四月廿五日課程補充參考資料

在開始正題之前我們複習泰勒定理。首先是單變數函數的泰勒定理：

Theorem 1.1. *Let $f:(a,b)\to\mathbb{R}$ be $(n+1)$-times differentiable. Then for all c, $x\in(a,b)$ satisfying $x\neq c$, there exists ξ between c and x such that*

$$f(x)=\sum_{k=0}^{n}\frac{f^{(k)}(c)}{k!}(x-c)^k+\frac{f^{(n+1)}(c)}{(n+1)!}(x-c)^{n+1}\,.$$

接下來是多變數函數的泰勒定理：

Theorem 1.2. *Let $\Omega\subseteq\mathbb{R}^n$ be an open set, and $f:\Omega\to\mathbb{R}$ be $(n+1)$-times differentiable. Then for all $c=(c_1,\cdots,c_n)$, $x=(x_1,\cdots,x_n)\in\Omega$ satisfying $x\neq c$ and $\overline{cx}\subseteq\Omega$, there exists $\xi=(\xi_1,\cdots,\xi_n)\in\overline{cx}$ such that*

$$f(x)=\sum_{k=0}^{n}\sum_{|\alpha|=k}\frac{(D^\alpha f)(c)}{\alpha!}(x-c)^\alpha+\sum_{|\alpha|=n+1}\frac{(D^\alpha f)(\xi)}{\alpha!}(x-c)^\alpha\,,$$

where for a given multi-index $\alpha=(\alpha_1,\cdots,\alpha_n)$,

$$\alpha!=\alpha_1!\cdots\alpha_n!\,,\qquad |\alpha|=\alpha_1+\cdots+\alpha_n\,,$$
$$D^\alpha f=\frac{\partial^{|\alpha|}f}{\partial x_1^{\alpha_1}\cdots\partial x_n^{\alpha_n}}\,,\qquad (x-c)^\alpha=(x_1-c_1)^{\alpha_1}\cdots(x_n-c_n)^{\alpha_n}\,.$$

接下來我們用泰勒定理證明一個我們上學期學過的 second derivative test 的另一個（條件給得比較嚴格的）版本。

Theorem 1.3. *Let $f:(a,b)\to\mathbb{R}$ be a twice differentiable function, $c\in(a,b)$, and $f'(c)=0$. Suppose that f'' is continuous at c.*

1. *If $f''(c)>0$, then f attains its relative minimum at c.*
2. *If $f''(c)<0$, then f attains its relative maximum at c.*

Proof. 1. Since $f'(c)>0$ and f'' is continuous at c, there exists $h>0$ such that

$$f''(x)>0\qquad\forall\,x\in(c-h,c+h)\,.$$

Let $x\in(c-h,c+h)$ and $x\neq c$. By Taylor's Theorem, there exists ξ between c and x such that

$$f(x)=f(c)+f'(c)(x-c)+\frac{f''(\xi)}{2}(x-c)^2\,.$$

Since $\xi\in(c-h,c+h)$ if $x\in(c-h,c+h)$, we have $f''(\xi)>0$; thus the fact that $f'(c)=0$ shows that

$$f(x)=f(c)+\frac{f''(\xi)}{2}(x-c)^2>f(c)\,.$$

Therefore, f attains its relative minimum at c.

2. By changing $>$ to $<$, we obtain the proof for the second case. □

上述定理與其證明可以幫助我們理解接下來要敘述的多變數函數的 second derivative test。

Theorem 1.4. *Let $\mathcal{R}\subseteq\mathbb{R}^n$ be an open region, $f:R\to\mathbb{R}$ be a twice differentiable, $(a,b)\in R$ and $f_x(a,b)=f_y(a,b)=0$. Suppose that f_{xx}, f_{xy}, f_{yx} and f_{yy} are continuous at (a,b).*

1. *If the matrix* $\begin{bmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{bmatrix}$ *is positive definite; that is,*

$$\begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} > 0 \qquad \forall\,(u,v) \neq (0,0)\,,$$

   then f attains its relative minimum at (a,b).

2. *If the matrix* $\begin{bmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{bmatrix}$ *is negative definite; that is,*

$$\begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} < 0 \qquad \forall\,(u,v) \neq (0,0)\,,$$

   then f attains its relative maximum at (a,b).

