MFDS - Session 7

1 $f : \mathbb{R} \to \mathbb{R}$

$f(x) = \sin(x)$ $\quad \frac{d}{dx} f(x) = \cos x$

$f(x) = \cos(x)$ $\quad \frac{d}{dx} f(x) = -\sin x$

2 $f : \mathbb{R}^2 \to \mathbb{R}$ $\quad f(\underbrace{x_1, x_2}_{\in \mathbb{R}^2}) = \boxed{\underbrace{x_1 + x_2}_{\in \mathbb{R}}}$

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

$\frac{\partial f}{\partial x_1} = 1 = \frac{\partial f}{\partial x_2}$

$f(\underbrace{x_1, x_2, \ldots x_n}_{\in \mathbb{R}^n}) = \underbrace{x_1^2 + x_2^2 + x_1 x_3 + \ldots + x_{n-1} x_n}_{\in \mathbb{R}}$

$\frac{\partial f}{\partial x_i}$ $1 \leq i \leq n$, $\boxed{\frac{\partial^2 f}{\partial x_i \partial x_j} \quad 1 \leq i, j \leq n}$

$f(\underbrace{x_1, x_2}) = x_1^2 + x_1 x_2 + x_2^2$ $\quad f : \mathbb{R}^2 \to \mathbb{R}$

$\frac{\partial f}{\partial x_1} = 2x_1 + x_2$, $\quad \frac{\partial^2 f}{\partial x_1^2} = 2$

$\frac{\partial^2 f}{\partial x_2 \partial x_1} = 1$ $\quad \frac{\partial^3 f}{\partial x_i \partial x_j \partial x_k}$ $\quad 1 \leq i, j, k \leq n$

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

$f(x_1, x_2, \ldots x_n) = (y_1, y_2, \ldots y_m)$

f $\mathbb{R}^2 \to \mathbb{R}^3$

$f(x_1, x_2) = (\underbrace{x_1 + x_2}, \underbrace{x_2}, \underbrace{x_1 - x_2})$

$$f(\underbrace{x_1, x_2}_{\in \mathbb{R}^2}) = \underbrace{(x_1 + x_2,\ x_2,\ x_1 - x_2)}_{\in \mathbb{R}^3}$$

$$g(\underbrace{x}_{\in \mathbb{R}^n}) = A_{m \times n}\, x_{n \times 1} \quad \text{(output is } \underbrace{m \times 1}\text{)}$$

$$g : \mathbb{R}^2 \to \mathbb{R}^3$$

$$g\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \overset{A}{\begin{pmatrix} 2 & -1 \\ 3 & 1 \\ 4 & 2 \end{pmatrix}} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2x_1 - x_2 \\ 3x_1 + x_2 \\ 4x_1 + 2x_2 \end{pmatrix}$$

$$\boxed{a_{11} x_1 + a_{12} x_2 + \quad + a_{1n} x_n} = b_1$$

(blood, urine) → (FBS, PPBS, HbA1c3, Sr the . . .

$\in \mathbb{R}^2$ . . . . , Albm, prth . . .)

$\in \mathbb{R}^{25}$

Conventions vectors $a, \underline{v}, w$ are all column vectors

∴ a^T, v^T, w^T row vectors

$$\underset{\sim}{a} = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad \text{\& so on}$$

1. Gradient of a linear function

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

$$f(w) = a^T w \quad \text{for some fixed } a$$

Properties $f : V \to W$ is linear if

Properties $f : V \to W$ is linear if

$$f(v_1 + v_2) = f(v_1) + f(v_2)$$
$$f(c v_1) = c f(v_1)$$

$= $ Scalar

$$\frac{d}{dx}\big(f(x) + g(x)\big) = \frac{d}{dx} f(x) + \frac{d}{dx} g(x)$$

$$\frac{d}{dx}\big(c f(x)\big) = c \frac{d}{dx} f(x)$$

Calculate the gradient of f

$$\nabla f = \left[\frac{\partial f}{\partial w_1}, \frac{\partial f}{\partial w_2}, \ldots \frac{\partial f}{\partial w_n} \right]$$

$$f = a^T w = [a_1, a_2 \ldots a_n] \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_n \end{bmatrix}$$

$$= a_1 w_1 + a_2 w_2 + \ldots + a_n w_n$$

$$\frac{\partial f}{\partial w_i} = a_i \qquad 1 \le i \le n$$

$$\nabla f = [a_1, a_2, \ldots a_n] = a^T$$

[As per the text book]

