THE CHAIN RULE

① $f(x,y)$. $c(t) = (x(t), y(t))$.

matrix vector

$\star$ $(f \circ c)'(t) = \underline{f_x}\, x'(t) + \underline{f_y}\, y(t) = [f_x \ f_y] \begin{bmatrix} x' \\ y' \end{bmatrix}$

SE102:Multivariable Calculus

② $f : \mathbb{R}^n \to \mathbb{R}^m$, $g : \mathbb{R}^m \to \mathbb{R}^l$.

$D(g \circ f) = Dg \cdot Df$.

Hyosang Kang[1]

[1]Division of Mathematics
School of Interdisciplinary Studies
DGIST

Week 04

Df or f' is a linear map that sends the velocity vector in domain space to the corresponding velocity " in target space.

☆ Theorem (Clairaut)

If the partial derivatives f_{xy}, f_{yx} are continuous at (x_0, y_0), then

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

(thus on a small region containing (x_0, y_0))

Definition

Suppose that $f(x, y)$ has continuous second partial derivatives at (x_0, y_0). Then the polynomial

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

is called the **Taylor polynomial of second degree** of f at (x_0, y_0).

PF of CLAIRAUT.

"NON-PROOF"

$$f_{xy}(x_0,y_0) = (f_x)_y(x_0,y_0) = \lim_{t\to 0} \frac{1}{t}\left(f_x(x_0,y_0+t) - f_x(x_0,y_0)\right)$$

$$= \lim_{t\to 0}\frac{1}{t}\left(\lim_{h\to 0}\frac{1}{h}\left(f(x_0+h,y_0+t) - f(x_0,y_0+t)\right) - \lim_{h\to 0}\frac{1}{h}\left(f(x_0+h,y_0) - f(x_0,y_0)\right)\right)$$

NOT OK.

$$= \lim_{\substack{t\to 0\\ h\to 0}} \frac{1}{t}\cdot\frac{1}{h}\left(f(x_0+h,y_0+t) - f(x_0,y_0+t) - f(x_0+h,y_0) + f(x_0,y_0)\right).$$

$$= \lim_{\substack{t\to 0\\ h\to 0}} \frac{1}{h}\cdot\left(\frac{1}{t}\left(f(x_0+h,y_0+t) - f(x_0+h,y_0)\right) - \frac{1}{t}\cdot\left(f(x_0,y_0+t) - f(x_0,y_0)\right)\right).$$

$$= \lim_{h\to 0}\frac{1}{h}\cdot\left(f_y(x_0+h,y_0) - f_y(x_0,y_0)\right)$$

$$= (f_y)_x(x_0,y_0) = f_{yx}(x_0,y_0)$$

PF. $F(t,h) = \frac{1}{t}\frac{1}{h}\cdot\left(f(x_0+h,y_0+t) - f(x_0+h,y_0) - f(x_0,y_0+t) + f(x_0,y_0)\right)$

NTS: $\lim_{\substack{t\to 0\\ h\to 0}} F(t,h) = (f_{xy})(x_0,y_0) = (f_{yx})(x_0,y_0)$.

$g_1(x) = f(x, y_0+t) - f(x,y_0)$.

$(g_1)' = f_x(x,y_0+t) - f_x(x,y_0)$

conti.

$F(t,h) = \frac{1}{t}\cdot\frac{1}{h}\cdot\left(g_1(x_0+h) - g_1(x_0)\right)$.

By mean-value theorem. $\Rightarrow$ there exist $x_1 \in (x_0, x_0+h)$ such that

depend on h

$g_1(x_0+h) - g_1(x_0) = (g_1)'(x_1)\,h$

$$F(t,h) = \frac{1}{t} \cdot g_1'(x_1) = \frac{1}{t} \cdot \left(f_x(x_1, y_0+t) - f_x(x_1, y_0) \right)$$

$$\lim_{\substack{t\to 0\\ h\to 0}} F(t,h) = \lim_{\substack{t\to 0\\ (h\to 0)}} \frac{1}{t} \cdot \left(f_x(x_1, y_0+t) - f_x(x_1, y_0) \right) = \lim_{h\to 0} (f_x)_y (x_1, y_0)$$

