Continuity & Derivatives

1) Continuous fn
2) Uniformly continuous
3) Lipschitz continuous

① Assume $x, y \in \text{dom} f$. f is continuous at x if $\forall\, \varepsilon > 0\ \exists\ \delta > 0$ s.t.

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

→ δ in general depends on ε, x, y.

② If δ is only a fn of ε, not x, y then f is uniformly continuous.

③ There is some constant L s.t.

$$\|f(x) - f(y)\| \leq L \|x - y\| \quad \forall\, x, y \in \underset{\text{some} \subset \text{Dom}}{\underbrace{N}}$$

e.g. $f(x) = \sqrt{1 - x^2}$ for $x \in [-1, 1]$.

→ Domain $[-1, 1]$. choose $x, y \in [-1, 1]$

Let $|x - y| \leq \delta$. Choose $\underline{\varepsilon = 2\delta}$.

Consider $|f(x) - f(y)|$

→ $\left| \sqrt{1 - x^2} - \sqrt{1 - y^2} \right|^2 = \left| \sqrt{1 - x^2} - \sqrt{1 - y^2} \right| \underbrace{\left| \sqrt{1 - x^2} - \sqrt{1 - y^2} \right|}$

$a, b > 0 \quad |a - b| \leq |a + b|$

$$\leq \left| \sqrt{1 - x^2} - \sqrt{1 - y^2} \right| \left| \sqrt{1 - x^2} + \sqrt{1 - y^2} \right|$$

$$= |x^2 - y^2| = |x - y||x + y| \leq 2\delta = \varepsilon$$

$\therefore$ fn is conts & unif. conts.

$\rightarrow$ Is it Lipschitz conts? We need to find L s.t.

$|f(x) - f(y)| < L|x-y| \quad \forall x, y \in [-1, 1]$

Choose $y = 1$. $f(y) = 0$.

$|f(x)| < L|x-1|$. Consider $x \to 1^-$

$\boxed{\sqrt{1-x^2} < L|x-1|}$

$$\lim_{x \to 1^+} \frac{\sqrt{1-x^2}}{1-x} = \frac{\sqrt{(1-x)(1+x)}}{\sqrt{(1-x)(1-x)}} = \sqrt{\frac{1+x}{1-x}} \to \infty$$

$\Rightarrow$ There is no value L for this fn.

$\hookrightarrow$ Derivatives

Scalar fns $\quad \displaystyle\lim_{y \to 0} \frac{f(x+y) - f(x)}{y} = f'(x) = g$

Rewrite: $\quad \displaystyle\lim_{y \to 0} \frac{f(x+y) - f(x) - gy}{y} = 0$

If this limit exists, g is the derivative of f.

Consider a fn $f: \mathbb{R}^n \to \mathbb{R}$

We say f is differentiable at x if $g \in \mathbb{R}^n$ s.t.

$$\lim_{y \to 0} \frac{f(x+y) - f(x) - g^T y}{\|y\|} = 0 \quad \text{for any } \|\cdot\|.$$

Called the Frechet derivative.

$\nabla f(x) = \begin{bmatrix} \partial f/\partial x_1 \\ \vdots \\ \partial f/\partial x_n \end{bmatrix}$

Simply $f(x,y) = y^2 e^x$

$\nabla f = \begin{bmatrix} \partial f/\partial x \\ \partial f/\partial y \end{bmatrix} = \begin{bmatrix} y^2 e^x \\ 2y e^x \end{bmatrix}$

$\nabla_z f(x,y,z) \longleftrightarrow \frac{\partial}{\partial z} f$

Consider a fn: $f: \mathbb{R}^n \to \mathbb{R}^m$. Extension:

Jacobian J: $J_{ij} = \frac{\partial f_i(x)}{\partial x_j}$ i → row, j → col

$$\begin{array}{c|cccc} & \partial/\partial x_1 & \partial/\partial x_2 & \cdots & \partial/\partial x_n \\ \hline f_1 & & & & \\ f_2 & & & & \\ \vdots & & & & \\ f_m & & & & \end{array}$$

$(x,y) \longrightarrow (r,\theta)$

<u>Chain rule</u>: Say $x = x(t)$. ~~Let h(t)~~

Let $h(t) = f(x(t))$. $\nabla h(t)$?

$$\nabla h(t) = \sum_i \frac{\partial f}{\partial x_i}\frac{\partial x_i}{\partial t} = \underbrace{\nabla x(t)}_{\text{row vect}} \times \underbrace{\nabla f(x(t))}_{\text{col vec}}$$

→ e.g. $f: \mathbb{R}^3 \to \mathbb{R}$, $f(u,v,w) = uv + vw - uw$.

$g: \mathbb{R}^2 \to \mathbb{R}^3$, $g(x,y) = (x+y,\ x+y^2,\ x^2+y)$.

