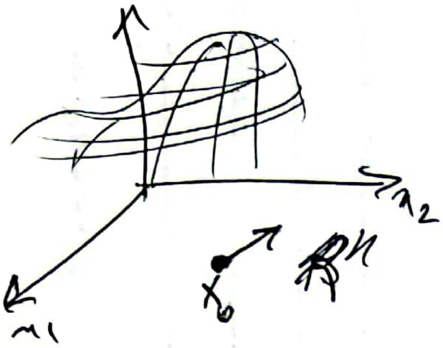
$f: \mathbb{R}^n \to \mathbb{R}$ $\quad f(x_1, x_2, \ldots, x_n) = f\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) = f(x)$



$$D[u]f\Big|_{x_0} = D[u]f(x_0)$$

$$D[u]f(x_0) = m^T u = m(x_0)^T u$$

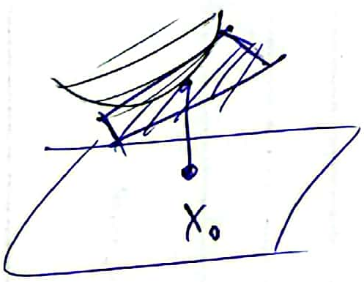
بردار گرادیان

$$= g^T u = \nabla^T u$$

$$\nabla = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x} \end{bmatrix}\Bigg|_{x = x_0}$$

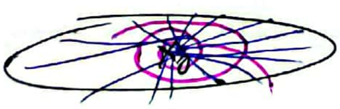
$f: \mathbb{R}^n \to \mathbb{R}$ is differentiable at $x_0 \in \mathbb{R}^n$ if there exists a linear function $\ell: \mathbb{R}^n \to \mathbb{R}$ such that

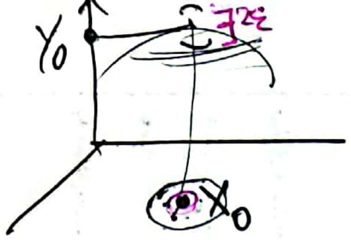
$$\lim_{x \to x_0} \frac{|f(x) - f(x_0) - \ell(x - x_0)|}{\|x - x_0\|} = 0$$



$\exists\, m \in \mathbb{R}^n$

$$\lim_{x \to x_0} \frac{|f(x) - f(x_0) - m^T(x - x_0)|}{\|x - x_0\|} = 0$$





$$\lim_{X \to X_0} f(X) = Y_0$$

MA 18 (II)

$$\forall \varepsilon > 0 \quad \exists \delta \quad \|X - X_0\| < \delta \implies |f(X) - Y_0| < \varepsilon$$

$f : \mathbb{R}^3 \to \mathbb{R}$

$$f(\vec{X}) = f\left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}\right) = x_1 x_2 + x_3 \sin x_2 + x_1 e^{x_3 x_2}$$

$$\partial f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \frac{\partial f}{\partial x_3} \end{bmatrix} = \begin{bmatrix} x_2 + e^{x_3 x_2} \\ x_1 + x_3 \cos x_2 + x_1 x_3 e^{x_3 x_2} \\ \sin x_2 + x_1 x_2 e^{x_3 x_2} \end{bmatrix}$$

~~$f(x)$~~ $f : \mathbb{R}^n \to \mathbb{R}$ $\qquad a \in \mathbb{R}^n$

$$f(X) = a^T X = \begin{bmatrix} a_1 & a_2 & \cdots & a_n \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = a_1 x_1 + a_2 x_2 + \cdots + a_n x_n$$

$$\frac{\partial f}{\partial x_i} = a_i$$

$$\nabla = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ \vdots \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = \vec{a}$$

$\mathrm{Diag} : \mathbb{R}^n \to \mathbb{R}^{n \times n}$

$$\mathrm{Diag}\left(\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) = \begin{bmatrix} x_1 & & \emptyset \\ & x_2 & \\ \emptyset & & \ddots & x_n \end{bmatrix}$$

