Page No.
Date :

Assignment-2

Ishit Bajpai
2020880

Q1

(a) Min $x_1^2 + 2x_2^2 + 3x_3^2 + (x_1 - x_2 + x_3 - 1)^2 + (-x_1 - 4x_2 + 2)^2$

$(1\cdot x_1 + 0x_2 + 0x_3)^2 + (0\cdot x_1 + \sqrt{2}x_2 + 0\cdot x_3)^2 + (0\cdot x_1 + 0\cdot x_2 + \sqrt{3}x_3)^2$
$+ (x_1 - x_2 + x_3 - 1)^2 + (-x_1 - 4x_2 + 0\cdot x_3 + 2)^2 = 0$

equaling all the square terms to 0 so as to minimize the expression.

$x_1 = 0,\ x_2\sqrt{2} = 0,\ \sqrt{3}x_3 = 0,\ x_1 - x_2 + x_3 = 1,\ -x_1 - 4x_2 = -2$

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & \sqrt{2} & 0 \\ 0 & 0 & \sqrt{3} \\ 1 & -1 & +1 \\ -1 & -4 & 0 \end{bmatrix} \qquad b = \begin{bmatrix} 0 \\ 0 \\ 0 \\ +1 \\ -2 \end{bmatrix}$$

(b) Minimize $(-6x_2 + 4)^2 + (-4x_1 + 3x_2 - 1)^2 + (x_1 + 8x_2 - 3)^2$

$0\cdot x_1 - 6x_2 + 4 = 0$
$-4x_1 + 3x_2 - 1 = 0$
$x_1 + 8x_2 - 3 = 0$

$$A = \begin{bmatrix} 0 & -6 \\ -4 & 3 \\ 1 & 8 \end{bmatrix} \qquad b = \begin{bmatrix} -4 \\ 1 \\ 3 \end{bmatrix}$$

(c) Minimize $2(-6x_2 + 4)^2 + 3(-4x_1 + 3x_2 - 1)^2 + 4(x_1 + 8x_2 - 3)^2$

Taking outside terms inside brackets

$0\cdot x_1 - 6\sqrt{2}x_2 + 4\sqrt{2} = 0$
$-4\sqrt{3}x_1 + 3\sqrt{3}x_2 - \sqrt{3} = 0$
$2x_1 + 16x_2 - 6 = 0$

Page No.
Date :

$$A = \begin{bmatrix} 0 & -6\sqrt{2} \\ -4\sqrt{3} & 3\sqrt{3} \\ 2 & 16 \end{bmatrix} \qquad b = \begin{bmatrix} -4\sqrt{2} \\ \sqrt{3} \\ -6 \end{bmatrix}$$

(d) $\min \; x^T x + \|Bx - d\|_2^2$ $\quad B \in R^{p \times n}$

$x^T x = \|x\|^2$ $\quad d \in R^p$

- $\min\left(\|x\|^2 + \|Bx - d\|^2\right)$

- Since $\|x\| \geq 0$ & $\|x\| = 0$ iff $x = 0$ $(0 \in R^n)$
  $Ix = x = 0$

Equating $\quad x = 0$

Similarly, $\underline{\underline{Bx - d = 0}}$ $\quad (0 \in R^p)$
$\Rightarrow Bx = d$

$$A = \begin{bmatrix} I \\ B \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} 0 \\ d \end{bmatrix} = b$$

(e) $\min\left(x^T D x + \|Bx - d\|_2^2\right)$

$D \in R^{n \times m}$

Page No.
Date:

$$\begin{bmatrix} x_1 & \cdots & x_n \end{bmatrix} \begin{bmatrix} d_1 & & \\ & \ddots & \\ & & d_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$= \begin{bmatrix} x_1 d_1 & x_2 d_2 & \cdots & x_n d_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$$

$$= x_1^2 d_1 + x_2^2 d_2 \cdots x_n^2 d_n$$

(Since it is a Linear least square problem)

$$= (x_1\sqrt{d_1})^2 + (x_2\sqrt{d_2})^2 \cdots + (x_n\sqrt{d_n})^2$$

$$x_i\sqrt{d_i} = 0 \quad i = 1, 2 - n$$

$$\overbrace{\begin{bmatrix} \sqrt{d_1} & & & \\ & \sqrt{d_2} & & \\ & & \ddots & \\ & & & \sqrt{d_n} \end{bmatrix}}^{D_1} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} 0 \end{bmatrix}$$

