SE102:Multivariable Calculus

Hyosang Kang[1]

[1]Division of Mathematics
School of Interdisciplinary Studies
DGIST

Lecture 04
Maxima and Minima

Definition

An **open ball** $B_\varepsilon(\mathbf{a})$ of the radius $\varepsilon > 0$ at a point $\mathbf{a} \in \mathbf{R}^n$ is the set defined by

$$B_\varepsilon(\mathbf{a}) = \{\mathbf{x} \in \mathbf{R}^n \mid \|\mathbf{x} - \mathbf{a}\| < \varepsilon\}$$

Definition

A function $f(x_1, \ldots, x_n) = (y_1, \ldots, y_m)$ is said to be **continuously differentiable**, or of **class** $\mathcal{C}^1$, at $\mathbf{a} \in \mathbf{R}^n$ if there is $\varepsilon > 0$ such that all the partial derivatives $\dfrac{\partial y_j}{\partial x_i}$ are continuous on an open ball $B_\varepsilon(\mathbf{a})$.

Theorem

If a function $z = f(x, y)$ is continuously differentiable at $\mathbf{a} = (x_0, y_0)$, then it is differentiable at $\mathbf{a}$.

Proof.

Let $B_\varepsilon(\mathbf{a})$ be an open ball on which the partial derivatives f_x, f_y are continuous. Let us choose a point $\mathbf{b} = (x, y) \in B_\varepsilon(\mathbf{a})$ and define a sequence

$$\mathbf{p}_0 = \mathbf{a}, \quad \mathbf{p}_1 = (x, y_0), \quad \mathbf{p}_2 = \mathbf{b}$$

Without loss of generality, we will assume that $x_0 \leq x$ and $y_0 \leq y$. Define a function $\phi_1(t) = f(t, y_0)$ on $[x_0, x]$. Note that $\phi_1'(t) = f_x(t, y_0)$. By the mean value theorem on one-variable function, there is $t_1 \in [x_0, x]$ such that

$$\phi_1'(t_0)(x - x_0) = f(x, y_0) - f(x_0, y_0)$$

Proof.

Likewise, let us define $\phi_2(t) = f(x,t)$ on $[y_0, y]$, then there exists $t_2 \in [y_0, y]$ satisfying

$$\phi_2'(t_2)(y - y_0) = f(x,y) - f(x,y_0)$$

Then,

$$\begin{aligned}
&f(\mathbf{b}) - f(\mathbf{a}) - f_x(\mathbf{a})(x - x_0) - f_y(\mathbf{a})(y - y_0) \\
&= (\phi_1'(t_1) - f_x(\mathbf{a}))(x - x_0) + (\phi_2'(t_2) - f_y(\mathbf{a}))(y - y_0) \\
&= (f_x(t_1, y_0) - f_x(\mathbf{a}))(x - x_0) + (f_y(x, t_2) - f_y(\mathbf{a}))(y - y_0)
\end{aligned}$$

Note that as $\mathbf{b} \to \mathbf{a}$, we have $t_1 \to x_0$ and $t_2 \to y_0$, which means the above goes to 0. Since $(x - x_0)/\|\mathbf{b} - \mathbf{a}\|$ and $(y - y_0)/\|\mathbf{b} - \mathbf{a}\|$ bounded, we proved the differentiability of f at $\mathbf{a}$.

Definition

A function $z = f(x, y)$ is said to be of **class** $\mathcal{C}^2$ at $\mathbf{a} \in \mathbf{R}^2$ if there exists $\varepsilon > 0$ such that all second-order partial derivatives are continuous on an open ball $B_\varepsilon(\mathbf{a})$.

