$f: \mathbb{R} \to \mathbb{R}$

f is said to be differentiable at c whenever the following limit exists:

$$\lim_{x \to c} \frac{f(x) - f(c)}{x - c} \triangleq f'(c) = \left.\frac{d f(x)}{dx}\right|_{x=c}.$$

$f: \mathbb{R}^n \to \mathbb{R}$ for $x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$, $c = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_n \end{pmatrix} \in \mathbb{R}^n$.

Suppose that $x_i = c_i$ for $i \neq k$. $x_k \neq c_k$

$$\lim_{x_k \to c_k} \frac{f(x) - f(c)}{x_k - c_k} \triangleq \frac{\partial f}{\partial x_k}(c)$$

If the limit exists, it is called the partial derivative of f with respect to the kth coordinate.

Here we simply treat f as a function of x_k, hold the others fixed.

Second-order partial derivative can also be defined similarly. $\frac{\partial^2 f}{\partial x_r \partial x_k} \overset{?}{=} \frac{\partial^2 f}{\partial x_k \partial x_r}$

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

$$\frac{\partial f(x,y)}{\partial x} = \begin{cases} \frac{y(x^4-y^4+4x^2y^2)}{(x^2+y^2)^2}, & \text{if } (x,y) \neq (0,0). \\ 0, & \text{if } (x,y) = (0,0) \end{cases}$$

$$\left(\lim_{x \to 0} \frac{f(x,0)-f(0,0)}{x-0} = 0 \right)$$

$$\frac{\partial}{\partial x}\left(\frac{A(x)}{B(x)}\right) = \frac{A'(x)B(x) - A(x)B'(x)}{B^2(x)}.$$

$$\left.\frac{\partial f(x,y)}{\partial x}\right|_{x=0} = -y, \ \forall y \Rightarrow \left.\frac{\partial^2 f(x,y)}{\partial y \partial x}\right|_{x=0} = -1 \ \forall y.$$

On the other hand,

$$\frac{\partial f(x,y)}{\partial y} = \begin{cases} \frac{x(x^4-y^4-4x^2y^2)}{(x^2+y^2)^2}, & \text{if } (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$$

$$\left.\frac{\partial f(x,y)}{\partial y}\right|_{y=0} = x \quad \forall x \Rightarrow \frac{\partial^2 f(x,y)}{\partial x \partial y} = +1 \quad \forall x.$$

$$\therefore \left.\frac{\partial^2 f(x,y)}{\partial x \partial y}\right|_{(x,y)=(0,0)} \neq \left.\frac{\partial^2 f(x,y)}{\partial y \partial x}\right|_{(x,y)=(0,0)}.$$

A function $f: \mathbb{R}^n \to \mathbb{R}^m$ is continuous at $x \in \operatorname{dom} f$ if $\forall \epsilon > 0, \exists \delta > 0$ such that for $y \in \operatorname{dom} f$, $\|y-x\|_2 \leq \delta$

$$\Rightarrow \|f(y) - f(x)\|_2 \leq \varepsilon$$

$\left(\begin{array}{c} |y-x| < \delta \\ \Rightarrow |f(x) - f(y)| < \epsilon \\ \text{if } f: \mathbb{R} \to \mathbb{R} \end{array} \right)$

Theorem If both $\frac{\partial f}{\partial x_1}$ & $\frac{\partial f}{\partial x_2}$ exist in a ball $B(c, \epsilon)$, $\epsilon > 0$ and if both $\frac{\partial f}{\partial x_i \partial x_j}$, $\frac{\partial f}{\partial x_j \partial x_i}$ are continuous at c,

then
$$\left.\frac{\partial^2 f}{\partial x_i \partial x_j}\right|_{x=c} = \left.\frac{\partial^2 f}{\partial x_j \partial x_i}\right|_{x=c}.$$

$f: \mathbb{R}^n \to \mathbb{R}^m$. $x \in \mathbb{R}^n$

$$f(x) = \begin{pmatrix} f_1(x) \\ f_2(x) \\ \vdots \\ f_m(x) \end{pmatrix}, \text{ where } f_i(x): \mathbb{R}^n \to \mathbb{R}.$$

The function f is differentiable at $x \in \operatorname{int} \operatorname{dom} f$ if there exists a matrix $Df(x) \in \mathbb{R}^{m \times n}$ that satisfies

$$\lim_{\substack{z \in \operatorname{dom} f \\ z \neq x,\ z \to x}} \frac{\| f(z) - f(x) - Df(x)(z-x) \|_2}{\| z - x \|_2} = 0$$

$$\lim_{z \to x} \frac{f(z) - f(x)}{z - x} = f'(x)$$

$$\Rightarrow \lim_{z \to x} \frac{f(z) - f(x) - f'(x)(z-x)}{z - x} = 0$$

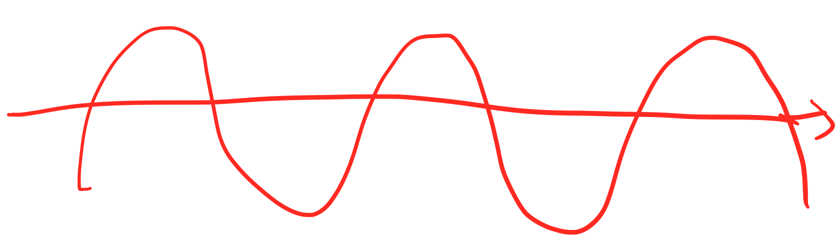
$Df(x)$: the derivative of f at x.

