Partial Derivatives

→ Derivative of multivariable functions are defined only if we specify point the point and direction in which derivative is to be calculated.

→ The default direction of x axis & y axis for a variable functions is called partial derivatives.

Let $z = 2x^2 - xy + y^2$, with $x = s+t$, $y = st$.

Use the Chain Rule to find $\dfrac{\partial z}{\partial s}$ and $\dfrac{\partial z}{\partial t}$.

$$\frac{\partial z}{\partial x} = 4x - y \quad \Big| \quad \frac{\partial x}{\partial s} = 1 \quad \Big| \quad \frac{\partial x}{\partial t} = 1$$

$$\frac{\partial z}{\partial y} = -x + 2y \quad \Big| \quad \frac{\partial y}{\partial s} = t \quad \Big| \quad \frac{\partial y}{\partial t} = s$$

$$\frac{\partial z}{\partial s} = (4x-y)\,1 + (-x+2y)\,t$$

$$\frac{\partial z}{\partial t} = (4x-y)\cdot 1 + (x+2y)\,s$$

Definition

The **partial derivatives** of a two-variable function $f(x, y)$ are the functions

$$f_x(x,y) = \lim_{h\to 0} \frac{f(x+h,y) - f(x,y)}{h}$$

and

$$f_y(x,y) = \lim_{h\to 0} \frac{f(x,y+h) - f(x,y)}{h}.$$

Chain Rule

$z = f(x,y)$ $\quad x = g(t,v)$, $y = h(t,v)$

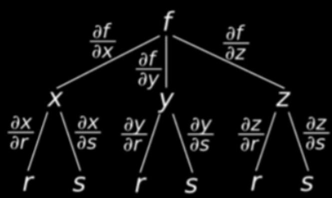
$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial t}$$

$$\frac{dz}{dt} = \frac{\partial z}{\partial x}\cdot\frac{dx}{dt} + \frac{\partial z}{\partial y}\cdot\frac{dy}{dt}$$

$z = f(x,y)$

$x = g(t)$

$y = h(t)$



To compute the derivative of f with respect to a variable in the bottom row, we follow every path to that variable, multiplying as we go, and add the results.

Suppose that f is a differentiable function of x and y, and

$$g(s,t) = f(s^2 - 3t, 4s - t).$$

	$f(x,y)$	$g(s,t)$	$f_x(x,y)$	$f_y(x,y)$
$(-2,3)$	5	1	2	3
$(1,1)$	4	5	6	7

(a) Use the table of values to compute $g_s(1,1)$.

(b) If $g_t(0,-1) = 3$ and $f_x(3,1) = 2$, use the table of values to compute $f_y(3,1)$.

a) $z = g(s,t) = f(x = s^2 - 3t,\ y = 4s - t)$

$g_s(1,1) = f(x=-2, y=3) = f(-2,3)$

$$g_s(s,t) = \frac{\partial z}{\partial s} = \frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial s} + \frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial s}$$

$$= 2\cdot 2 + 3\cdot 4$$

$$= \underline{\underline{16}}$$

b) $s = 0$, $t = -1$ $\Rightarrow$ $x = 3$, $y = 1$

$$\frac{\partial z}{\partial t} = \frac{\partial z}{\partial x}\cdot\frac{\partial x}{\partial t} + \frac{\partial z}{\partial y}\cdot\frac{\partial y}{\partial t}$$

$$3 = 2\cdot -3 + f_y\big|_{3,1}\ -1$$

$$3 + 6 = -f_y\big|(3,1) \Rightarrow f_y(3,1) = \underline{\underline{-9}}$$

Gradient $\nabla f(\bar{x})$

1) $\nabla f(x,y) = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix}$

2) $\nabla f(\bar{x}) = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix}$ mapping from $R^n \to R^n$

$\bar{x} = (x_1, x_2, \dots x_n)^T$ where $f(\bar{x})$ $R^n \to R$

$\bar{x} \in R^n$

3) $\nabla f(\bar{x}) \Big|_{\bar{x}=u} \in R^n$

where $f(\bar{x})$ is $R^n \to R$

Instead of Limit, directional derivative can be calculated directly from the gradient using dot product.

$$\nabla_{\hat{u}} f(\bar{x}) \Big|_{\bar{x}=\bar{v}} = \nabla f(\bar{x}) \Big|_{\bar{x}=\bar{v}} \cdot \hat{u}$$