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

$x_1 \to x_0$ as $h \to 0$

$$g_2(y) = f(x_0+h, y) - f(x_0, y)$$

$$\to \cdots \to \lim_{\substack{t\to 0\\ h\to 0}} F(t,h) = \cdots = (f_{yx})(x_0, y_0) \qquad \blacksquare$$

$$(\Delta \vec{x}) \circ \nabla = (\Delta x, \Delta y) \circ \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y} \right) = (x-x_0)\frac{\partial}{\partial x} + (y-y_0)\frac{\partial}{\partial y}$$

inner prod

$$(\Delta \vec{x} \circ \nabla)^0 = 1.$$

$$(\Delta \vec{x} \circ \nabla)^1 = \Delta \vec{x} \circ \nabla = (x-x_0)\frac{\partial}{\partial x} + (y-y_0)\frac{\partial}{\partial y}.$$

$$(\Delta \vec{x} \circ \nabla)^2 = \left((x-x_0)\frac{\partial}{\partial x} + (y-y_0)\frac{\partial}{\partial y} \right) \cdot \left((x-x_0)\frac{\partial}{\partial x} + (y-y_0)\frac{\partial}{\partial y} \right)$$

$$= (x-x_0)^2 \frac{\partial^2}{\partial x^2} + (y-y_0)(x-x_0)\frac{\partial}{\partial y}\frac{\partial}{\partial y}$$

$$+ (y-y_0)(x-x_0)\frac{\partial}{\partial x}\frac{\partial}{\partial y}$$

$$+ (y-y_0)^2 \frac{\partial^2}{\partial y^2}$$

use diff operators $\frac{\partial}{\partial x}, \frac{\partial}{\partial y}$ as if they are value (expression)

$$= (x-x_0)^2 \frac{\partial^2}{\partial x^2} + 2(y-y_0)(x-x_0)\frac{\partial^2}{\partial x \partial y} + (y-y_0)^2 \frac{\partial^2}{\partial y^2}.$$

if f_{xy}, f_{yx} are conti

☆ $f_{xyx} = f_{xxy}$.

☆ **Remark** Clairaut. thm if $f_{xyx}, f_{xxy}, f_{yxx}$ are conti.

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

bold 'x'.

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

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

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

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

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: since $(\Delta x \cdot \nabla)^n f(x_0) = \Delta x^n \nabla^n f(x_0)$,

$$P_k(x) = \sum_{n=0}^{k} \frac{f^{(n)}(x_0)\Delta x^n}{n!}$$

Theorem (Taylor)

Q approximates f better than L. (L from the definition of differentiability)

Let $f(x, y)$ is a function whose third partial derivatives $f_{xxx}, f_{xyx}, \cdots, f_{yyy}$ are all continuous on a rectangular region

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

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

$$f(x, y) = Q(x, y) + \underbrace{R_2(x, y)}_{\text{error.}}$$

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 generalizes the Taylor theorem for single-variable function:

> *Let f be a $\mathcal{C}^{k+1}$-function 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*

$$f(x) = P_k(x) + \frac{f^{(k+1)}(\xi)}{(k+1)!}(x-c)^{k+1}$$

Here $x = x_0 + \Delta x$ and $\xi = x_0 + c\Delta x$ for some $0 \leq c \leq 1$. Note that the choice of c *depends* on the choice of x, x_0. We can approximate *any* function by $Q_2(x, y)$ if it has partial derivatives upto 3rd order. If (x, y) is sufficiently close to (x_0, y_0), the error $|R_n(x, y)|$ decreases to zero.

Example

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

- $f(x, y) = xy - x^2 - 5y^2 + y - 1$ ← deg 2 poly.
- $f(x, y) = \cos x \cos y$ ← ☆.

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

$f(1,0) = -2.$ $\quad f_x = y - 2x \quad f_y = x - 10y + 1.$

$f_x(1,0) = -2 \quad f_y(1,0) = 2.$

$f_{xx} = -2 \quad f_{xy} = 1. \quad f_{yy} = -10$

$$Q = -2 + (-2)(x-1) + 2(y-0) + \frac{1}{2}\left(-2(x-1)^2 + 2\cdot 1\cdot(x-1)\cdot(y-0) - 10(y-0)^2\right)$$

$= f(x,y).$

Definition

Let $f(x, y)$ be a function $f(x, y)$ defined on a region D.

- A point (x_0, y_0) is said to be **local maximal** (**minimal**, respectively) at $\mathbf{x}_0 = (x_0, y_0)$ if there exists a (sufficiently small) $\epsilon > 0$ such that for all $\mathbf{x} = (x, y)$ satisfying $\|\mathbf{x} - \mathbf{x}_0\| < \epsilon$, we have $f(x_0, y_0) \geq f(x, y)$ ($f(x_0, y_0) \leq f(x, y)$, respectively). A local maximal / minimal is often called an **extremal**.
- A point (x_0, y_0) is called a **critical point** if it satisfies one of the following.

