5/10/20

EC 414

Ivan Isakov

Midterm 1

Problem 1

$f(z) = z^2$, $A[-1, 1]$

$\text{argmin} \frac{1}{13 + \sqrt{1+2f(z)}}$

$z^* =$

$= \text{argmin}\left(\frac{1}{13+\sqrt{1+2f(z)}}\right)$

$= \frac{d}{dz}\left(\frac{1}{13+\sqrt{1+2f(z)}}\right)$

$= -\left(13+\sqrt{1+2z^2}\right)^{-2} \times$

$\times \frac{d}{dz}\left(13+\sqrt{1+2z^2}\right)$

$\downarrow$

$$\frac{1}{\cancel{2}}(1+2z^2)^{-\frac{1}{2}} \times \cancel{4}z$$

$$\Rightarrow -\left(13+\sqrt{1+2z^2}\right)^{-2} \times \left(\frac{2z}{\sqrt{1+2z^2}}\right)$$

$$= -\frac{2z}{\sqrt{1+2z^2}\left(13+\sqrt{1+2z^2}\right)^2} = 0$$

$z_1 = 0 \rightarrow f(z_1) = 0$

$z_2 = -1 \rightarrow f(z_2) = 5.32 \times 10^{-3}$

$z_3 = 1 \rightarrow f(z_3) = -5.32 \times 10^{-3}$ ✓

so $\operatorname{argmin}(f(z)) = 1$

Problem 2

$Y = \{1.5, 2.0, 3.5, 6.0\}$

$y_1 = 1.5$

$y_2 = 1.5$

$y_3 = 2.0$

$y_4 = 3.5$

$y_5 = 3.5$

$y_6 = 6.0$

a) $\text{argmin}\left(\frac{1}{6}\sum_{j=1}^{6} 1[y \neq y_i]\right)$

$\rightarrow$ 0/1 loss

Mode of the set

$x^* = \{1.5, 3.5\}$

b) $\text{argmin}\left(\frac{1}{6}\sum_{j=1}^{6} (y - y_i)^2\right)$

Mean of the set

$$x^* = \frac{2 \times 1.5 + 2 + 2 \times 3.5 + 6}{6}$$

$= 3.0$

c) $\text{argmin}\left(\frac{1}{6}\sum_{j=1}^{6} |y - y_i|\right.$

Median of the set

$x^* \in [2.0, 3.5]$

<u>Problem 3</u>

$b = 3$

$w = [1, -4, 8]^T$

a) $w^T x + b = 0$

i) $x_1 = [-2, 2, 1]^T$

$\Rightarrow w^T x_1 + b = \begin{bmatrix} 1 & -4 & 8 \end{bmatrix} \begin{bmatrix} -2 \\ 2 \\ 1 \end{bmatrix} + 3 =$

$= -2 \times 1 - 2 \times 4 + 8 \times 1 + 3 = 1 \neq 0$

x_1 does <u>not</u> lie on the plane

ii) $x_2 = [0, 1, 0]^T$

$\Rightarrow \begin{bmatrix} 1 & -4 & 8 \end{bmatrix} \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix} + 3 = 0$

$= -4 \times 1 + 3 = -1 \neq 0$

so x_2 does not lie on

the plane

iii) $x_3 = [1, 3, 1]^T$

$$[1 \ -4 \ 8] \begin{bmatrix} 1 \\ 3 \\ 1 \end{bmatrix} + 3 =$$

$$= 1 - 4 \times 3 + 8 + 3 = 0$$

so x_3 does lie on the plane

b) $$d = \frac{|w^T x_4 + b|}{\|w\|_2} =$$

$$= \frac{[1 \ -4 \ 8] \begin{bmatrix} -1 \\ -1 \\ -1 \end{bmatrix} + 3}{\sqrt{1^2 + 4^2 + 8^2}}$$

$$= -\frac{2}{9} \longrightarrow \frac{2}{9}$$

c) $(x - x) // w$ ①

$c)\ (x_4 - x_p) \parallel \omega$

$$\omega^T x_p + b = 0 \quad (2)$$

(1)