Remark

Given a function $\mathbf{f} : \mathbf{R}^n \to \mathbf{R}^m$, the derivative $\mathbf{Df}$ is a function $\mathbf{Df} : \mathbf{R}^n \to \mathbf{R}^{nm}$ whose coordinate functions are all the first derivatives of $\mathbf{f}$. The function $\mathbf{f}$ being a class $\mathcal{C}^1$ means that $\mathcal{D}\{$ is continuous. Similarly, being a class $\mathcal{C}^2$ means that $\mathbf{Df}$ is $\mathcal{C}^1$, which equivalent to saying that $\mathbf{D}^2\mathbf{f}$ is continuous. We say the function $\mathbf{f}$ is of class $\mathcal{C}^n$ if $\mathbf{D}^n\mathbf{f}$ is continuous. Moreover, the function $\mathbf{f}$ is of class $\mathcal{C}^\infty$ if $\mathbf{D}^n\mathbf{f}$ is continuous for all $n > 0$.

Theorem (Clairaut)

If $z = f(x, y)$ is of class $\mathcal{C}^2$ at (x_0, y_0), then

$$f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0).$$

Example

Check if the following function satisfies Clairaut's theorem at $(0, 0)$.

$$f(x, y) = \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2} & (x, y) \neq (0, 0) \\ 0 & (x, y) = (0, 0) \end{cases}$$

Proof of Clairaut's Theorem.

Let us define $\mu : B_\varepsilon(\mathbf{a}) \to \mathbf{R}$ as

$$\mu(x,y) = f(x,y) - f(x_0,y) - f(x,y_0) + f(x_0,y_0)$$

Let us define $\phi(t) = f(t,y) - f(t,y_0)$ on $[x_0,x]$. Since f is of class $\mathcal{C}^1$, ϕ is differentiable on $[x_0,x]$. Note that

$$\mu(x,y) = \phi(x) - \phi(x_0).$$

By the mean value theorem, there exists $t_1 \in [x_0,x]$ satisfying

$$\mu(x,y) = \phi'(t_1)(x-x_0) = (f_x(t_1,y) - f_x(t_1,y_0))(x-x_0).$$

Since f_x is of class $\mathcal{C}^1$, we can apply the mean value theorem on $\psi(h) = f_x(t_1,h)$, which means that there exists $t_2 \in [y_0,y]$ satisfying

$$f_x(t_1,y) - f_x(t_1,y_0) = \psi'(t_2)(y-y_0) = f_{xy}(t_1,t_2)(y-y_0)$$

Proof.

Thus we showed that there exists $(t_1, t_2) \in [x_0, x] \times [y_0, y]$ satisfying

$$\mu(x, y) = f_{xy}(t_1, t_2)(x - x_0)(y - y_0).$$

If we switch the role of x and y when we defined ϕ, we would obtain the similar result: there exists a point $(s_1, s_2) \in [x_0, x] \times [y_0, y]$ satisfying

$$\mu(x, y) = f_{yx}(s_1, s_2)(x - x_0)(y - y_0).$$

Let us replace $x = x_0 + h$ and $y = y_0 + h$. Since we assume that f is of class $\mathcal{C}^2$, the following function converges as $(x, y) \to (x_0, y_0)$.

$$\frac{\mu(x_0 + h, y_0 + h)}{h^2} = f_{xy}(t_1, t_2) = f_{yx}(s_1, s_2)$$

Since $(t_1, t_2), (s_1, s_2) \to (x_0, y_0)$ as $(x, y) \to (x_0, y_0)$, we are done.

Definition

Suppose that $f(x, y)$ is of class $\mathcal{C}^2$ at (x_0, y_0). Then the following polynomial is called the **Taylor polynomial of second degree** 2 of f at (x_0, y_0).

$$
\begin{aligned}
Q(x, y) = {} & f(x_0, y_0) + f_x(x_0, y_0)(x - x_0) \\
& + f_y(x_0, y_0)(y - y_0) + \frac{1}{2} f_{xx}(x_0, y_0)(x - x_0)^2 \\
& + f_{xy}(x_0, y_0)(x - x_0)(y - y_0) \\
& + \frac{1}{2} f_{yy}(x_0, y_0)(y - y_0)^2.
\end{aligned}
$$

Remark

If all n-th order partial derivatives of a function $f(x, y)$ are continuous, then

- f is differentiable, and
- the formulae of n-order partial derivatives does not depend on the order of partial derivatives.