$f(x)$

$f: \mathbb{R}^n \to \mathbb{R}$ $\qquad u, v \in \mathbb{R}^n$

$$f(x) = \underbrace{u^T}_{1\times n} \underbrace{\mathrm{Diag}(x)}_{n\times n} \underbrace{v}_{n\times 1}$$

$$= [u_1 \; u_2 - u_n] \begin{bmatrix} x_1 & & \emptyset \\ & x_2 & \\ \emptyset & & \ddots \; x_n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix} = x_1 u_1 v_1 + x_2 u_2 v_2 + \cdots + x_n u_n v_n$$

$$\frac{\partial f}{\partial x_1} = u_1 v_1 \qquad \frac{\partial f}{\partial x_i} = u_i v_i$$

$$\nabla = \begin{bmatrix} u_1 v_1 \\ u_2 v_2 \\ \vdots \\ u_n v_n \end{bmatrix} = \mathrm{Diag}(u)\, v = \mathrm{Diag}(v)\, u = u \odot v$$

$f: \mathbb{R}^n \to \mathbb{R}$ $\qquad f(x) = \underbrace{\underbrace{x^T}_{1\times n} \underbrace{A}_{n\times n} \underbrace{x}_{n\times 1}}_{1\times 1}$ $\qquad A \in \mathbb{R}^{n\times n}$

$$f(x) = [x_1 \; x_2 - x_n] \begin{bmatrix} a_{11} & a_{12} & a_{1n} \\ | & \vdots & | \\ a_{n1} & a_{n2} & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ | \\ x_n \end{bmatrix} = \sum_{i=1}^{n} \sum_{j=1}^{n} x_i x_j a_{ij}$$

$$\frac{\partial f}{\partial x_k} = \frac{\partial}{\partial x_k}\Big(\cancel{\sum_{\substack{i=1\\ i\neq k}}^{n} \sum_{\substack{j=1\\ j\neq k}}^{n} x_i x_j a_{ij}} + \sum_{\substack{j=1\\ j\neq k}}^{n} x_k x_j a_{kj} + \sum_{\substack{i=1\\ i\neq k}}^{n} x_i x_k a_{ik} + x_k^2 a_{kk} \Big)$$

$\emptyset$

$$= \sum_{\substack{j=1\\ j\neq k}}^{n} x_j a_{kj} + \sum_{\substack{i=1\\ i\neq k}}^{n} x_i a_{ik} + \underbrace{2 a_{kk} x_k}_{a_{kk}x_k + a_{kk}x_k}$$

$$= \sum_{j=1}^{n} a_{kj} x_j + \sum_{i=1}^{n} a_{ik} x_i$$

MA18 IV

$$\frac{\partial f}{\partial k} = \begin{bmatrix} a_{k1} & a_{k2} & \cdots & a_{kn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ | \\ x_n \end{bmatrix} + \begin{bmatrix} a_{1k} & a_{2k} & \cdots & a_{nk} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ | \\ x_n \end{bmatrix}$$

$$\nabla f = \begin{bmatrix} \frac{\partial f}{\partial x_1} \\ \frac{\partial f}{\partial x_2} \\ | \\ \frac{\partial f}{\partial x_n} \end{bmatrix} = \begin{bmatrix} a_{11} & a_{12} & \text{—} & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ | & | & & | \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ | \\ x_n \end{bmatrix} + \begin{bmatrix} a_{11} & a_{21} & \text{—} & a_{n1} \\ a_{12} & a_{22} & & a_{n2} \\ & & & \\ a_{1n} & a_{2n} & \text{—} & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ | \\ x_n \end{bmatrix}$$

$$= Ax + A^T x = (A + A^T)x$$

$f(x) = x^T A x$ $\quad \nabla f = \underbrace{\underbrace{(\underbrace{A}_{n\times n} + \underbrace{A^T}_{n\times n})}_{n\times n} \underbrace{x}_{n\times 1}}_{n\times 1}$

A symmetric $\Rightarrow \nabla f = 2Ax$

$$x^T A x = x^T \left(\frac{A + A^T}{2}\right) x$$

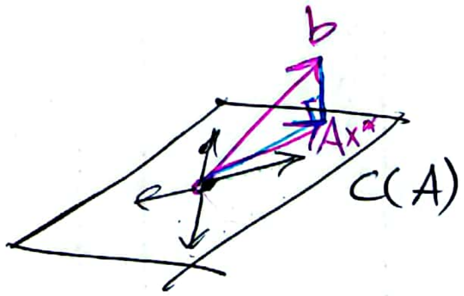
Least Squares Problem

$Ax = b$

A: $m \times n$

$m > n$

$$\begin{bmatrix} \\ A \\ \\ \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} \\ b \\ \\ \end{bmatrix}$$



$$x^* = \underset{x}{\operatorname{argmin}} \|Ax - b\| \Rightarrow x^* = (A^T A)^{-1} A^T b$$