Gradient $(f \circ g)(x,y)$? $\mathbb{R}^2 \to \mathbb{R}$

$$f(u(x,y),\ v(x,y),\ w(x,y)) \quad \text{— fn of } x,y$$

$$\frac{\partial f}{\partial x} = \frac{\partial f}{\partial u}\frac{\partial u}{\partial x} + \frac{\partial f}{\partial v}\frac{\partial v}{\partial x} + \frac{\partial f}{\partial w}\frac{\partial w}{\partial x}$$

$$\frac{\partial f}{\partial y} = \quad " \quad y \qquad y \qquad y$$

$$\begin{pmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{pmatrix}^T = \underbrace{\begin{pmatrix} \frac{\partial f}{\partial u} & \frac{\partial f}{\partial v} & \frac{\partial f}{\partial w} \end{pmatrix} \begin{pmatrix} \frac{\partial u}{\partial x} & \frac{\partial u}{\partial y} \\ \frac{\partial v}{\partial x} & \frac{\partial v}{\partial y} \\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y} \end{pmatrix}}_{x,y}$$

$$\frac{\partial f}{\partial u} = v - w = (x+y^2) - (x^2+y)$$

⋮

→ Quadratic form: $\phi(x): \mathbb{R}^n \to \mathbb{R}$

$$\phi(x) = x^T A x + b^T x + c$$

$c \in \mathbb{R}$, $b \in \mathbb{R}^n$, $A \in \mathbb{R}^{n\times n}$

$$\nabla \phi(x) = (A^T + A)x + b.$$

e.g → linear term $b^T x = \sum_i b_i x_i$

$$\nabla f = \begin{pmatrix} \partial f / \partial x_1 \\ \vdots \\ \partial f / \partial x_n \end{pmatrix} \Rightarrow \nabla (b^T x) = \begin{pmatrix} b_1 \\ \vdots \\ b_i \\ \vdots \\ b_n \end{pmatrix} = b$$

↳ Let $f(x) = Ax \quad : \mathbb{R}^n \to \mathbb{R}^n$. Use Jacobian.

$$f(x) = \begin{pmatrix} \sum A_{1i} x_i \\ \sum A_{2i} x_i \\ \vdots \\ \sum A_{ni} x_i \end{pmatrix} \qquad J_{ij} = \frac{\partial f_i}{\partial x_j}$$

$$\boxed{\nabla (Ax) = A} \quad \Rightarrow \quad \nabla (x) = I.$$

$\nabla (x^T A x)$. Product rule:

$$f(g(x), h(x)) = g^T(x) h(x) \quad \text{defn}$$

Then

$$\nabla f = \nabla h \, g + \nabla g \, h$$

$n \times 1$

Use $g(x) = x, \; h(x) = Ax$

$$\nabla (x^T A x) = x^T A + (Ax)^T \times I = x^T (A + A^T)$$

I WILL FIX!

$$\Rightarrow \boxed{\nabla \phi(x) = (A^T + A)x + b.}$$

Directional derivative.

$$D(f(x), p) = \lim_{\varepsilon \to 0} \frac{f(x+\varepsilon p) - f(x)}{\varepsilon}$$

unit vector

$$= \underbrace{\nabla f(x)^T p}$$

Hessian $\nabla^2 f = \begin{bmatrix} \frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \vdots & & \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \cdots & \frac{\partial^2 f}{\partial x_n} \end{bmatrix}$

<u>Updated proof for $\nabla(x^T A x)$.</u>

say $f(x) = g^T(x)\, h(x)$, both $g, h \in \mathbb{R}^n$

Consider : $g = (g_1(x) \cdots g_n(x))^T$ (col vec)

$$\nabla g = \begin{bmatrix} \partial g_1/\partial x_1 & \cdots & \partial g_1/\partial x_n \\ \partial g_2/\partial x_1 & \cdots & \partial g_2/\partial x_n \\ \vdots & & \vdots \end{bmatrix}$$

Jacobian defn

$$\begin{bmatrix} \vdots & & \vdots \\ \partial g_n/\partial x_1 & \cdots & \partial g_n/\partial x_n \end{bmatrix} \qquad J_{ij} = \frac{\partial f_i}{\partial x_j}$$

Then $\dfrac{\partial f}{\partial x_i} = \dfrac{\partial}{\partial x_i}\left(\sum_j g_j(x)\, h_j(x) \right)$

prod rule →

$$= \sum_j \Bigg[\underbrace{\frac{\partial g_j(x)}{\partial x_i}}_{i^{th}\text{ col of } \nabla g} \times h_j(x) + g_j(x) \underbrace{\frac{\partial h_j(x)}{\partial x_i}}_{i^{th}\text{ col of } \nabla h.} \Bigg]$$

$$= h^T (\nabla g)_{i^{th}\text{ col}} + g^T (\nabla h)_{i^{th}\text{ col}}$$

$\therefore$

$$\left(\frac{\partial f}{\partial x_1} \quad \frac{\partial f}{\partial x_2} \quad \cdots \quad \frac{\partial f}{\partial x_n} \right) = \nabla f^T = h^T \nabla g + g^T \nabla h$$

$$\Rightarrow \quad \nabla f = \nabla g^T h + \nabla h^T g.$$

$\therefore$ with $f = g^T h$ and $g = x$, $h = Ax$

$$\nabla f = (\underbrace{\nabla x}_{I})^T A x + [\underbrace{\nabla (Ax)}_{A}]^T x$$

$$= A x + A^T x.$$

Thus $\nabla \phi(x) = (A + A^T)x + b$

— x —