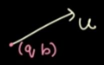
Directional derivative in unit vector direction $\hat{u}$ at point V

gradient @ point $\bar{v}$

Directional derivative $\nabla_{\vec{u}} f(\bar{x})$

Let

- f differentiable function at (a, b)
- $\mathbf{u} = \langle u_1, u_2 \rangle$ unit vector



Then the directional derivative of f in the direction of $\mathbf{u}$ at (a, b) is

$$\nabla_{\vec{u}} f = D_{\mathbf{u}} f(a, b) = \lim_{h \to 0} \frac{f(a + hu_1, b + hu_2) - f(a, b)}{h}$$

We can express this change as the difference quotient

$$\lim_{h \to 0} \frac{f(\vec{X} + h\vec{u}) - f(\vec{X})}{h}$$

1) Consider the curve $f(x, y) = x^2 - y^2$. Compute the rate of change in f at the point $(3, 2)$ in the direction of the vector $\vec{Y} = \begin{pmatrix} 2 \\ 1 \end{pmatrix}$.

$$\nabla_{\vec{u}} f(\bar{x}) \Big|_{x,y} = \underset{h \to 0}{Lt} \frac{f(x + hu_1, y + hu_2) - f(x, y)}{h}$$

$\hat{u} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$ $\quad \nabla_{\vec{u}} f(\bar{x}) \Big|_{a,b} = \underset{h \to 0}{Lt} \frac{(a + hu_1)^2 - (b + hu_2)^2 - a^2 + b^2}{h}$

$$\nabla_{\vec{u}} f(\bar{x}) \Big|_{a,b} = \frac{\cancel{a^2} + h^2 u_1^2 + 2ahu_1 - \cancel{b^2} - h^2 u_2^2 - 2bhu_2 - \cancel{a^2} + \cancel{b^2}}{h}$$

$- 2bu_2$

$\times 2 \cdot \frac{1}{\sqrt{5}} = \frac{8}{\sqrt{5}}$ //

Properties of the the gradient vector

Remark: If θ is the angle between ∇f and $\mathbf{u}$, then holds

$$D_{\mathbf{u}} f = \nabla f \cdot \mathbf{u} \quad \Rightarrow \quad D_{\mathbf{u}} f = |\nabla f| \cos(\theta).$$

The formula above implies:

- The function f increases the most rapidly when $\mathbf{u}$ is in the direction of ∇f, that is, $\theta = 0$. The maximum increase rate of f is $|\nabla f|$.
- The magnitude of the ∇f vector is equal to the largest rate of change.

→ In the perpendicular direction to gradient the function value does not change i.e. directional derivative in the perpendicular direction of derivative is 0.

$$\nabla_{\vec{u}} f(\bar{x} = v) = |\nabla f(\bar{x} = v)| \cos \theta$$

$\theta \to$ angle btw $\vec{u}$ & $\nabla f(\bar{x} = v)$

Make a change of variable that transforms the quadratic form

$$Q(\mathbf{x}) = x_1^2 - 8x_1x_2 - 5x_2^2$$

into a quadratic form with no cross-product term.

$Q(x) = x^T A x$

$A = \begin{bmatrix} 1 & -4 \\ -4 & -5 \end{bmatrix}$

$\begin{vmatrix} 1-\lambda & -4 \\ -4 & -5-\lambda \end{vmatrix} = 0$

$\Rightarrow \lambda^2 + 4\lambda - 21 = 0$

$(\lambda + 7)(\lambda - 3) = 0$

$\lambda = -7, \lambda = 3$

$X_{\lambda=-7}$ $\quad \begin{bmatrix} 8 & -4 \\ -4 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$

$8x - 4y = 0$

$2x - y = 0$

$2x = y$

$X_{\lambda=-7} = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$

$X_{\lambda=3}$

$\begin{bmatrix} -2 & -4 \\ -4 & -8 \end{bmatrix}$

$-2x - 4y = 0$

$-x = 2y$

$X_{\lambda=3} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$

$$A = \frac{\begin{bmatrix} 1 & -2 \\ 2 & 1 \end{bmatrix}}{\sqrt{5}} \begin{bmatrix} -1 & 0 \\ 0 & 3 \end{bmatrix} \frac{\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}}{\sqrt{5}} =$$

$$Q(x) = x^T P^T A P x$$

$$Q(y) = (Px)^T A (Px)$$

$$y = Px = \frac{\begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}}{\sqrt{5}} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \frac{\begin{bmatrix} x_1 + 2x_2 \\ x_2 - 2x_1 \end{bmatrix}}{\sqrt{5}}$$