$$\frac{x_{4,1} - x_{p,1}}{\omega_1} = \frac{x_{4,2} - x_{p,2}}{\omega_2} =$$

$$= \frac{x_{4,3} - x_{p,3}}{\omega_3} = \alpha$$

$$\frac{-1 - x_{p,1}}{1} = \frac{-1 - x_{p,2}}{-4} = \frac{-1 - x_{p,3}}{8} =$$

$$= \alpha$$

$$\Rightarrow \quad \begin{aligned} -1 - x_{p,1} &= \alpha \\ -1 - x_{p,2} &= -4\alpha \\ -1 - x_{p,3} &= 8\alpha \end{aligned}$$

$$\Rightarrow \quad \begin{aligned} x_{p,1} &= -\alpha - 1 \\ x_{p,2} &= 4\alpha - 1 \\ x_{p,3} &= -8\alpha - 1 \end{aligned}$$

$$(2) \rightarrow \begin{bmatrix} 1 & -4 & 8 \end{bmatrix} \begin{bmatrix} -\alpha - 1 \\ \ldots \end{bmatrix} + 5 = 0$$

$$\begin{bmatrix} 4\alpha-1 \\ -8\alpha-1 \end{bmatrix}$$

$$\Rightarrow (-\alpha-1)-4(4\alpha-1)+8(-8\alpha-1)+$$
$$+3=0$$

$$\Rightarrow -81\alpha-2=0$$
$$\alpha=-\frac{2}{81}$$

$$\Rightarrow x_p=\begin{bmatrix} -\frac{79}{81} \\ -\frac{89}{81} \\ -\frac{65}{81} \end{bmatrix}$$

d) $x_5=\begin{bmatrix} 1/2 \\ 0 \\ 0 \end{bmatrix}$ $\quad x_6=\begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}$

$$x_7=\begin{bmatrix} -1 \\ 1 \\ -1 \end{bmatrix}$$

$$w^T x_5 + b = 0$$
$$w^T x_6 + b = 0 \qquad w=\begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

$$w^T x_7 + b = 0$$

① $\frac{1}{2} w_1 + b = 0$

② $w_1 + w_2 + b = 0$

③ $-w_1 + w_2 - w_3 + b = 0$

3 eqns, 4 unknowns so
set $b = 1$

$$\Rightarrow \quad w_1 = -2$$
$$w_2 = 1$$
$$w_3 = 4$$

$$w = \begin{bmatrix} -2 \\ 1 \\ 4 \end{bmatrix}, \quad b = 1$$

Problem 4

$$x_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad x_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \quad x_3 = \begin{bmatrix} -3 \\ 1 \end{bmatrix}$$

$$y_1 = 4 \quad y_2 = 2 \quad y_3 = -8$$

$$x_4 = \begin{bmatrix} 4 \\ 3 \end{bmatrix}$$

$$y_4 = 2$$

a) $b = 0$

$$\hat{y} = w^T x + b = w^T x$$

$$\underline{X} = \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ -3 & 1 \\ 4 & 3 \end{bmatrix} \qquad y = \begin{bmatrix} 4 \\ 2 \\ -8 \\ 2 \end{bmatrix}$$

$$\underline{X}^T \underline{X} = \begin{bmatrix} 1 & 2 & -3 & 4 \\ -1 & 1 & 1 & 3 \end{bmatrix} \begin{bmatrix} 1 & -1 \\ 2 & 1 \\ -3 & 1 \\ 4 & 3 \end{bmatrix}$$

$$= \begin{bmatrix} 30 & 10 \\ 10 & 12 \end{bmatrix}$$

$$\left(\underline{X}^T \underline{X}\right)^{-1} = \frac{1}{260} \begin{bmatrix} 12 & -10 \\ -10 & 30 \end{bmatrix}$$

$$\left(\underline{X}^T \underline{X}\right)^{-1} \underline{X}^T = \frac{1}{260} \begin{bmatrix} 22 & 14 & -46 & 18 \\ -40 & 10 & 60 & 50 \end{bmatrix}$$

$$(X^T X)^{-1} X^T y = w_{OLS}$$

$$= \frac{1}{260}\begin{bmatrix} 520 \\ -520 \end{bmatrix}$$

b) $\hat{y} = w_{OLS}^{*T} x + \cancel{b}^{\,0} = w_{OLS}^{*T} x$

$$x_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\hat{y} = \frac{1}{260}\begin{bmatrix} 520 \\ -520 \end{bmatrix}^T \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$= \frac{1}{260} \times 1040 = 4$$