Minimizing for $\|Bx-d\|_2^2$, $Bx - d = 0 \Rightarrow Bx = d$

$$\begin{bmatrix} B \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} d \end{bmatrix}$$

Combining

$$\begin{bmatrix} D_1 \\ B \end{bmatrix} \begin{bmatrix} x \end{bmatrix} = \begin{bmatrix} 0 \\ d \end{bmatrix} \quad 0 \in \mathbb{R}^n$$

Page No.

Date :

$D_1 \in R^{n \times n}$

$B \in R^{p \times n}$

$$A = \begin{bmatrix} D_1 \\ B \end{bmatrix} \qquad A \in R^{(p+n) \times n}$$

$$b = \begin{bmatrix} 0 \\ d \end{bmatrix}$$

(Q2) Given:

$$\overset{A}{\begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}} \overset{x}{\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}} = \overset{b}{\begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix}}$$

Using necessary condition for existence of minimum, normal equation were obtained

i.e $A^T A x = A^T b$

$$A^T A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 2 \end{bmatrix}$$

$$A^T b = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 2 \\ 3 \end{bmatrix}$$

$$x = (A^T A)^{-1} (A^T b)$$

$$= \begin{bmatrix} 2 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Page No.

Date

Verification:

$$A = \begin{bmatrix} 1 & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$$

for sufficiency - condition for minima is also that ~~Rank~~ A is full rank i.e all columns are linearly independent.

$$c_1 a_1 + c_2 a_2 = 0$$

$$\begin{bmatrix} c_1 \\ 0 \\ 0 \end{bmatrix} + \begin{bmatrix} c_2 \\ c_2 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

$$c_1 + c_2 = 0$$

$$c_2 = 0$$

$$\Rightarrow c_1 = 0$$

$c_1 a_1 + c_2 a_2 = 0$ iff $c_1 = 0$ and $c_2 = 0$

$\Rightarrow$ a_1, a_2 are linearly independent

$\Rightarrow$ Rank$(A) = 2$ $\Rightarrow$ A is full rank

Hence proved.

Q5

a) $y_i \approx \dfrac{e^{\alpha t_i+\beta}}{1+e^{\alpha t_i+\beta}}$

$y_i \approx \dfrac{1}{1+e^{-(\alpha t_i+\beta)}}$

$1+\left(e^{-(\alpha t_i+\beta)}\right) \approx \left(\dfrac{1}{y_i}\right)$

$e^{-(\alpha t_i+\beta)} = \dfrac{1-y_i}{y_i}$

taking log

$-(\alpha t_i+\beta) = \ln\left(\dfrac{1-y_i}{y_i}\right)$

$(\alpha t_i+\beta) = -\ln\left(\dfrac{1-y_i}{y_i}\right) = \ln\left(\dfrac{y_i}{1-y_i}\right)$

$\Rightarrow \alpha t_i+\beta = \ln\left(\dfrac{y_i}{1-y_i}\right) \quad i=1,\ldots 50$

$\alpha t_1+\beta = \ln\left(\dfrac{y_1}{1-y_1}\right) = C_1$ (Constant)