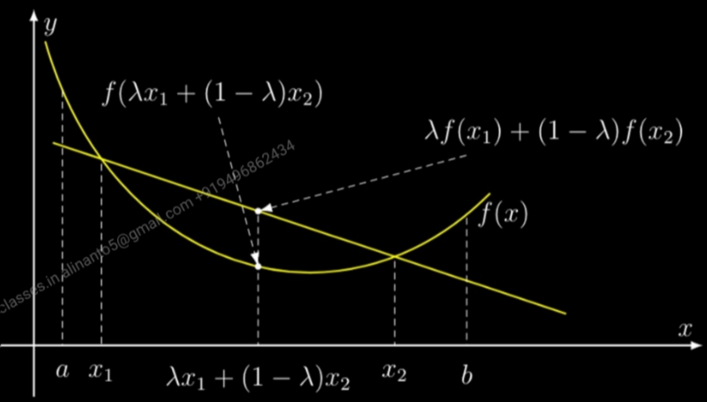


Figure 2: A convex function $f(x)$

$f(x)$ convex $\longleftrightarrow$ $-f(x)$ concave

$f(x) = k \;\; \forall x$ $\longrightarrow$ constant function is both convex & concave.

Condition for convex function

$f(\lambda x_1 + (1-\lambda)x_2) \leq \lambda f(x_1) + (1-\lambda) f(x_2)$ convex function

for $f(x) : \mathbb{R}^n \to \mathbb{R}$

$x_1, x_2 \in \mathbb{R}^n$

$\lambda \in [0,1]$

$f(\lambda x_1 + (1-\lambda)x_2) < \lambda f(x_1) + (1-\lambda) f(x_2)$ strict convex

$\forall \; x_1, x_2 \in \mathbb{R}^n$

$\forall \; \lambda \in [1,0]$

then $f(x) \mid \mathbb{R}^n \to \mathbb{R}$ is strictly convex

---

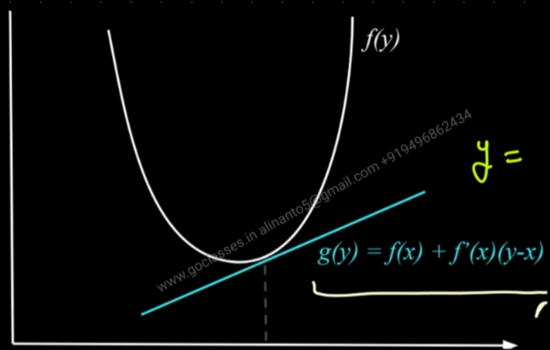
Convexity: First-order condition

A real-valued differentiable function is **convex** *iff*

$$f(y) \geq f(x) + \nabla f(x)^T (y - x),$$

for all $x, y \in \mathbb{R}^n$.

- The function is globally *above the tangent* at y.
- Show that any stationary point is a global minimum.



$y =$

Tangent eq. always below fn

Convex : Concave up

Concave: Concave down

Single Variable fn: convex $\longleftrightarrow$ $\frac{\partial^2 f}{\partial^2 x} \geq 0$

strictly convex $\longleftrightarrow$ $\frac{\partial^2 f}{\partial^2 x} > 0$

Convexity: Second-order condition

A real-valued twice-differentiable function is **convex** *iff*

$$\underbrace{\nabla^2 f(x)}_{\text{Hessian}} \succeq 0$$

for all $x \in \mathbb{R}^n$.

- The function is *flat or curved upwards* in every direction.

$\succeq$ Symbol for PSD

equalent to writing $x^T \nabla f x \geq 0$, $\forall x \in \mathbb{R}^n$

Maxima, Minima & Saddle point

For Single Valued function $f(x): \mathbb{R} \to \mathbb{R}$

① Stationary $\Rightarrow f'(x) = 0$
- → ① Maxima $f''(x) < 0$
- → ② Minima $f''(x) < 0$
- → ③ Sta. Inflexion point $f''(x) = 0$

② Inflexion point $f''(x) = 0$; $f'(x) \neq 0$ changes sign

③ Critical point (x_0) $\begin{cases} f'(x_0) = 0 \text{ or} \\ f'(x_0) \text{ doesn't exist } \pm\infty \end{cases}$

all stationary points are critical points

For convex function $f(x)$
- Hessian is PSD → necessary & sufficient
- Hessian is PD → sufficient not necessary