1. $f_x(x_0, y_0) = f_y(x_0, y_0) = 0$; $\to L = f(x_0, y_0)$. tangent plane is parallel to xy-plane.
2. f_x or f_y does not exist at (x_0, y_0);
3. f is discontinuous at (x_0, y_0).

A critical point which is *not* an extremal 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. Let $f(x, y)$ is differentiable at (x_0, y_0). Suppose that (x_0, y_0) is a critical point $f(x, y)$ of the type 1 in the definition. Then the linear approximation (x_0, y_0) is the plane $z = f(x_0, y_0)$.
2. Suppose that (x_0, y_0) is a saddle point. Then there exists a curve $c : (-\epsilon, \epsilon) \to \mathbf{R}^2$, $c(0) = (x_0, y_0)$ such that composition $F(t) = (f \circ c)(t)$ is a inflection point.

Q. HOW CAN WE CLASSIFY CRITICAL POINT SYSTEMATICALLY. (OR ANALYTICALLY) WITHOUT LOOKING AT THE GRAPH?

Definition

Let $R \subset \mathbf{R}^2$ be a domain of $f(x, y)$ and $(x_0, y_0) \in R$. Suppose that all second partial derivatives of $f(x, y)$ are continuous on a region R. Then

$$\Delta_f = f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0)^2 = \det \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$$

$(= f_{xy})$

"Hessian" of f.

is called the **discriminant** of f.

Example

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

l. max l. min. saddle

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

$\det \begin{bmatrix} -2 & 0 \\ 0 & -2 \end{bmatrix} = 4.$ $\det \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 4$ $\det \begin{bmatrix} 2 & 0 \\ 0 & -2 \end{bmatrix} = -4.$

Theorem (Hesse)

Let (x_0, y_0) be a critical point of the type 1 of $f(x, y)$.

- *If $\Delta > 0$ and $f_{xx}(x_0, y_0) > 0$, then (x_0, y_0) is a local minimum point.* ex $z = x^2 + y^2$
- *If $\Delta > 0$ and $f_{xx}(x_0, y_0) < 0$, then (x_0, y_0) is a local maximum point.* ex $z = -x^2 - y^2$
- *If $\Delta < 0$, then $\cancel{f}(x_0, y_0)$ is a saddle point.*
- *If $\Delta = 0$, then we cannot determine local extremity by this method.* ex $f = x^4 + y^4$. $f_{xx}(0,0) = 0$

$\det \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix} = 0$ $f_{yy}(0,0) = 0$

Remark

Let $\Delta x = x - x_0, \Delta y = y - y_0$. The degree-2 summands 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.

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

Say (x_0, y_0) is a critical point of type 1.

(i.e. $f_x(x_0,y_0) = f_y(x_0,y_0) = 0$).

Then $Q = f(x_0, y_0) + \frac{1}{2} f_{xx}(x_0,y_0)(x-x_0)^2 + f_{xy}(x_0,y_0)(x-x_0)(y-y_0)$

$$+ \frac{1}{2} f_{yy}(x_0,y_0)(y-y_0)^2.$$

$$= \underbrace{f(x_0,y_0)}_{\text{determines the "height", NOT determines the classification of critical pt.}} + \underbrace{\frac{1}{2}\begin{bmatrix} \overset{X}{x-x_0} & \overset{Y}{y-y_0} \end{bmatrix} \begin{bmatrix} f_{xx} & f_{xy} \\ f_{xy} & f_{yy} \end{bmatrix} \begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix}}_{\text{DETERMINES THE EXTREMALITY OF } (x_0,y_0)}$$

→ Ignore when classify critical pt (x_0,y_0)

$$= \begin{bmatrix} X & Y \end{bmatrix} \begin{bmatrix} a & b \\ b & c \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix}$$

$$= \begin{bmatrix} aX+bY & bX+cY \end{bmatrix} \begin{bmatrix} X \\ Y \end{bmatrix} = (aX+bY)X + (bX+cY)Y$$

$$= aX^2 + 2bXY + cY^2$$

$$= a\left(X^2 + \frac{2b}{a}XY + \frac{b^2}{a^2}Y^2\right) + \left(c - \frac{b^2}{a}\right)Y.$$

$$= a\Big(\underbrace{X + \frac{b}{a}Y}_{X'}\Big)^2 + \frac{ac-b^2}{a}\underbrace{Y}_{Y'}$$

$$= a(X')^2 + \frac{ac-b^2}{a}(Y')^2 = f(X', Y').$$

$a > 0$, $\frac{ac-b^2}{a} > 0$ $\Rightarrow$ $(X', Y') = (0,0)$ is a min.