Proof. We mimic the proof of the previous theorem.

1. First, the continuity of f_{xy} and f_{yx} at (a,b) implies that $f_{xy}(a,b) = f_{yx}(a,b)$. Note that the matrix $\begin{bmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{bmatrix}$ is positive definite if and only if

$$f_{xx}(a,b) > 0 \quad \text{and} \quad \begin{vmatrix} f_{xx}(a,b) & f_{xy}(a,b) \\ f_{yx}(a,b) & f_{yy}(a,b) \end{vmatrix} = f_{xx}(a,b)f_{yy}(a,b) - f_{xy}(a,b)^2 > 0\,.$$

   By the continuity of f_{xx}, f_{xy}, f_{yx} and f_{yy} at (a,b), we find that there exists $\delta > 0$ such that

$$f_{xx}(x,y) > 0 \quad \text{and} \quad f_{xx}(x,y)f_{yy}(x,y) - f_{xy}(x,y)^2 > 0 \quad \forall\,(x,y) \in D((a,b),\delta);$$

   thus we obtain $\delta > 0$ that

$$\begin{bmatrix} f_{xx}(x,y) & f_{xy}(x,y) \\ f_{yx}(x,y) & f_{yy}(x,y) \end{bmatrix} \text{ is positive definite for all } (x,y) \in D((a,b),\delta).$$

   Let $(x,y) \in D((a,b),\delta)$ and $(x,y) \neq (a,b)$. Then the segment joining (a,b) and (x,y) is a subset of $D((a,b),\delta)$ (since $D((a,b),\delta)$ is convex); thus by Taylor's Theorem there exists (ξ,η) on the segment joining (a,b) and (x,y) such that

$$\begin{aligned} f(x,y) &= \sum_{k=0}^{1} \sum_{|\alpha|=k} \frac{(D^\alpha f)(a,b)}{\alpha!}((x,y)-(a,b))^\alpha + \sum_{|\alpha|=2} \frac{(D^\alpha f)(\xi,\eta)}{\alpha!}((x,y)-(a,b))^\alpha \\ &= f(a,b) + \sum_{|(\alpha_1,\alpha_2)|=1} \frac{(D^{(\alpha_1,\alpha_2)} f)(a,b)}{\alpha_1!\alpha_2!}(x-a,y-b)^{(\alpha_1,\alpha_2)} \\ &\quad + \sum_{|(\alpha_1,\alpha_2)|=2} \frac{(D^{(\alpha_1,\alpha_2)} f)(\xi,\eta)}{\alpha!}(x-a,y-b)^{(\alpha_1,\alpha_2)} \\ &= f(a,b) + \frac{(D^{(1,0)} f)(a,b)}{1!0!}((x-a,y-b))^{(1,0)} + \frac{(D^{(0,1)} f)(a,b)}{0!1!}((x,y)-(a,b))^{(0,1)} \\ &\quad + \frac{(D^{(2,0)} f)(\xi,\eta)}{2!0!}((x-a,y-b))^{(2,0)} + \frac{(D^{(1,1)} f)(\xi,\eta)}{1!1!}((x,y)-(a,b))^{(1,1)} \\ &\quad + \frac{(D^{(0,2)} f)(\xi,\eta)}{0!2!}((x,y)-(a,b))^{(0,2)}\,. \end{aligned}$$

   Since

$$D^{(1,0)}f = f_x, \quad D^{(0,1)}f = f_y, \quad D^{(2,0)}f = f_{xx}, \quad D^{(1,1)}f = f_{xy}, \quad D^{(0,2)}f = f_{yy}\,,$$

the fact that $f_x(a,b) = f_y(a,b) = 0$ shows that

$$\begin{aligned} f(x,y) &= f(a,b) + f_x(a,b)(x-a) + f_y(a,b)(y-b) \\ &\qquad + \frac{1}{2} f_{xx}(\xi,\eta)(x-a)^2 + f_{xy}(\xi,\eta)(x-a)(y-b) + \frac{1}{2} f_{yy}(\xi,\eta)(y-b)^2 \\ &= f(a,b) + \frac{1}{2} \begin{bmatrix} x-a & y-b \end{bmatrix} \begin{bmatrix} f_{xx}(\xi,\eta) & f_{xy}(\xi,\eta) \\ f_{yx}(\xi,\eta) & f_{yy}(\xi,\eta) \end{bmatrix} \begin{bmatrix} x-a \\ y-b \end{bmatrix}. \end{aligned}$$