Problem 5

a) $\nabla g(w) = \frac{\partial}{\partial w} g(w)$

$$\frac{\partial}{\partial w}\left(\sum_{i=1}^{d} w_i^2 + \frac{1}{2m}\sum_{j=1}^{m}\left(y_j - \sum_{i=1}^{d} x_{j,i}\, w_i\right)^2\right)$$

$$= 2w_i + \frac{1}{\cancel{2}m}\sum \cancel{2}\left(y_j - \sum x_{j,i}\, w_i\right)(-x_{j,i})$$

$$= 2w_i - \sum_1^d x_{ji}(y_j - \sum_1^1 x_{j,i} w_i)$$

$$= 2w - X^T 2(y - Xw)$$

$$= 2w - 2X^T(y - Xw)$$

b) Find $w^* = \arg\min g(w)$ with GD

Initialise w_0

for $t = 1, \dots T$

Calculate gradient wrt w

update $w \to w_t = w_{t-1} -$

$- \eta \nabla_w g(w)$

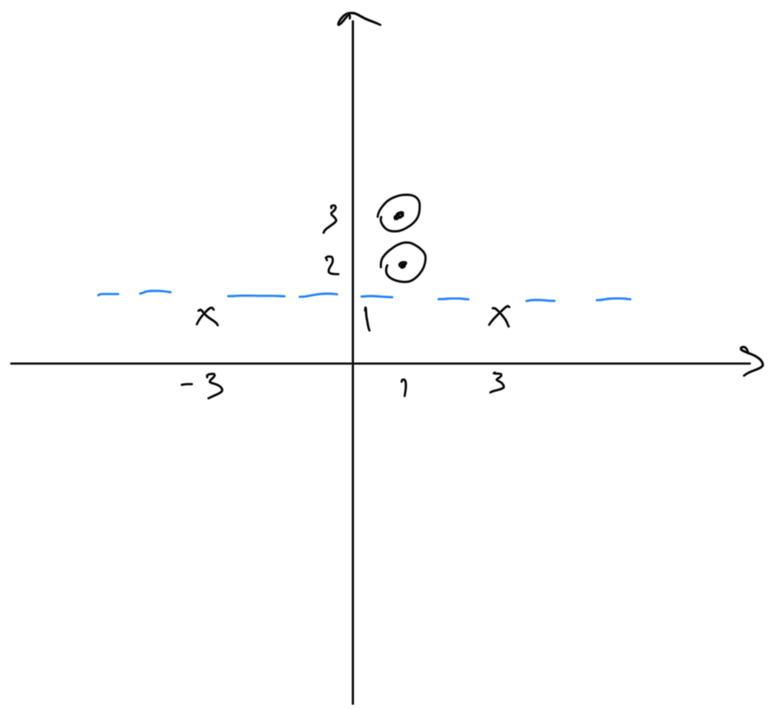
end for

return w

c) Yes, it can happen if the algorithm has reached the minimum, where gradient is 0. So, it won't need to update anymore

Problem 6

a)



b) It is possible, but it will be difficult since the margin is quite small

c) Depends, since there is a possibility that you can separate the dataset, but the margin is small

d) $\tilde{x}_1 = \begin{bmatrix} 1 \\ x_1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix}$

$\tilde{w}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \qquad y_1 = -1$

$y_1 \tilde{w}_1^T \tilde{x}_1 = -1 \begin{bmatrix} 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} = 0$

$\hookrightarrow \tilde{w}_2 = \tilde{w}_1 + y_1 \tilde{x}_1 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 2 \end{bmatrix} =$

$= \begin{bmatrix} -1 \\ -1 \\ -2 \end{bmatrix}$

Problem 7

$f(x \quad x \;) = \frac{1}{\;} x^2 + \frac{1}{\;} x^2 +$

$$f(x_1, x_2) = \frac{1}{2}x_1^2 + \frac{1}{5}x_2^2 +$$

$$+ \frac{1}{4}\sin(2x_1)$$

This function is not convex. $\frac{1}{2}x_1^2$ and $\frac{1}{5}x_2^2$ are but the $\frac{1}{4}\sin(2x)$ is not convex which makes whole $f(x, x_2)$ non-convex