微積分 MA1002-A 上課筆記（精簡版）

2019.05.14.

Ching-hsiao Arthur Cheng 鄭經斅

Definition 13.58: Relative Extrema

Let f be a function defined on a region R containing (x_0, y_0).

1. The function f has a relative minimum at (x_0, y_0) if $f(x, y) \geqslant f(x_0, y_0)$ for all (x, y) in an open disk containing (x_0, y_0).

2. The function f has a relative maximum at (x_0, y_0) if $f(x, y) \leqslant f(x_0, y_0)$ for all (x, y) in an open disk containing (x_0, y_0).

Definition 13.59: Critical Points

Let f be defined on an open region R containing (x_0, y_0). The point (x_0, y_0) is a critical point of f if one of the following is true.

1. $f_x(x_0, y_0) = 0$ and $f_y(x_0, y_0) = 0$;

2. $f_x(x_0, y_0)$ or $f_y(x_0, y_0)$ does not exist.

Theorem 13.60

Let R be an open region in the plane, and $f : R \to \mathbb{R}$ be continuous. If f has a relative extremum at (x_0, y_0) on an open region R, then (x_0, y_0) is a critical point of f.

Definition 13.62

Let f be a function of two variables. A point (x_0, y_0) is a saddle point of f if (x_0, y_0) is a critical point of f but f does not attain its extrema at (x_0, y_0).

Theorem 13.63

Suppose that a function f of two variables has continuous second partial derivatives on an open region containing a point (x_0, y_0) for which $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$. Let

$$D = f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2 = \begin{vmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{yx}(x_0, y_0) & f_{yy}(x_0, y_0) \end{vmatrix}.$$

1. If $D > 0$ and $f_{xx}(x_0, y_0) > 0$, then f has a relative minimum at (x_0, y_0).

2. If $D > 0$ and $f_{xx}(x_0, y_0) < 0$, then f has a relative maximum at (x_0, y_0).

3. If $D < 0$, then (x_0, y_0) is a saddle point of f.

4. The test is inconclusive if $D = 0$.

13.8.3 Absolute Extrema

Theorem 13.65: Extreme Value Theorem

Let f be a continuous function of two variables x and y defined on a closed bounded region R in the plane.

1. There is at least one point in R at which f takes on a minimum value.
2. There is at least one point in R at which f takes on a maximum value.

A minimum is also called an absolute minimum and a maximum is also called an absolute maximum.

Example 13.66. Find the absolute extrema of the function $f(x,y) = \sin(xy)$ on the closed region given by $0 \leqslant x \leqslant \pi$ and $0 \leqslant y \leqslant 1$.

From the partial derivatives

$$f_x(x,y) = y\cos(xy) \quad \text{and} \quad f_y(x,y) = x\cos(xy)\,,$$

we find that each point on the hyperbola $xy = \dfrac{\pi}{2}$ is a critical point of f. The value of f at each of these points is $\sin\dfrac{\pi}{2} = 1$ which is the maximum of the sine function. Therefore, the maximum of f is 1.

The minimum of f occurs at the boundary of the region.

1. $x = 0$ and $0 \leqslant y \leqslant 1$: then $f(x,y) = 0$.
2. $x = \pi$ and $0 \leqslant y \leqslant 1$: then $f(x,y) = \sin(\pi y)$. The critical points of the function $g(y) = \sin(\pi y)$ occurs at $y = \dfrac{1}{2}$ since $g'\big(\dfrac{1}{2}\big) = \pi\cos\big(\dfrac{\pi}{2}\big) = 0$. Since $g\big(\dfrac{1}{2}\big) = 1$ and $g(0) = g(1) = 0$, we find that the minimum of g is 0.
3. $y = 0$ and $0 \leqslant x \leqslant \pi$: then $f(x,y) = 0$.
4. $y = 1$ and $0 \leqslant x \leqslant \pi$: then $f(x,y) = \sin x$ whose minimum on $[0,\pi]$ is 0.

Therefore, the minimum of f is 0.