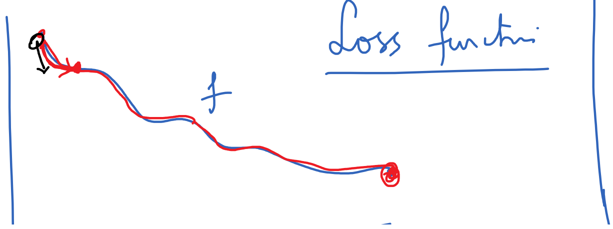
In other books $\nabla f = \begin{bmatrix} \partial f / \partial w_1 \\ \partial f / \partial w_2 \\ \partial f / \partial w_n \end{bmatrix} \to$ Column vector

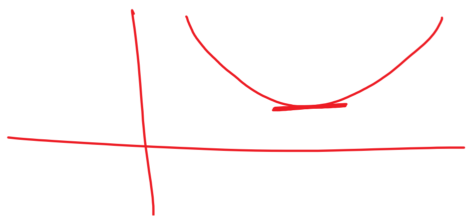
$$\nabla f = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} = a$$

$\underline{\nabla f} \to$ direction of maximum increase.

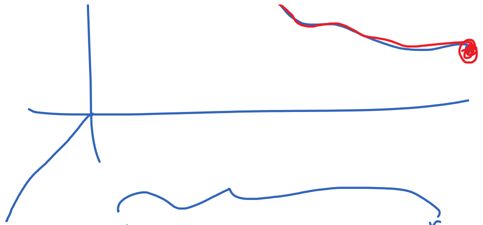
$-\nabla f$

Loss function





Convex optimization



x_1	x_2	x_3	.	x_7	x_8
5	2	1 . . .		3	11 cr
2	½	½	.	1	24 cr
:					
2.6	1.32	0.67	.	0.22	?

1000

$y_1 = 17.5$

$y_2 = 17.5$

Predicted output for x_1 is

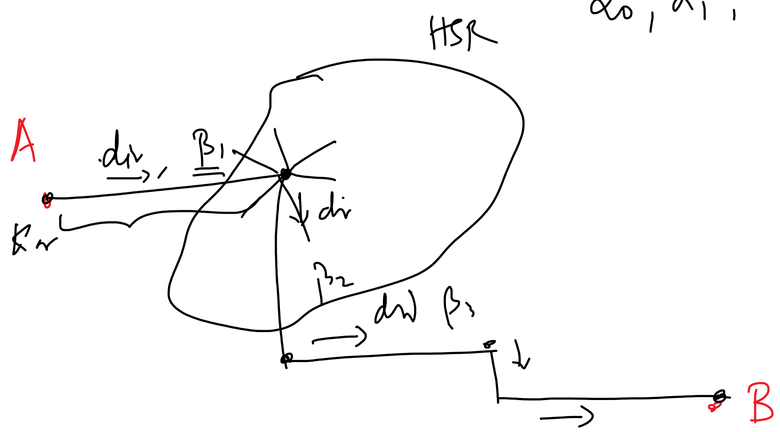
$$\boxed{\underbrace{\alpha_0 + \alpha_1 x_1 + \alpha_2 x_2 + \dots + \alpha_7 \cdot x_7}} = y_1$$

$\alpha_0 + \alpha_1 \cdot 5 + \alpha_2 \cdot 1 + \dots + \alpha_7 \cdot 3 = y_1$

$$\text{error}^2 = (y_1 - 11)^2 + (y_2 - 24) + \dots + (y_{10000} - 28)$$

Min error w.r.t the parameters

$\alpha_0, \alpha_1, \dots \alpha_7$



2

$$f(w) = \underline{w^T A w} \qquad A_{n \times n}$$

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

$$A = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{pmatrix} \qquad w = \begin{pmatrix} w_1 \\ \vdots \\ w_n \end{pmatrix}$$

