HW 3 Josh Liu

zhaoliu8@illinois.

1-a proof

1.a. Let $\underset{80\times 2}{\underline{x}}$ = the data matrix,

let $\underset{80\times1}{\underline{1}_{80}} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

Let $\underline{X} = (\underline{x}\ \underline{1}_{80})$. $\underset{80\times1}{\underline{Y}} = \begin{pmatrix} 1 \\ \vdots \\ 1 \\ -1 \\ \vdots \\ -1 \end{pmatrix}$

In the objective function

$$\frac{1}{2}\underline{\beta}'\underline{D}\underline{\beta} - \underline{d}'\beta ,$$

we don't have a linear term

for $\min \|\underline{\beta}\|^2$,

so we set

$$\underline{d} = \begin{pmatrix} 0 \\ 0 \\ 0 \end{pmatrix}$$

Note that $\underset{3\times1}{\underline{\beta}}$ ~~is~~ has 3 elements, and $\underline{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_0 \end{pmatrix}$

However, because we only want to restrict $\beta_1^2+\beta_2^2$, we set

$$\underline{D} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 10^{-5} \end{pmatrix}$$

Now, let's have a look at the constraint

$$y_i(\underline{x}_i'\underline{\beta} + \beta_0) \geq 1$$

for $i = 1, 2, \ldots, 80$

Construct another matrix $\underset{80\times 80}{\underline{G}}$ such that $\underline{G}$ is a diagonal matrix, and $\mathrm{diag}(\underline{G}) = \underline{Y}$

Then $y_i(\underline{x}_i'\underline{\beta} + \beta_0) \geq 1$ can be written, with our 3-di $\underline{\beta}$, as:

$$y_i\underline{x}_i'\underline{\beta} + y_i\beta_0 \geq 1$$

$$\Rightarrow y_i\underline{X}_i'\underline{\beta} \geq 1$$

where $\underline{\beta} = \begin{pmatrix} \beta_1 \\ \beta_2 \\ \beta_0 \end{pmatrix}$ a

$$\underset{80\times3}{\underline{X}} = (\underline{x}\ \ \underline{1}_8$$

$$\underline{G}\,\underline{X}\,\underline{\beta} \geq \cancel{1}\begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$$

Therefore, $\underline{A} = (\underline{G}\underline{X})'$ and $\underline{b}^0 = \underline{1}$

1-b.

Let $\underset{80\times1}{\underline{\alpha}} = \begin{pmatrix}\alpha_1\\ \alpha_2\\ \vdots\\ \vdots\\ \alpha_{80}\end{pmatrix}$. Given that

$$\underset{80\times1}{\underline{y}} = \begin{bmatrix} \left.\begin{matrix}1\\ \vdots\\ 1\end{matrix}\right\}40 \\ \left.\begin{matrix}-1\\ -1\\ \vdots\\ -1\end{matrix}\right\}40 \end{bmatrix}$$, and $\sum y_i \cdot \alpha_i = 0$

we have $\underline{y}'\underline{\alpha} = 0$

Also, construct a 80×80 matrix $\underline{G}$, such that $diag(\underline{G}) = \underline{y}$, and all off-degond elements are 0.

Also, let $\underset{80\times1}{\underline{1}_{80}} = \begin{bmatrix}1\\ \vdots\\ 1\end{bmatrix}$

Then the objective function becomes

$\frac{1}{2}(\underline{G}\underline{\alpha})'\underline{x}\underline{x}'\underline{G}\alpha - \underline{1}_{80}'\alpha$

~~$\Rightarrow \underline{\beta} = \underline{\alpha}$~~;

$= \frac{1}{2}\underline{\alpha}'\underline{G}'\underline{x}\underline{x}'\underline{G}\underline{\alpha}$

$\Rightarrow \underline{\beta} = \underline{\alpha}$; $\underline{D} = \underline{G}'\underline{x}\underline{x}'\underline{G}$

$\underline{d} = \underline{1}_n$

For the constraints, we have

$\underline{y}'\underline{\alpha} = 0$, and $\underline{b} = \underline{0}_{81}$

$\underline{\alpha} \geqslant \underline{0}$

That is: $\underset{81\times80}{A'} = \begin{bmatrix} y_1 & y_2 & \cdots & y_{80} \\ & \underline{I}_{80} & & \end{bmatrix}$

~~$A_1 = \underline{y}$~~; ~~$A_2 = \underline{I}_{80}$~~ $\underline{A}$ is 80x

At last, we ~~to~~ calculate

$\hat{\underline{\beta}} = (\underline{G}\underline{\alpha})'\underline{x} = \underline{\alpha}'\underline{G}'\underline{x}$

$\equiv$ ~~$\begin{pmatrix}\beta_1\\ \beta_2\end{pmatrix}$~~ $(\beta_1 \;\; \beta_2)$

and $\beta_0 = \frac{1}{2}[\max(-1)\underline{x}_i^T\hat{\underline{\beta}} + \max \underline{x}_j^T\hat{\underline{\beta}}]$, where

$i = 1, 2, \ldots, 40$

$j = 41, 42, \ldots, 80$

1.c.