For concave function $f(x)$
- Hessian is NSD → necessary & sufficient
- Hessian is ND > sufficient not necessary

multivariate $f(\bar{x}): \mathbb{R}^n \to \mathbb{R}$ $\quad \bar{x} = [x_1, x_2, x_3 \dots x_n]^T$

$$\nabla f(\bar{x}) = \left[\frac{\partial f(\bar{x})}{\partial x_1} \quad \frac{\partial f(\bar{x})}{\partial x_2} \quad \dots \quad \frac{\partial f(\bar{x})}{\partial x_n}\right]$$

gradient

$$\nabla f(\bar{x}) : \mathbb{R}^n \to \mathbb{R}^n$$

Jacobian = gradient of Vector valued function $f(\bar{x}) \to \mathbb{R}^n$ to $\mathbb{R}^n$

Hessian = Jacobian of gradient

$$\text{Hessian } \nabla^2 f(\bar{x}) = \begin{bmatrix} \frac{\partial^2 f(\bar{x})}{\partial x_1^2} & \frac{\partial^2 f(\bar{x})}{\partial x_1 x_2} & \dots & \frac{\partial^2 f(\bar{x})}{\partial x_1 x_n} \\ \frac{\partial^2 f(\bar{x})}{\partial x_2 \partial x_1} & & & \\ \vdots & & & \\ \frac{\partial^2 f(\bar{x})}{\partial x_n x_1} & \dots & & \frac{\partial^2 f(\bar{x})}{\partial x_n x_n} \end{bmatrix}$$

$$\to \mathbb{R}^n \to \mathbb{R}^{n \times n}$$

Sufficient condition not necessary.

Maxima ← $\begin{cases} \text{gradient} = 0 \text{ vector } \& \\ \text{Hessian ND} \Leftrightarrow \text{all eigen values} < 0 \end{cases}$

Minima ← $\begin{cases} \text{gradient} = 0 \text{ vector } \& \\ \text{Hessian PD} \Leftrightarrow \text{all eigen value} > 0 \end{cases}$

Saddle point ← { both + & −ve eigen values

0 eigen value → test fails

Problem

$$\begin{aligned} &\text{minimize} \quad f(\mathbf{x}) \\ &\text{subject to:} \quad \mathbf{x} \in \mathbb{R}^n \end{aligned}$$

First Order Necessary Conditions

If $\mathbf{x}^*$ is a local minimizer of $f(\mathbf{x})$ and $f(\mathbf{x})$ is continuously differentiable in an open neighbourhood of $\mathbf{x}^*$, then

$$\nabla f(\mathbf{x}^*) = \mathbf{0}$$

That is, $f(\mathbf{x})$ is stationary at $\mathbf{x}^*$

Second Order Necessary Conditions

If $\mathbf{x}^*$ is a local minimizer of $f(\mathbf{x})$ and $\nabla^2 f(\mathbf{x})$ is continuously differentiable in an open neighbourhood of $\mathbf{x}^*$, then

gradient $\nabla f(\mathbf{x}^*) = \mathbf{0}$

Hessian $\nabla^2 f(\mathbf{x}^*)$ is positive semi definite $\quad H(x) \geq 0$

Necessary but not sufficient

Second Order Sufficient Conditions

Suppose that $\nabla^2 f(\mathbf{x})$ is continuously differentiable in an open neighbourhood of $\mathbf{x}^*$. If the following two conditions are satisfied, then $\mathbf{x}^*$ is a local minimum of $f(\mathbf{x})$.

$$\nabla f(\mathbf{x}^*) = \mathbf{0}$$

$\nabla^2 f(\mathbf{x}^*)$ is positive definite

Not neccessary but sufficient } For some thing to be local min this is not absolutely required. But if it is satisfied then it is absolutely a local min

for 2x2 matrix

$|\nabla^2 f(\bar{x})| > 0$ & $\nabla f(\bar{x}) = 0$ → Maxima

$|\nabla^2 f(\bar{x})| > 0$ & $\nabla f(\bar{x}) = 0$ → Minima

$|\nabla^2 f(\bar{x})| < 0$ & $\nabla f(x) = 0$ → saddle

given determinant & trace of a Hessian matrix + $\nabla f(\bar{x}) = 0$

Let $|\nabla^2 f(\bar{x})| = D$ $\quad$ Trace$(\nabla^2 f(\bar{x})) = T$ $\quad f(\bar{x}) \to \mathbb{R}^n \to \mathbb{R}$

Case ① $D < 0$
- → Atleast one −ve eigen value ⇒ Not Positive Definite
- → if n = even → Atleast one positive & negative eigen value → Not PD + Not ND
- → if $T > 0$ → Atleast one positive & negative eigen value → Not PD + Not ND

Case ② $D > 0$ → if n = odd & $T < 0$ → Both +ve & negative eigen value ⇒ Not PD + Not ND

saddle point