$w \in \mathbb{R}^n$

$$\begin{bmatrix} \quad w^T \quad \end{bmatrix} A_{n \times n} \begin{bmatrix} w \end{bmatrix}_n$$

$$f(w) = w^T (Aw) = w^T \begin{pmatrix} a_{11} w_1 + a_{12} w_2 + \dots + a_{1n} w_n \\ a_{21} w_1 + a_{22} w_2 + \dots + a_{2n} w_n \\ a_{n1} w_1 + a_{n2} w_2 + \dots a_{nn} w_n \end{pmatrix}$$

$$f(w) = w^T(Aw) = w^T \begin{pmatrix} \vdots \\ a_{21}w_1 + a_{22}w_2 + \cdots + a_{2n}w_n \\ a_{n1}w_1 + a_{n2}w_2 + \cdots a_{nn}w_n \end{pmatrix}$$

$$= w^T \begin{pmatrix} \sum_{j=1}^{n} a_{1j} w_j \\ \sum_{j=1}^{n} a_{2j} w_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj} w_j \end{pmatrix}$$

$$= \begin{pmatrix} \underline{w_1} & \underline{w_2} & \cdots & w_n \end{pmatrix} \begin{pmatrix} \sum_{j=1}^{n} a_{1j} w_j \\ \sum_{j=1}^{n} a_{2j} w_j \\ \vdots \\ \sum_{j=1}^{n} a_{nj} w_j \end{pmatrix}$$

$$= w_1 \sum_{j=1}^{n} a_{1j} w_j + w_2 \sum_{j=1}^{n} a_{2j} w_j + \cdots + w_n \sum_{j=1}^{n} a_{nj} w_j$$

$$= \sum_{i=1}^{n} w_i \sum_{j=1}^{n} a_{ij} w_j$$

$$= \sum_{i=1}^{n} \sum_{j=1}^{n} w_i a_{ij} w_j = \sum_{i=1}^{n} w_i^2 a_{ii} + \sum_{i \neq j} w_i a_{ij} w_j$$

$$w = \begin{pmatrix} 1 \\ 2 \end{pmatrix};\ A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

$w_1 = 1$

$w_2 = 2$

$$\sum_{i=1}^{2} \sum_{j=1}^{2} w_i a_{ij} w_j$$

$$\underline{w_1 a_{11} w_1} + \underline{w_1 a_{12} w_2} + \underline{w_2 a_{21} w_1} + \underline{w_2 a_{22} w_2}$$

$f(x) = x^2$

$+ \quad w_2 a_{21} w_1 + \cdots$

$f(x) = x^2$

$\frac{df}{dx} = 2x$

$$f(w) = \left(\sum_{i=1}^{n} w_i^2 a_{ii}\right) + \left(\sum_{\substack{j \neq i \\ j=1 \\ i=1}}^{n,n} w_i a_{ij} w_j\right)$$

$$\boxed{\frac{\partial f}{\partial w_k} = 2 w_k a_{kk} + \sum_{j \neq k} w_j a_{jk} + \sum_{k \neq j} a_{kj} w_j}$$

$(a^T w)$

$$f(w) = \sum_{i=1}^{n} a_i w_i$$

$$\frac{\partial f}{\partial w_k} = \sum_{i=1}^{n} a_i \delta_{ik} \qquad \delta_{ik} = \begin{cases} 1 & \text{if } i = k \\ 0 & \text{otherwise} \end{cases}$$

$$= a_k$$

$$\frac{\partial f}{\partial w_k} = \sum_{j=1}^{n} w_j a_{jk} + \sum_{j=1}^{n} a_{kj} w_j$$