$f:\mathbb{R}^n \to \mathbb{R} \qquad f(x)=\|Ax-b\|$ MAT 18

$f:\mathbb{R}^n \to \mathbb{R} \qquad f(x)=\|\underbrace{\underbrace{A}_{m\times n}\underbrace{x}_{n\times 1}-\underbrace{b}_{m\times 1}}_{m\times 1}\|^2 \qquad \|v\|=\sqrt{v^Tv}$

$$f(x)=(Ax-b)^T(Ax-b)=(x^TA^T-b^T)(Ax-b)$$
$$=x^TA^TAx-b^TAx-\underbrace{x^T}_{1\times n}\underbrace{A^T}_{n\times m}\underbrace{b}_{m\times 1}+b^Tb$$
$$=x^TA^TAx-b^TAx-b^TAx+b^Tb \qquad \nabla a^Tx=a$$

$$\nabla f = 2A^TAx-A^Tb-A^Tb=2A^TAx-2A^Tb$$

$$\nabla f=0 \Rightarrow A^TAx=A^Tb \Rightarrow x^*=(A^TA)^{-1}A^Tb$$

Second method of Calculating Gradient

$f(x) \qquad D[u]f(x)=u^Tg=\langle u, \underline{\underline{g}}\rangle$

$\cancel{f(x)=a^Tx} \qquad D[u]f(x)=\frac{d}{d\alpha}f(x+\alpha u)\Big|_{\alpha=0}$

$h(\alpha)=f(x+\alpha u) \qquad f:\mathbb{R}^n\to\mathbb{R}$

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

$f(x)=a^Tx \qquad f(x+\alpha u)=a^T(x+\alpha u)=a^Tx+\alpha a^Tu$

$$\frac{d}{d\alpha}f(x+\alpha u)=a^Tu=\langle a,u\rangle \qquad \nabla=a$$

$$\langle Ax,y\rangle=\langle x,A^Ty\rangle=(Ax)^Ty=x^TA^Ty$$
$$\langle A,B\rangle=\mathrm{trace}(A^TB)=\mathrm{trace}(B^TA)$$
$$\langle A,BC\rangle=\langle B^TA,C\rangle=\langle AC^T,B\rangle$$

$f(x) = u^T \mathrm{Diag}(x)\, v$ MA12 (VI)

$$D[d]f(x) = \frac{d}{d\alpha} f(x+\alpha d) = \frac{d}{d\alpha} u^T \mathrm{Diag}(x+\alpha \vec{d})\, v$$

$$= \frac{d}{d\alpha}\left(u^T \mathrm{Diag}(x)\, v + \alpha\, u^T \mathrm{Diag}(d)\, v\right)$$

$$= u^T \mathrm{Diag}(d)\, v = \underbrace{\underbrace{u^T}_{1\times n} \underbrace{\mathrm{Diag}(v)}_{n\times n}}_{1\times n} d$$

$$= (\mathrm{Diag}(v)\, u)^T d = \langle \underbrace{\mathrm{Diag}(v) u}_{\nabla}, d \rangle$$

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

$$\frac{d}{d\alpha} f(x+\alpha u) = \frac{d}{d\alpha}(x+\alpha u)^T A\, (x+\alpha u) \Big|_{\alpha=0}$$

$$= u^T A(x+\alpha u) + (x+\alpha u)^T A\, u \Big|_{\alpha=0}$$

$$= \underbrace{u^T A x}_{1\times 1} + x^T A u$$

$$= x^T A^T u + x^T A\, u = (Ax)^T u + (A^T x)^T u$$

$$= \langle Ax + A^T x, u \rangle$$

$$= \langle \underbrace{(A+A^T)x}_{\nabla}, u \rangle$$

$$f(x) = \|Ax-b\|^2 = (Ax-b)^T(Ax-b)$$

$$h(\alpha) = f(x+\alpha u) = (A(x+\alpha u) - b)^T (A(x+\alpha u) - b)$$