The concepts of relative extrema and critical points can be extended to functions of three or more variables. On the other hand, the second derivative test for functions of three or more variables are more tricky, and we will not talk about this until the course of Advance Calculus.

13.9 Applications of Extrema

Theorem 13.67

The least squares regression line for n points $\{(x_1,y_1),(x_2,y_2),\cdots,(x_n,y_n)\}$ is given by $y=ax+b$, where

$$a=\frac{n\sum\limits_{i=1}^{n}x_iy_i-\Big(\sum\limits_{i=1}^{n}x_i\Big)\Big(\sum\limits_{i=1}^{n}y_i\Big)}{n\sum\limits_{i=1}^{n}x_i^2-\Big(\sum\limits_{i=1}^{n}x_i\Big)^2}\quad\text{and}\quad b=\frac{1}{n}\Big(\sum_{i=1}^{n}y_i-a\sum_{i=1}^{n}x_i\Big). \tag{13.9.1}$$

Proof. For $a,b\in\mathbb{R}$, define $S(a,b)=\sum\limits_{i=1}^{n}(ax_i+b-y_i)^2$. Then

$$\frac{\partial S}{\partial a}(a,b)=2\sum_{i=1}^{n}(ax_i+b-y_i)x_i\,,$$
$$\frac{\partial S}{\partial b}(a,b)=2\sum_{i=1}^{n}(ax_i+b-y_i)\,.$$

The critical points (a,b) of S satisfies

$$a\sum_{i=1}^{n}x_i^2+b\sum_{i=1}^{n}x_i=\sum_{i=1}^{n}x_iy_i\,, \tag{13.9.2a}$$
$$a\sum_{i=1}^{n}x_i+b\sum_{i=1}^{n}1=\sum_{i=1}^{n}y_i \tag{13.9.2b}$$

which implies that (a,b) are given by (13.9.1). Clearly such (a,b) minimizes S. □

Remark 13.68. An easy way to memorize the equations (a,b) satisfies is given in this remark. We assume (even though in general it is a false assumption) that the line $y=ax+b$ passes through (x_1,y_1), (x_2,y_2), $\cdots$, (x_n,y_n). Then $y_i=ax_i+b$ for all $1\leqslant i\leqslant n$; thus in matrix form, we have

$$\begin{bmatrix}x_1 & 1\\ x_2 & 1\\ \vdots & \vdots\\ x_n & 1\end{bmatrix}\begin{bmatrix}a\\ b\end{bmatrix}=\begin{bmatrix}y_1 & 1\\ y_2 & 1\\ \vdots & \vdots\\ y_n & 1\end{bmatrix}.$$

Therefore,

$$\begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} x_1 & 1 \\ x_2 & 1 \\ \vdots & \vdots \\ x_n & 1 \end{bmatrix} \begin{bmatrix} a \\ b \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \\ 1 & 1 & \cdots & 1 \end{bmatrix} \begin{bmatrix} y_1 & 1 \\ y_2 & 1 \\ \vdots & \vdots \\ y_n & 1 \end{bmatrix}$$

which implies (13.9.2).

13.10 Lagrange Multipliers

The concept of this section is to find the extrema of a function of several variables subject to certain constraints:

Find extrema of the function $w = f(x_1, x_2, \cdots, x_n)$ when $(x_1, x_2, \cdots, x_n)$ satisfies $g_1(x_1, \cdots, x_n) = g_2(x_1, \cdots, x_n) = \cdots = g_m(x_1, \cdots, x_n) = 0$.

Theorem 13.69: Lagrange Multiplier Theorem - Simplest Version

Let f and g be continuously differentiable functions of two variables. Suppose that on the level curve $g(x, y) = c$ the function f attains its extrema at (x_0, y_0). If $(\nabla g)(x_0, y_0) \neq \mathbf{0}$, then there is a real value λ such that

$$(\nabla f)(x_0, y_0) = \lambda (\nabla g)(x_0, y_0) \,.$$

Remark 13.70. The scalar λ in the theorem above is called a Lagrange multiplier.