$$\nabla f = A^T w + A w = (A^T + A) w$$

$$f = w^T A w, \quad \nabla f = (A^T + A) w$$

If A is Sy $\quad \nabla f = 2Aw$

$$X = \boxed{\sum_{\substack{j \neq i \\ j=1 \\ i=1}}^{n\,n} w_i a_{ij} w_j} = w_1 (a_{12} w_2 + a_{13} w_3 + \cdots + a_{1n} w_n)$$

$$+ \, w_2 (a_{21} w_1 + a_{23} w_3 + \cdots + a_{2n} w_n)$$

$$+ \, w_3 (a_{31} w_1 + a_{32} w_2) + \cdots$$

$$+ \, w_n (a_{n1} w_1 + a_{n2} w_2 + \cdots + a_{n\,n-1} w_{n-1})$$

$$\frac{\partial X}{\partial w_2} = \underline{w_1 a_{12}} + \underline{a_{21} w_1 + a_{23} w_3 + \dots + a_{2n} w_n} + \underline{w_3 a_{32}} + \underline{w_4 a_{42}} + \dots + \underline{w_n a_{n2}}$$

$$\begin{bmatrix} c_{11} & c_{12} \\ c_{21} & c_{22} \end{bmatrix}_{2\times 2} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix}$$

$$= (a_{11} b_{12} + a_{12} b_{22})$$

$$B \in \mathbb{R}^{m \times n} \qquad K = B^T B \in \mathbb{R}^{n \times n}$$

~~Let $B = \begin{bmatrix} \leftarrow r_1 \rightarrow \\ \leftarrow r_2 \rightarrow \\ \leftarrow r_n \rightarrow \end{bmatrix}$~~

K_{pq} . ~~p^{th} row of B^T~~ $\times$ q^{th} column of B

$r_p^T \times q^{th}$ column of B

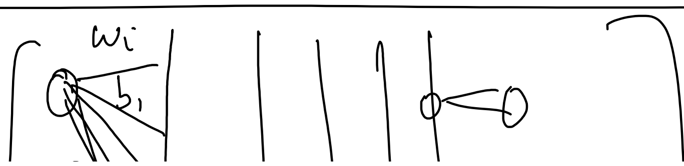
$r_p^T \times r_q$

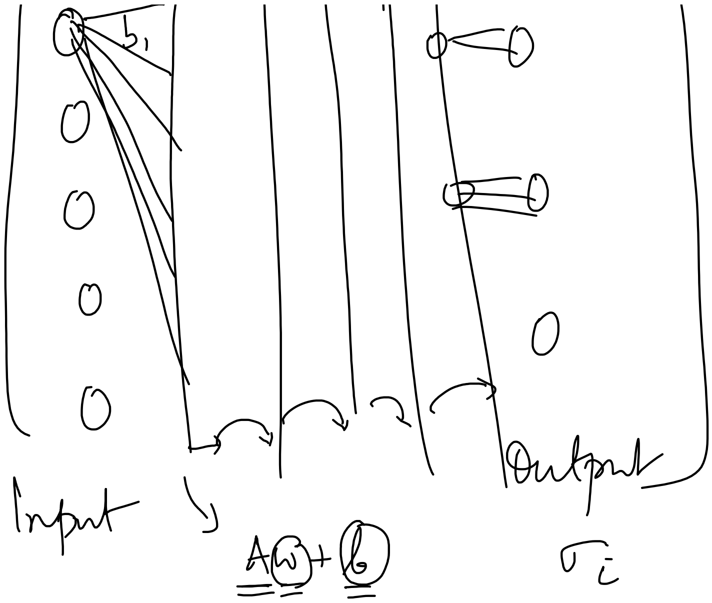
r_i be the i^{th} column of B

$$\begin{bmatrix} \uparrow & \uparrow & & \uparrow \\ r_1 & r_2 & \cdots & r_n \\ \downarrow & \downarrow & & \downarrow \end{bmatrix}$$

K_{pq} $\underline{p^{th} \text{ row of } B^T}$ $\times$ q^{th} col of B

$r_p^T \times r_q$




b1
Input
Output
AW + b
σi