The right side of (1) is positive, and we have found a minimum.

Theorem

$ax^2 + 2bxy + cy^2$ is positive definite if and only if $a > 0$ and $ac > b^2$. Any $F(x, y)$ has a minimum at a point (α, β) where $\frac{\partial F}{\partial x}(\alpha, \beta) = \frac{\partial F}{\partial y}(\alpha, \beta) = 0$ with

$$\frac{\partial^2 F}{\partial x^2}(\alpha, \beta) > 0 \text{ and } \left[\frac{\partial^2 F}{\partial x^2}(\alpha, \beta) \right] \left[\frac{\partial^2 F}{\partial y^2}(\alpha, \beta) \right] > \left[\frac{\partial^2 F}{\partial x \partial y}(\alpha, \beta) \right]^2$$

Here

$$a = \frac{1}{2} \frac{\partial^2 F}{\partial x^2}(\alpha, \beta), \quad b = \frac{1}{2} \frac{\partial^2 F}{\partial x \partial y}(\alpha, \beta), \quad c = \frac{1}{2} \frac{\partial^2 F}{\partial y^2}(\alpha, \beta).$$



Test for a Maximum

Since f has a maximum whenever $-f$ has a minimum, we just reverse the signs of a, b , and c .

- This actually leaves $ac > b^2$ unchanged: The quadratic form is negative definite if and only if $a < 0$ and $ac > b^2$.
- The same change applies for a maximum of $F(x,y)$.

Singular case $ac = b^2$

- (a) The second term in equation

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

disappears to leave the first square—which is either positive semidefinite, when $a > 0$, or negative semidefinite, when $a < 0$.

- (b) The prefix semi allows the possibility that f can equal zero, as it will at the point $x = b, y = -a$.
- (c) The surface $z = f(x,y)$ degenerates from a bowl into a valley. For $f = (x+y)^2$, the valley runs along the line $x+y=0$.

Saddle point

- (a) In one dimension, $F(x)$ has a minimum or a maximum, or $F'' = 0$.
- (b) In two dimensions, a very important possibility still remains: The combination $ac - b^2$ may be negative. This occurred in both examples, when b dominated a and c . It also occurs if a and c have opposite signs. Then two directions give opposite results—in one direction f increases, in the other it decreases.
- (c) It is useful to consider two special cases:

$$f_1(x, y) = 2xy \quad \text{and} \quad f_2(x, y) = x^2 - y^2$$

Saddle points at $(0, 0)$.

- (d) The saddles $2xy$ and $x^2 - y^2$ are practically the same; if we turn one through 45° we get the other. They are also hard to draw.

Quadratic Forms

For any symmetric matrix A , the product $x^T A x$ is a pure **quadratic form** $f(x_1, x_2, \dots, x_n)$:

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

$\sum_{i=1}^n \sum_{j=1}^n a_{ij} x_i x_j$ $x^T A x$ **对称矩阵**
 $= f(x)$

There are no higher-order terms or lower-order terms-only second-order. The function is zero at $x = (0, \dots, 0)$, and its first derivatives are zero.

$f(x) > 0$ for all $x \neq 0$ \Rightarrow 正定

$$[x_1 \ x_2 \ \dots \ x_n] \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & \ddots & & \\ \vdots & & & \\ a_{n1} & & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \left[\sum_{i=1}^n a_{1i} x_i \ \dots \ \sum_{i=1}^n a_{ni} x_i \right] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

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

每一项都是二次的
 $a_{ij} = a_{ji}$ 如果 $i \neq j$ 可以合并

Examples

Example 3 $f = 2x^2 + 4xy + y^2$ and $A = \begin{bmatrix} 2 & 2 \\ 2 & 1 \end{bmatrix} \rightarrow$ saddle point.

Example 4 $f = 2xy$ and $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \rightarrow$ saddle point.