Remark

Let

$$\Delta \mathbf{x} = (\Delta x, \Delta y) = (x - x_0, y - y_0)$$

and denote the *gradient operator* ∇ in vector notation:

$$\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right)$$

Let us define the *multiplication* of differential operators as follows:

$$\left(\frac{\partial}{\partial x}\frac{\partial}{\partial x}\right) f = \frac{\partial^2 f}{\partial x^2}, \quad \left(\frac{\partial}{\partial x}\frac{\partial}{\partial y}\right) f = \frac{\partial^2 f}{\partial x \partial y}$$

By the assumption in the definition of Taylor polynomial, we have $f_{xy}(x_0, y_0) = f_{yx}(x_0, y_0)$. Thus we can write

$$Q(x, y) = \sum_{n=0}^{2} \frac{1}{n!}(\Delta \mathbf{x} \cdot \nabla)^n f(x_0, y_0)$$

We can generalize this to the k-th order Taylor polynomial.

$$P_k f(\mathbf{x}) = \sum_{n=0}^{k} \frac{1}{n!}(\Delta \mathbf{x} \cdot \nabla)^n f(\mathbf{x}_0)$$

This is a generalization of the Taylor polynomial for single variable function:

Theorem (Taylor)

Let $z = f(x, y)$ *be a function of class* $\mathcal{C}^3$ *on a rectangular region*

$$D = \{(x, y) \mid |x - x_0|, |y - y_0| \le \epsilon\}.$$

Then for each $(x, y) \in D$*, there exists a constant* $0 \le c \le 1$ *satisfying*

$$f(x, y) = Q(x, y) + R_2(x, y)$$

where

$$R_2(x, y) = \frac{1}{3!}(\Delta\mathbf{x} \cdot \nabla)^3 f(\mathbf{x}_0 + c\Delta\mathbf{x})$$

Remark

This theorem is a generalization of Taylor theorem for one-variable function:

> *Let f be of class $\mathcal{C}^{k+1}$ on an interval $I = (x_0 - \epsilon, x_0 + \epsilon)$. Then for $x, c \in I$, there exists a constant ξ between x and c such that*
>
> $$f(x) = P_k(x) + \frac{f^{(k+1)}(\xi)}{(k+1)!}(x-c)^{k+1}$$

Note that the choice of c *depends* on the choice of x, x_0. Since all the third-order partial derivatives of f are bounded, the error $|R_n(x,y)|$ decreases to zero, as $(x,y) \to (x_0, y_0)$,

Example

Find $Q_2(x, y)$ at $(0, 0)$ for

- $f(x, y) = xy - x^2 - 5y^2 + y - 1$
- $f(x, y) = \cos x \cos y$

and compare the graphs of Q_2 and f near $(1, 0)$

Definition

Let $z = f(x, y)$ be a function defined on an open ball at $\mathbf{a} = (x_0, y_0)$. The point $\mathbf{a}$ is said to be **local maximal** (**minimal**, respectively) if there exists a sufficiently small $\varepsilon > 0$ such that for all $\mathbf{x} \in B_\varepsilon(\mathbf{a})$,

$$f(\mathbf{a}) \geq f(\mathbf{x}) \quad (f(\mathbf{a}) \leq f(\mathbf{x}), \text{ respectively}).$$

A local maximal or minimal is called an **extremal**.

Definition

A point $(\mathbf{a})$ is called a **critical point** if it satisfies one of the following conditions:

1. $f_x(\mathbf{a}) = f_y(\mathbf{a}) = 0$;
2. f_x or f_y does not exist at $\mathbf{a}$;
3. f is discontinuous at $\mathbf{a}$.

A critical point which is *not* an extremal point is called a **saddle point**.