1-c is simular to 1-b. The only two differences are;

① $\underline{Y}$ changed; and

② One more constraints $\underline{\alpha} \le C$

still, let $\underset{80\times 80}{\underline{G}}$ be a diagonal matrix such that

$$diag(\underline{G}) = \underline{Y}.$$ ~~an~~

And still,

$\underline{\beta} = \underline{\alpha}$; $\underline{D} = \underline{G}'\underline{x}\underline{x}'\underline{G}$;

$\underline{d} = \underline{1}_N'$

$$\underset{161\times 80}{\underline{A}'} = \begin{bmatrix} t(\underline{Y}) \\ \underline{I}_{80} \\ -\underline{I}_{80} \end{bmatrix}, \quad \underline{b} = \begin{bmatrix} \underline{0}_{81} \\ -\underline{C}_{80} \end{bmatrix}$$

1.d

Following previous sets up,

$\underline{\beta} = \underline{\alpha}$;

$\underset{n\times n}{\underline{G}}$ is such that

$$diag(\underline{G}) = \underline{Y};$$

$\underline{d} = \underline{1}_n$.

Then

$$\hat{\underline{\beta}} = (\underline{G}\underline{\alpha})'\underline{x} = \underline{\alpha}'\underline{G}'\underline{x}$$

and $\beta_0 = mean(\underline{G}\underline{x}\hat{\underline{\beta}})$

$= mean(\underline{G}\underline{x}\alpha'G'\underline{x})$

Question 2

$$f(x) = \sum_{j=0}^{3} \beta_j x^j + \sum_{k=1}^{K} \theta_k (x - \varepsilon_k)_+^3$$

$$= \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \sum_{k=1}^{K \theta_k} (x - \varepsilon_k)_+^3$$

$$= \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + \sum_{k=1}^{K} \theta_k (x - \varepsilon_k)^3$$

when $x < \varepsilon_1$ and $x > \varepsilon_k$, $f(x)$ is linear in ~~k~~ x. That is

when $x < \varepsilon_1$

$f(x) = \beta_0 + \beta_1 x + \beta_2 x^2 + \beta_3 x^3 + 0$ is linear in $x \Rightarrow \beta_2 = \beta_3 = 0$

when $x > \varepsilon_3$:

$$f(x) = \beta_0 + \beta_1 x + \sum_{k=1}^{K} \theta_k (x^3 - 3x^2 \varepsilon_k + 3 x \varepsilon_k^2 - \varepsilon_k^3)$$

is also linear in ~~θ~~ x

$\Rightarrow$ the coefficients of x^3 and x^2 are both 0.

$$\sum_{k=1}^{K} \theta_k = 0 \quad , \quad \sum_{k=1}^{K} \theta_k 3\varepsilon_k = 0 \Rightarrow \sum_{k=1}^{K} \theta_k \varepsilon_k = 0$$

Therefore:

$$f(x) = \beta_0 + \beta_1 x + \sum_{k=1}^{K} \theta_k (x - \varepsilon_k)_+^3 .$$

~~Since~~

$$= \beta_0 + \beta_1 x + \sum_{k=1}^{K-2} \theta_k (x - \varepsilon_k)_+^3$$

$$+ \theta_{K-1} (x - \varepsilon_{K-1})_+^3 + \theta_K (x - \varepsilon_K)_+^3$$

Because $\sum_{k=1}^{K} \theta_k = 0$,

$$\theta_K = -\sum_{k=1}^{K-1} \theta_k = -\sum_{k-1}^{K-2} \theta_k - \theta_{K-1} \quad ①$$

Because $\sum_{k=1}^{K} \theta_k \varepsilon_k = 0$

$$\theta_K \varepsilon_K = -\sum_{k=1}^{K-1} \theta_k \varepsilon_k \quad ②$$

$$= -\sum_{k-1}^{K-2} \theta_k \varepsilon_k - \theta_{K-1} \varepsilon_{K-1}$$

1)
2) PROVER
3)