$\alpha t_2+\beta = \ln\left(y_2/1-y_2\right) = C_2$

$\vdots$

$\alpha t_{50}+\beta = \ln\left(y_{50}/1-y_{50}\right) = C_{50}$

$$c_i = \ln\left(y_i/1-y_i\right)$$

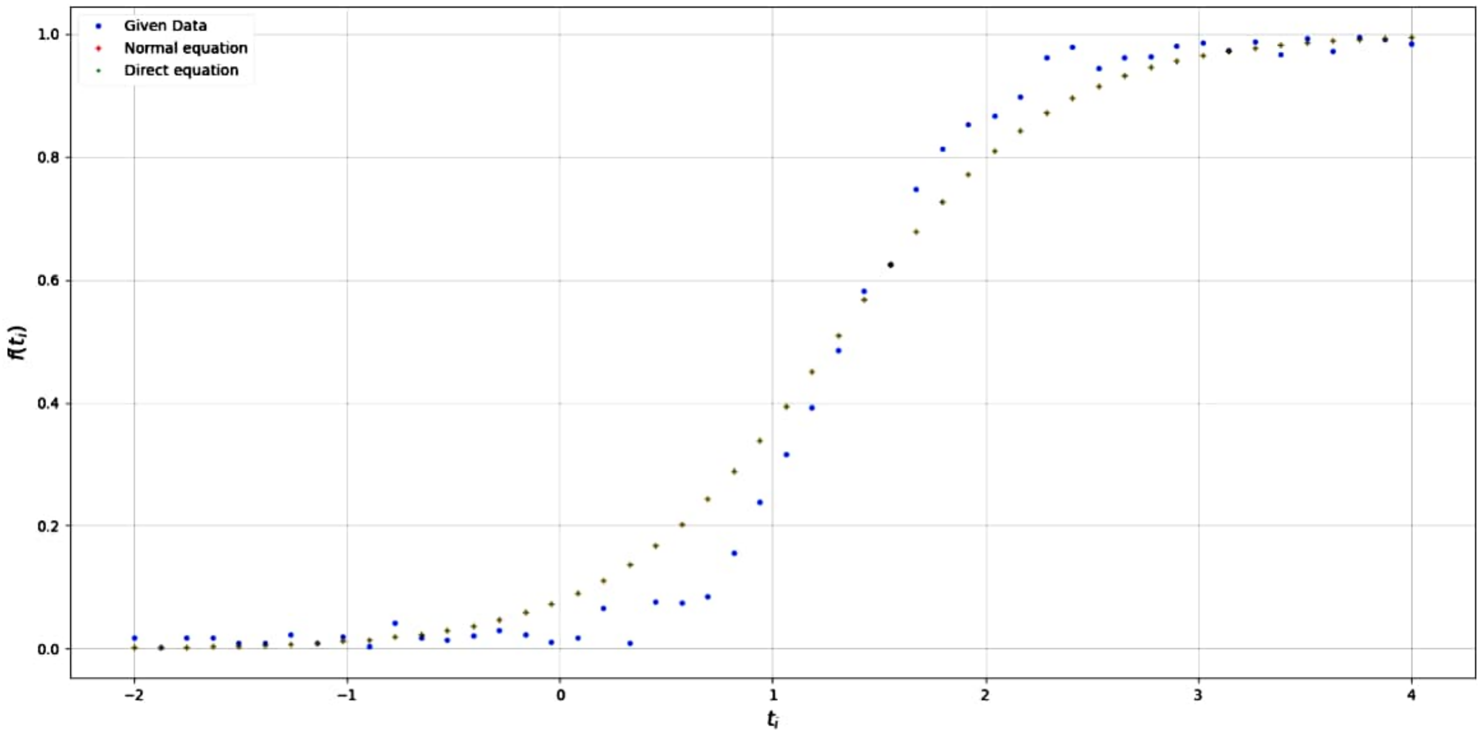
$$\cancel{\begin{bmatrix} 1 & t_1 \\ \vdots & t_2 \\ \vdots & \vdots \\ 1 & t_{50} \end{bmatrix} \begin{bmatrix} B \\ \alpha \end{bmatrix} = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_{50} \end{bmatrix}}$$