$a < 0$, $\dfrac{ac-b^2}{a} < 0 \Rightarrow (x', y') = (0, 0)$ is local max.

$a \cdot \dfrac{ac-b^2}{a} = ac - b^2 < 0 \Rightarrow (x', y') = (0, 0)$ is saddle.

$(x - x_0, y - y_0) \longrightarrow (x, y)$ translated (x_0, y_0) to $(0, 0)$

$x + 0\,y \to x'$, $y \to y'$: linear transformation squeezes the graph BUT NOT changes the classification of the critical point

y_i : estimate

$f(x_i) = y_i'$: actual value

linear

Example (Least square method)

Suppose that a set of data is given by

$$(x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n) \leftarrow y_i = f(x_i)$$

slope y-intercept.

We want to find a line $y = mx + b$ which approximates these data. If there are values m_0, b_0 such that the sum of squares of vertical distances between data and the line $y = m_0 x + b_0$ is the minimum among all possible lines, then we say that $y = m_0 x + b_0$ best approximates the data. In other words, we want to find m, b such that

actual value

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

estimate

is minimum.

2-var func on m, b.

$$d(m,b) = \sum y_i^2 - 2\cdot(mx_i + b)\cdot y_i + (mx_i + b)^2$$

$$= \sum y_i^2 - \left(2\sum y_i\right)b - \left(2\sum y_i x_i\right)m$$

$$+ \left(\sum x_i^2\right)m^2 + 2\left(\sum x_i\right)mb + n\cdot b^2$$

$$d_m = -2\left(\sum y_i x_i\right) + 2\left(\sum x_i^2\right)m + 2\left(\sum x_i\right)b = 0$$

$$d_b = -2\left(\sum y_i\right) + 2\left(\sum x_i\right)m + 2nb = 0.$$

$$\begin{cases} \cancel{2}\left(\sum x_i^2\right)m + \cancel{2}\left(\sum x_i\right)b = \cancel{2}\left(\sum y_i x_i\right) \\ \cancel{2}\left(\sum x_i\right)m + \cancel{2}nb = \cancel{2}\left(\sum y_i\right) \end{cases}$$

$$\underbrace{\begin{bmatrix} \sum x_i^2 & \sum x_i \\ \sum x_i & n \end{bmatrix}}_{A} \begin{bmatrix} m \\ b \end{bmatrix} = \underbrace{\begin{bmatrix} \sum y_i x_i \\ \sum y_i \end{bmatrix}}_{b}.$$

$$m = \frac{\det\begin{bmatrix} \sum y_i x_i & \sum x_i \\ \sum y_i & n \end{bmatrix}}{\det A} = \frac{n\sum y_i x_i - \left(\sum x_i\right)\left(\sum y_i\right)}{n\sum x_i^2 - \left(\sum x_i\right)^2}$$

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

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

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

The REGRESSION takes the RED (length of) lines as the error.

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

$$d_{mm} = 2\sum x_i^2 > 0.$$

Since the discriminant of f is

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

The point (m_0, b_0) is a local minimum point. (in fact a global minimum point, why?)

$$n\sum x_i^2 - \underbrace{\left(\sum x_i\right)^2} = n\cdot\sum x_i^2 - \left(\sum x_i^2 + \sum_{i\neq j} x_i x_j\right).$$

$$= \underbrace{(n-1)\sum x_i^2} - 2\sum_{i>j} x_i x_j = \sum_{i>j} x_i^2 - 2x_i x_j + x_j^2 = \sum (x_i - x_j)^2 > 0$$

has continuous 2nd order partial deriv.

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

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

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.*

  $-$ $+$ $-$
  d_1 d_2 d_3
- *In all other cases, (x_0, y_0, z_0) is a saddle point.*

critical point NOT loc. min/max.

$f = x^2 + y^2 + z^2.$

$H = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \Rightarrow \quad d_1 = 2 \quad d_2 = 4 \quad d_3 = 8$

$f = -x^2 - y^2 - z^2$

$H = \begin{bmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{bmatrix} \Rightarrow \quad d_1 = -2 \quad d_2 = 4 \quad d_3 = -8$

$f = x^2 - y^2 + z^2$

$H = \begin{bmatrix} 2 & & \\ & -2 & \\ & & 2 \end{bmatrix} \Rightarrow \quad d_1 = 2 \quad d_2 = 4 \quad d_3 = -8.$

"Level set $L_c(f) = \{(x,y) \mid f(x,y) = c\}$.
at c of f"

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)