(Jacobian of $f(x)$).

(There can be at most one such matrix).

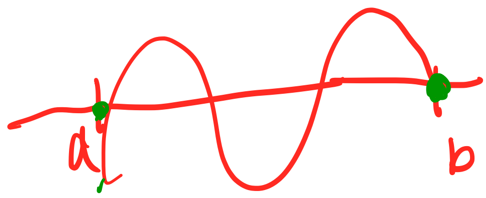
f is differentiable if $\operatorname{dom} f$ is open and is differentiable at every point in $\operatorname{dom} f$.

$f(x) = \sin x$



$\exists\, f'(x) \quad \forall (x)$

If $a \le x \le b$,



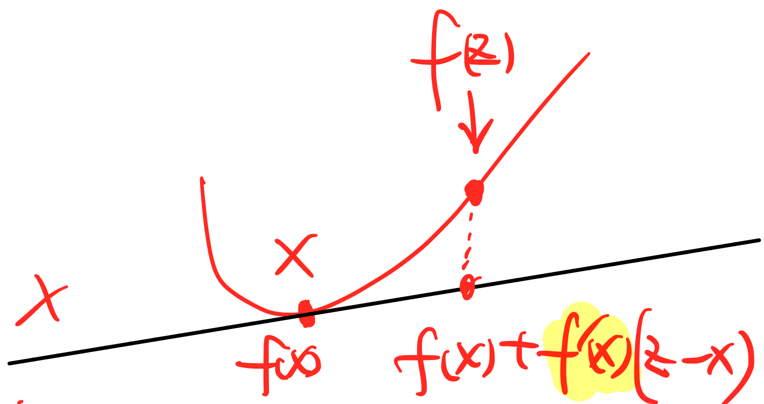
Ax : linear

$Ax+b$: affine

The affine function of z given by

$$f(x) + \underbrace{Df(x)}_{\text{matrix}}\underbrace{(z-x)}_{\text{vector}}$$

is called the first order approximation of f at x.



When z is close to x this affine function is very close to f.

It can be shown that

$$[Df(x)]_{ij} = \frac{\partial f_i(x)}{\partial x_j}$$

$i = 1, \dots, m, \quad j = 1, \dots, n.$

$$\begin{bmatrix} \frac{\partial f_1(x)}{\partial x_1} & \frac{\partial f_1(x)}{\partial x_2} & \cdots & \frac{\partial f_1(x)}{\partial x_n} \\ \vdots & & & \\ \frac{\partial f_m(x)}{\partial x_1} & \frac{\partial f_m(x)}{\partial x_2} & \cdots & \frac{\partial f_m(x)}{\partial x_n} \end{bmatrix}$$

When $f: \mathbb{R}^n \to \mathbb{R}$, $Df(x)$ is a $1 \times n$ matrix.

Its transpose is called the gradient of f

$$\nabla f(x) = Df(x)^T = \begin{pmatrix} \frac{\partial f(x)}{\partial x_1} \\ \frac{\partial f(x)}{\partial x_2} \\ \vdots \\ \frac{\partial f(x)}{\partial x_n} \end{pmatrix} \in \mathbb{R}^n.$$

The first order approximation of f

at $x \in \operatorname{int} \operatorname{dom} f$ is

$$f(x) + \nabla f(x)^T (z - x).$$

Example. $f: \mathbb{R}^n \to \mathbb{R}$. $f(x) = \frac{1}{2} x^T P x + q^T x + r$

where $P \in S^n$, $q \in \mathbb{R}^n$. $r \in \mathbb{R}$

$$Df(x) = x^T P + q^T.$$

$q^T x = \sum_i q_i x_i \qquad \frac{\partial q^T x}{\partial x_j} = q_j \Rightarrow \nabla q^T x = q.$

$x^T P x = \sum_{j=1}^n \sum_{i=1}^n P_{ij} x_i x_j. \qquad \frac{\partial x^T P x}{\partial x_k} = \sum_{j=1}^n P_{kj} x_j + \sum_{i=1}^n P_{ik} x_i$

$$\Rightarrow \nabla x^T P x = Px + Px = 2Px.$$