Since $(x,y) \neq (a,b)$, $(x-a, y-b) \neq (0,0)$; thus the fact that $(\xi,\eta) \in D((a,b),\delta)$ shows that

$$\begin{bmatrix} x-a & y-b \end{bmatrix} \begin{bmatrix} f_{xx}(\xi,\eta) & f_{xy}(\xi,\eta) \\ f_{yx}(\xi,\eta) & f_{yy}(\xi,\eta) \end{bmatrix} \begin{bmatrix} x-a \\ y-b \end{bmatrix} > 0\,;$$

thus $f(x,y) > f(a,b)$ for all $(x,y) \in D((a,b),\delta)$ satisfying $(x,y) \neq (a,b)$. Therefore, f attains its relative minimum at (a,b).

2. Again by changing $>$ to $<$ (and positive to negative), we obtain the proof for the second case. □

Remark 1.5. 我們如何判斷一個 2×2 對稱矩陣 $\begin{bmatrix} A & B \\ B & C \end{bmatrix}$ 是正定的呢？由定義可知 $M \equiv \begin{bmatrix} A & B \\ B & C \end{bmatrix}$ 是正定的若且唯若

$$Au^2 + 2Buv + Cv^2 = \begin{bmatrix} u & v \end{bmatrix} \begin{bmatrix} A & B \\ B & C \end{bmatrix} \begin{bmatrix} u \\ v \end{bmatrix} > 0 \qquad \forall\, (u,v) \neq (0,0)\,.$$

注意到 $(u,v) \neq (0,0)$ 表示 u, v 至少其一非零。

1. 若 $u \neq 0$，我們定義 $x = v/u$ 並得到 M 為正定若且唯若對所有 $x \in \mathbb{R}$ 我們有 $A+2Bx+Cx^2 > 0$。這個觀察證明了 $C > 0$ 且判別式 $B^2 - AC < 0$ 兩條件同時滿足等價於 M 為正定矩陣。
2. 若 $v \neq 0$，我們定義 $x = u/v$ 並得到 M 為正定若且唯若對所有 $x \in \mathbb{R}$ 我們有 $Ax^2+2Bx+C > 0$。這個觀察證明了 $A > 0$ 且判別式 $B^2 - AC < 0$ 兩條件同時滿足等價於 M 為正定矩陣。

上述兩條件幫助我們如何判斷一個 2×2 矩陣是否正定。判斷負定的方法，唯一的區別是係數 A 或 C 得小於零（但是判別式 $B^2 - AC$ s 一樣要小於零）。

In general, (by almost the same proof of the theorem above) for a twice differentiable function f of n-variables defined on an open set Ω with

$$\frac{\partial f}{\partial x_k}(c_1, \cdots, c_n) = 0 \qquad \forall\, k \in \{1, 2, \cdots, n\}\,,$$

where $c \equiv (c_1, \cdots, c_n) \in \Omega$ at which $f_{x_j x_k} \equiv \dfrac{\partial^2 f}{\partial x_k \partial x_j}$ is continuous for all $1 \leqslant j,k \leqslant n$, we have

1. f attains its relative minimum at c if the matrix $\begin{bmatrix} f_{x_1x_1}(c) & \cdots & f_{x_1x_n}(c) \\ \vdots & \ddots & \vdots \\ f_{x_nx_1}(c) & \cdots & f_{x_nx_n}(c) \end{bmatrix}$ is positive definite.
2. f attains its relative maximum at c if the matrix $\begin{bmatrix} f_{x_1x_1}(c) & \cdots & f_{x_1x_n}(c) \\ \vdots & \ddots & \vdots \\ f_{x_nx_1}(c) & \cdots & f_{x_nx_n}(c) \end{bmatrix}$ is negative definite.

一般而言用來判斷一個對稱矩陣正負定的方法可參考 Sylvester's criterion（請同學自行查 wiki）。