$(f_x(x_0,y_0), f_y(x_0,y_0))$

2-dim'l vector

pf

$d(0) = (x_0, y_0)$

Let $d(t) = (x(t), y(t))$ be a curve in $L_c(f)$. $(f \circ d)(t) = c$ (constant)

$$(f \circ d)'(0) = f_x(d(0)) \cdot x'(0) + f_y(d(0)) \cdot y'(0) = 0.$$

Chain rule

$$\Rightarrow \nabla f(x_0,y_0) \circ d'(0) = 0.$$

Inner prod

"target function"

or continuous partial derivative

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

ON THE $L_c(g)$

"constraint"

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

$\Leftrightarrow$ "∇f, ∇g are parallel"

OBJECTIVE: WE WANT TO FIND A POINT (x_0, y_0) ON $L_c(g)$ SUCH THAT $f(x_0, y_0)$ IS LOCALLY MIN/MAX ON $L_c(g)$.

WE WANT TO FIND LOC EXTREMALS ON THE $L_c(g)$ (NOT ON ENTIRE $\mathbb{R}^2$).

① Check f, g are diff.

② find ∇g, ∇f (vector field)

③ find x, y (and λ) satisfying

$$\lambda \nabla g = \nabla f \quad \text{AND} \quad g(x,y) = c$$

④ Then such (x_0, y_0) is A CANDIDATE of loc. min/max of f restricted on $L_c(g)$.

$\nabla f = (3,1) = \lambda(2x, 2y) = \lambda \nabla g$ AND $x^2+y^2=10$

$x = \frac{3}{2\lambda}$ $y = \frac{1}{2\lambda}$.

$\frac{10}{4\lambda^2} = 10 \Rightarrow \lambda = \pm\frac{1}{2}$.

$\lambda = \frac{1}{2} \Rightarrow (x,y) = (3,1)$, $\lambda = -\frac{1}{2} \Rightarrow (x,y) = (-3,-1)$

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.

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

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).

WHY LAGRANGE MULTIPLIER WORKS?

If we change d gradually the level set $L_c(f)$ "touches" the level set $L_c(g)$ at some point (x_0, y_0)

at (x_1, y_1), $\underset{\text{"}d'}{f(x_1,y_1)} < d$.

THE LOCAL EXTREMA HAPPENS WHEN $L_c(g)$ AND $L_d(f)$ "TOUCHES", OR TANGENT EACH OTHERS.

since $\nabla g \parallel L_c(g)$

$\nabla f \parallel L_d(f)$

and $L_c(g) \parallel L_d(f)$

thus $\nabla g \parallel \nabla f$

"parallel"

$\Rightarrow \nabla g, \nabla f$ are constant multiple of each others,

then $\nabla f = \lambda \nabla g$

(You can also use $\nabla g = \lambda \cdot \nabla f$)

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.

If the critical point is of type 1
↑ classify the extremality.

Problem

Find all critical points and classify them

1. $f(x, y) = xy + \frac{2}{x} + \frac{2}{y}$ $\longrightarrow$ $(0, y_0)$, $(x_0, 0)$: critical.
2. $e^y(x^2 + y^2 - z^2)$

$f_x = y - \frac{2}{x^2} = 0$

$f_y = x - \frac{2}{y^2} = 0$

Problem

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

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

$x^2 + \frac{y^2}{2} = 1$

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

$\nabla f = (2+2x, 2y) = \lambda \nabla g = \lambda(4x, 2y)$.

$\cancel{2}^{1} + \cancel{2}^{1}x = \cancel{4}^{2}\lambda x \Rightarrow$ if $\lambda = 1$. $1 + x = 2x \Rightarrow x = 1$. if $y = 0$.

$\cancel{2}y = \cancel{2}\lambda y \Rightarrow \lambda = 1$ or $y = 0$.

$1 + x = 2\lambda x$

$\Rightarrow x = \frac{1}{2\lambda - 1}$.

$(1, y, 1)$ or $\left(\frac{1}{2\lambda-1}, 0, \lambda\right)$

$\Downarrow$

$2 + y^2 = 2 \Rightarrow y = 0$

$\cancel{2} \cdot \left(\frac{1}{2\lambda-1}\right)^2 + 0^2 = \cancel{2} \Rightarrow 2\lambda - 1 = \pm 1 \Rightarrow \lambda = \frac{1}{2}(1 \pm 1) = 0, 1$.

$(1,0)$, $(-1,0)$, $\cancel{(1,0)}$.

$f(1,0) = 3$, $f(-1,0) = -1$.

loc. max loc. min

Problem

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

constraints. $g(x,y,z) = xy^2z^3$

Problem

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

WANT TO MINIMIZE $x^2+y^2+z^2 = f(x,y,z)$

"target function"