Example

Find the critical points of

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

and classify them. Also, find $Q_2(x, y)$ at each critical points and compare their graphs.

Remark

1. Suppose that $f(x, y)$ is differentiable at $\mathbf{a}$ and $f_x(\mathbf{a}) = f_y(\mathbf{a}) = 0$. Then the linear approximation of $f(x, y)$ at $\mathbf{a}$ is the plane $z = f(\mathbf{a})$ which is parallel to xy-plane.
2. Suppose that $\mathbf{a}$ is a saddle point of $f(x, y)$. Then there exists a curve

$$c : (-\epsilon, \epsilon) \to \mathbf{R}^2, \quad c(0) = (\mathbf{a})$$

such that composition $F(t) = (f \circ c)(t)$ has an inflection point at $t = 0$.

Definition

Suppose that $f(x, y)$ is of class $\mathcal{C}^2$ at $\mathbf{a} \in \mathbf{R}^2$. Then

$$\Delta_f = f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2$$

is called the **discriminant** of f.

Example

Graph the following function at $(0, 0)$ and compare their discriminants.

$$z = -x^2 - y^2, \quad z = x^2 + y^2, \quad z = x^2 - y^2$$

Theorem (Hesse)

Let (x_0, y_0) *be a critical point satisfying* $f_x(\mathbf{a}) = f_y(\mathbf{a}) = 0$.

- *If* $\Delta_f > 0$ *and* $f_{xx}(x_0, y_0) > 0$, *then* (x_0, y_0) *is a local minimal point.*
- *If* $\Delta_f > 0$ *and* $f_{xx}(x_0, y_0) < 0$, *then* (x_0, y_0) *is a local maximal point.*
- *If* $\Delta_f < 0$, *then* $f(x_0, y_0)$ *is a saddle point.*
- *If* $\Delta_f = 0$, *then we cannot determine local extremity by this method.*

Remark

Let $\Delta x = x - x_0, \Delta y = y - y_0$. The summands of the second degrees of $Q(x, y)$ can be written as

$$\frac{1}{2}\begin{bmatrix}\Delta x & \Delta y\end{bmatrix}\underbrace{\begin{bmatrix} f_{xx}(x_0, y_0) & f_{xy}(x_0, y_0) \\ f_{xy}(x_0, y_0) & f_{yy}(x_0, y_0)\end{bmatrix}}_{=A}\begin{bmatrix}\Delta x \\ \Delta y\end{bmatrix}$$

Note that $\Delta_f = \det A$. Using linear transformation of x, y, we can coorrespond the matrix A to one of three matrices below, wihout changing the classifications of extremals.

$$\begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}, \quad \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad \begin{bmatrix} \pm 1 & 0 \\ 0 & \mp 1 \end{bmatrix}$$

Example (Least square method)

Suppose that a set of data is given by

$$(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n)$$

We want to find a linear function $f(x) = mx + b$ which approximates these data. More precisely, we want to minimize the sum of all squares of errors $f(x_i) - y_i$ to be minimum. we say that $y = m_0x + b_0$ best approximates the data. In other words, we want to find m, b such that

$$d(m, b) = \sum_{i=1}^{n} (y_i - (mx_i + b))^2$$

is minimum.

Consider $d(m,b)$ as two-variable function on m, b. The critical point is

$$m_0 = \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}, \quad b_0 = \frac{\sum_{i=1}^{n} x_i^2 \sum_{i=1}^{n} y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} x_i y_i}{n\sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i\right)^2}$$

The Hessian of $d(m,b)$ at (m_0, b_0) is $\begin{bmatrix} 2\sum_{i=1}^{n} x_i^2 & 2\sum_{i=1}^{n} x_i \\ 2\sum_{i=1}^{n} x_i & 2n \end{bmatrix}$

Since the discriminant of f is

$$\Delta_f(m_0, b_0) = 4n \sum_{i=1}^{n} x_i^2 - 4 \left(\sum_{i=1}^{n} x_i \right)^2 > 0$$

The point (m_0, b_0) is a local minimal point. In fact, it is a global minimal point. (why?)