Therefore:

~~$$f(x) = \beta_0 + \beta_1 x + \sum_{k=1}^{K-2} \theta_k (x - \varepsilon_k)_+^3$$~~

~~$$+ \theta_{K-1} (x - \varepsilon_{K-1})_+^3$$~~

~~$$- \sum_{k=1}^{K-2} \theta_k (x - \varepsilon_k)^3 - \theta_{K-2} (x - \varepsilon_K)_+^3$$~~

It is given that for $k=1,2,..,K-2$

$$d_k(x) = \frac{(x-\varepsilon_k)^3_+ - (x-\varepsilon_K)^3_+}{\varepsilon_K - \varepsilon_k}$$

Therefore, ~~when~~

$$d_k(x) = \begin{cases} 0, \text{ when } x<\varepsilon_k, \\ \cancel{0, \text{ when } x \geq \varepsilon_K,} \\ \cancel{3x^2 (} 3x^2 - 3x(\varepsilon_k+\varepsilon_K) \\ \quad - (\varepsilon_K^2 + \varepsilon_k^2 + 2\varepsilon_K\varepsilon_k), \\ \quad x \geq \varepsilon_K \\ (x-\varepsilon_k)^3, \varepsilon_k \leq x \leq \varepsilon_K \end{cases}$$

$$d_{K-1}(x) = \begin{cases} 0, \text{ when } x<\varepsilon_{K-1}, \\ \vdots \\ \vdots \text{ replacing } k \text{ with } K-1 \end{cases}$$

in $d_k(x)$

Then

$$\sum_{k=1}^{K-2} \alpha_k [d_k(x) - d_{K-1}(x)]$$

$$= \sum_{k=1}^{K-2} \alpha_k \left[\frac{(x-\varepsilon_k)^3_+ - (x-\varepsilon_K)^3_+}{\varepsilon_K - \varepsilon_k} - \frac{(x-\varepsilon_{K-1})^3_+ - (x-\varepsilon_K)^3_+}{\varepsilon_K - \varepsilon_{K-1}} \right]$$

$$= \sum_{k=1}^{K-2} \cancel{\alpha}\cdot\alpha_k (x-\varepsilon_K)^3 \left(\frac{1}{\varepsilon_K-\varepsilon_{K-1}} - \frac{1}{\varepsilon_K-\varepsilon_k}\right) - \cancel{\sum_{k=1}^{K-2}\left[\alpha_k\right.} \frac{(x-\varepsilon_{K-1})^3}{\varepsilon_K-\varepsilon_{K-1}} \sum_{k=1}^{K-2}\alpha_k + \sum_{k=1}^{K-2}\theta_k(x-\varepsilon_k)^3_+$$

$$= \frac{(x-\varepsilon_K)^3_+ - (x-\varepsilon_{K-1})^3_+}{\varepsilon_K - \varepsilon_{K-1}} \sum_{k=1}^{K-2} \alpha_k - (x-\varepsilon_K)^3 \sum_{k=1}^{K-2} \frac{\alpha_k}{\varepsilon_K - \varepsilon_k}$$

$$+ \sum_{k=1}^{K-2} \theta_k (x-\varepsilon_k)^3_+$$

$$= \frac{(x-\varepsilon_K)^3_+ - (x-\varepsilon_{K-1})^3_+}{\varepsilon_K - \varepsilon_{K-1}} (-\theta_{K-1}(\varepsilon_K - \varepsilon_{K-1}))$$

$$-(x-\varepsilon_K)^3_+ \sum_{k=1}^{K-2} \theta_k + \sum_{k=1}^{K-2} \theta_k (x-\varepsilon_k)^3_+$$

$$= -\theta_{K-1}\left[(x-\varepsilon_K)^3_+ - (x-\varepsilon_{K-1})^3_+\right] + (x-\varepsilon_K)^3\left[\theta_K + \theta_{K-1}\right]$$

$$+ \sum_{k=1}^{K-2} \theta_k (x-\varepsilon_k)^3$$

$$= \sum_{k=1}^{K-2} \theta_k (x-\varepsilon_k)^3_+ + \theta_{K-1}(x-\varepsilon_{K-1})^3_+ + \theta_K (x-\varepsilon_K)^3_+$$

$$= \sum_{k=1}^{K} \theta_k (x-\varepsilon_k)^3_+$$

by ~~$\sum_{k=1}^{K-2}$~~ ①, ②, ~~③~~