Example 5 A is 3 by 3 for $2x_1^2 - 2x_1x_2 + 2x_2^2 - 2x_2x_3 + 2x_3^2$:

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

f has a minimum at $(0, 0, 0)$.

$$\mathbf{x}^T A \mathbf{x}$$

(A 对称 矩阵)

A 对称 (对称矩阵)

$$\mathbf{x}^T A \mathbf{x} = \mathbf{x}^T Q^T \Lambda Q \mathbf{x} = \mathbf{y}^T \Lambda \mathbf{y} = \lambda_1 y_1^2 + \dots + \lambda_n y_n^2$$

$$\text{例 5: } = 2(x_1^2 - x_1x_2 + (x_1x_2)^2) - \frac{1}{3}x_2^2 + 2x_3^2 - 2x_2x_3 + 2x_3^2$$

$$\text{期末船方} = 2(x_1 - \frac{1}{2}x_2)^2 + \frac{2}{3}(x_2^2 - \frac{4}{3}x_2x_3 + (\frac{2}{3}x_3)^2) - \frac{2}{3}x_3^2 + 2x_3^2$$

$$\text{由\gamma 船方} = \underline{2(x_1 - \frac{1}{2}x_2)^2} + \underline{\frac{2}{3}(x_2^2 - \frac{4}{3}x_2x_3)} + \underline{\frac{4}{3}x_3^2} > 0$$

for all $(x_1, x_2, x_3) \neq (0, 0, 0)$

$f(x_1, x_2, x_3)$ positive definite

$$A = \begin{bmatrix} 2 & 1 & 0 \\ 1 & 2 & 1 \\ 0 & 1 & 2 \end{bmatrix} \rightarrow \begin{bmatrix} 2 & 1 & 0 \\ 0 & \frac{3}{2} & 1 \\ 0 & 0 & 4 \end{bmatrix} \quad \text{LDU 分解}$$

用逆乘法

$$A = \begin{bmatrix} 1 & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \\ 0 & -\frac{2}{3} & 1 \end{bmatrix} \begin{bmatrix} 2 & \frac{1}{2} & 0 \\ 0 & \frac{3}{2} & 1 \\ 0 & 0 & 4 \end{bmatrix} \begin{bmatrix} 1 & -\frac{1}{2} & 0 \\ 0 & \frac{1}{2} & -\frac{1}{3} \\ 0 & 0 & 1 \end{bmatrix} \quad \text{LOU}$$

正定是否立元中都 >0 ?

$f(x_1, x_2, x_3)$

" $x^T A x > 0$ for all $x \neq 0$
 $\Rightarrow A$ positive definite

$$A_{nn} = LDL^T$$

$$x^T A x = x^T L D L^T x$$

$$y = L^T x$$

$$\Rightarrow = y^T D y$$

$$= d_1 y_1^2 + \dots + d_n y_n^2 > 0$$

for all $(y_1, y_2, \dots, y_n) \neq (0, 0, \dots, 0)$

$$d_1, d_2, \dots, d_n > 0$$

与 (x_1, x_2, \dots, x_n) 等价

$\neq (0, 0, \dots, 0)$
因为 L^T 正定

$$f(x_1, x_2, x_3) = 2x_1x_2 - 6x_1x_3 + 2x_2x_3$$

无平方项

技巧: $x_1 = y_1 + y_2$

$x_2 = y_1 - y_2$ 产生平方项

$x_3 = y_3$

$$= 2(y_1^2 - y_2^2) - 6y_1y_3 - 6y_2y_3 + 2y_1y_3 - 2y_2y_3$$

...

Test for positive definiteness

For function $F(x_1, x_2, \dots, x_n)$: F has a minimum when the pure quadratic $x^T Ax$ is positive definite. These second-order terms control F near the stationary point: 相当于“凹槽”

$$F(x) = F(0) + x^T (\text{grad } F) + \frac{1}{2} x^T Ax + \text{higher order terms}$$

-凹槽 - 等效于 - 对称性 - 特征值 - 阶数 - 只为 0

The next section contains the tests to decide whether $x^T Ax$ is positive (the bowl goes up from $x = 0$). Equivalently, the tests decide whether the matrix A is positive definite—which is the main goal of the chapter.

Homework Assignment 27

6.1: 5, 7, 11, 13, 16, 17, 19.