Theorem

Let (x_0, y_0, z_0) *be a critical point of* $f(x, y, z)$ *where* f_x, f_y, f_z *are all zero. Let* H *be the* 3×3 *matrix defined by*

$$\begin{bmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{bmatrix}$$

Let d_1, d_2, d_3 *be the determinants of the* 1×1, 2×2, 3×3 *sub-matrices on the left-top corner of* H*.*

- *If* $d_i > 0$ *for all* i*, then* (x_0, y_0, z_0) *is a local minimum point.*
- *if* $d_1, d_3 < 0$ *and* $d_2 > 0$*, then* (x_0, y_0, z_0) *is a local maximal point.*
- *In all other cases,* (x_0, y_0, z_0) *is a saddle point.*

Proposition

Let $L_c(f)$ *be the level curve at* $c = f(x_0, y_0)$. *on the* xy*-plane. Then the gradient vector* $\nabla f(x_0, y_0)$ *is perpendicular to the curve* $L_c(f)$ *at* (x_0, y_0)

Theorem (Lagrange multiplier)

Let $g(x, y)$, $f(x, y)$ be differentiable functions. Let $L_c(g)$ be a level curve at c. Let us retrict the domain of f onto $L_c(g)$. If (x_0, y_0) is an extremal point of f and $\nabla g(x_0, y_0) \neq \mathbf{0}$, there exists λ such that

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

Example

The Lagrange multiplier finds the maxima or minima of a **target** function $f(x, y)$ under the **constraint** $g(x, y) = c$. Find the point on the circle $x^2 + y^2 = 10$ where the function $f(x, y) = 3x + y$ attains maximal or minimal.

Corollary

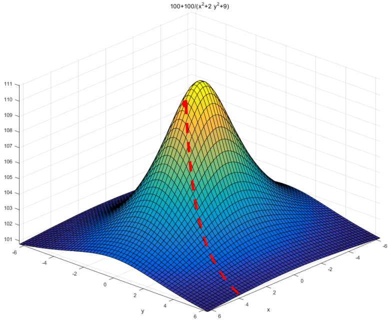
The gradient vector $\nabla f(x_0, y_0)$ has the direction where the value of function $f(x, y)$ increases the most from (x_0, y_0).

Example

Let

$$f(x,y) = 100 + \frac{100}{x^2 + 2y^2 + 9}$$

be a function whose graph represent the contour of a mountain. Suppose that a water flow from the point $(1, 0, 110)$ down the valley in the steepest direction. Find the trajectory of the water path.



Theorem

Let $g(x, y, z)$, $f(x, y, z)$ be differentiable functions. Suppose that (x_0, y_0, z_0) is a local extremal of $f(x, y, z)$ restricted the level set $L_c(g)$. If $\nabla g(x_0, y_0) \neq \mathbf{0}$, then there exists λ such that

$$\nabla f(x_0, y_0, z_0) = \lambda \nabla g(x_0, y_0, z_0)$$

Example

Find the minimal and maximal value of $f(x, y, z) = x^3 + y^3 + z^3$ on the sphere $x^2 + y^2 + z^2 = 1$ on the first octant.

Problem

Find all critical points and classify them

1. $f(x, y) = xy + \dfrac{2}{x} + \dfrac{2}{y}$
2. $e^y(x^2 + y^2 - z^2)$

Problem

Find all local extremes of $f(x, y)$ with the give contraints.

1. $f(x, y) = 2x + y^2 + x^2, 2x^2 + y^2 = 2$
2. $f(x, y, z) = xy + yz, x^2 + y^2 = 1, yz = 1$

Problem

Find the local extremes of $f(x,y) = x^2 + xy + y^2$ on the disk $D = \{(x,y) \mid x^2 + y^2 \leq 1\}$.

Problem

Find the point on the graph $xy^2z^3 = 2$ which is the closest to the origin.