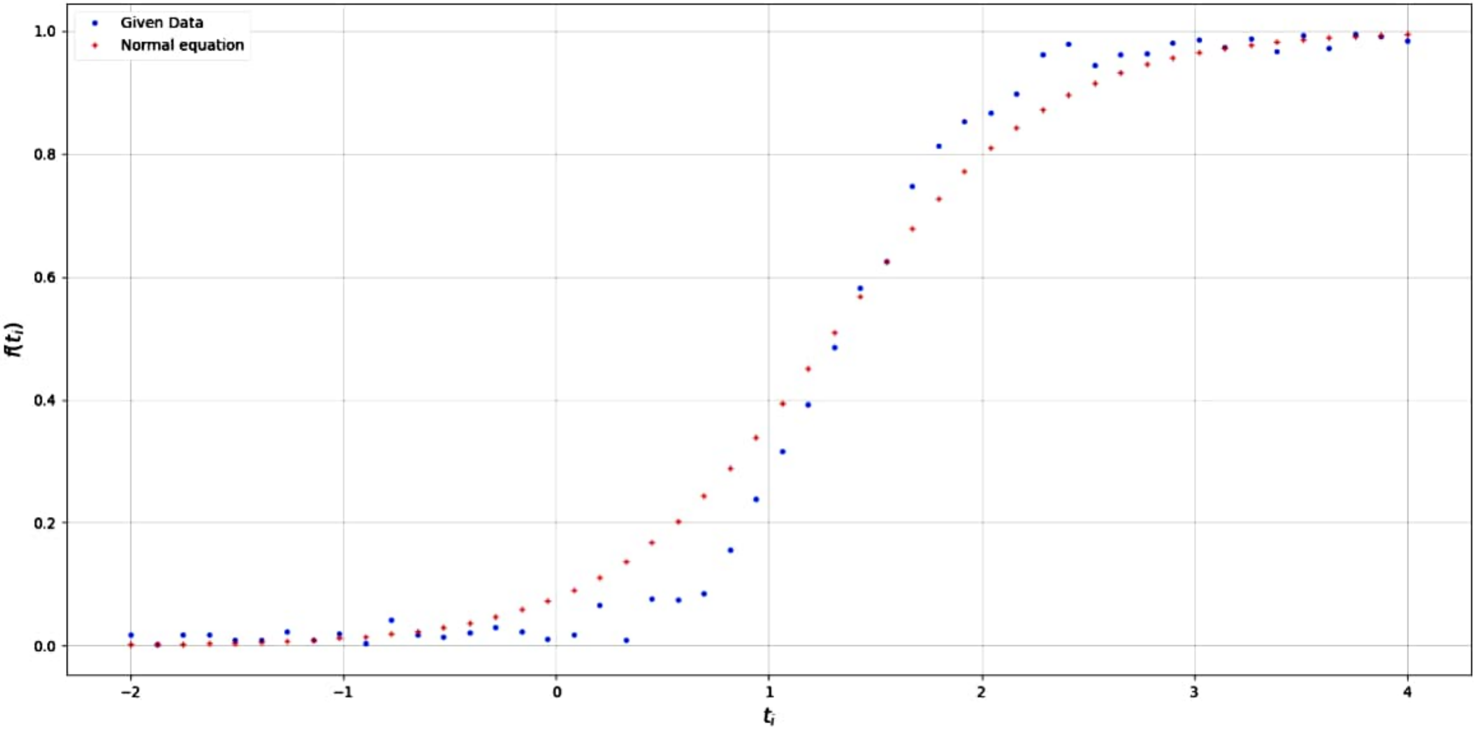
$$\underset{50\times 2}{\begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_{50} \end{bmatrix}} \underset{2\times 1}{\begin{bmatrix} B \\ \alpha \end{bmatrix}} = \underset{50\times 1}{\begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_{50} \end{bmatrix}}$$

$$A = \begin{bmatrix} 1 & t_1 \\ 1 & t_2 \\ \vdots & \vdots \\ 1 & t_{50} \end{bmatrix} \qquad b = \begin{bmatrix} C_1 \\ C_2 \\ \vdots \\ C_{50} \end{bmatrix}$$

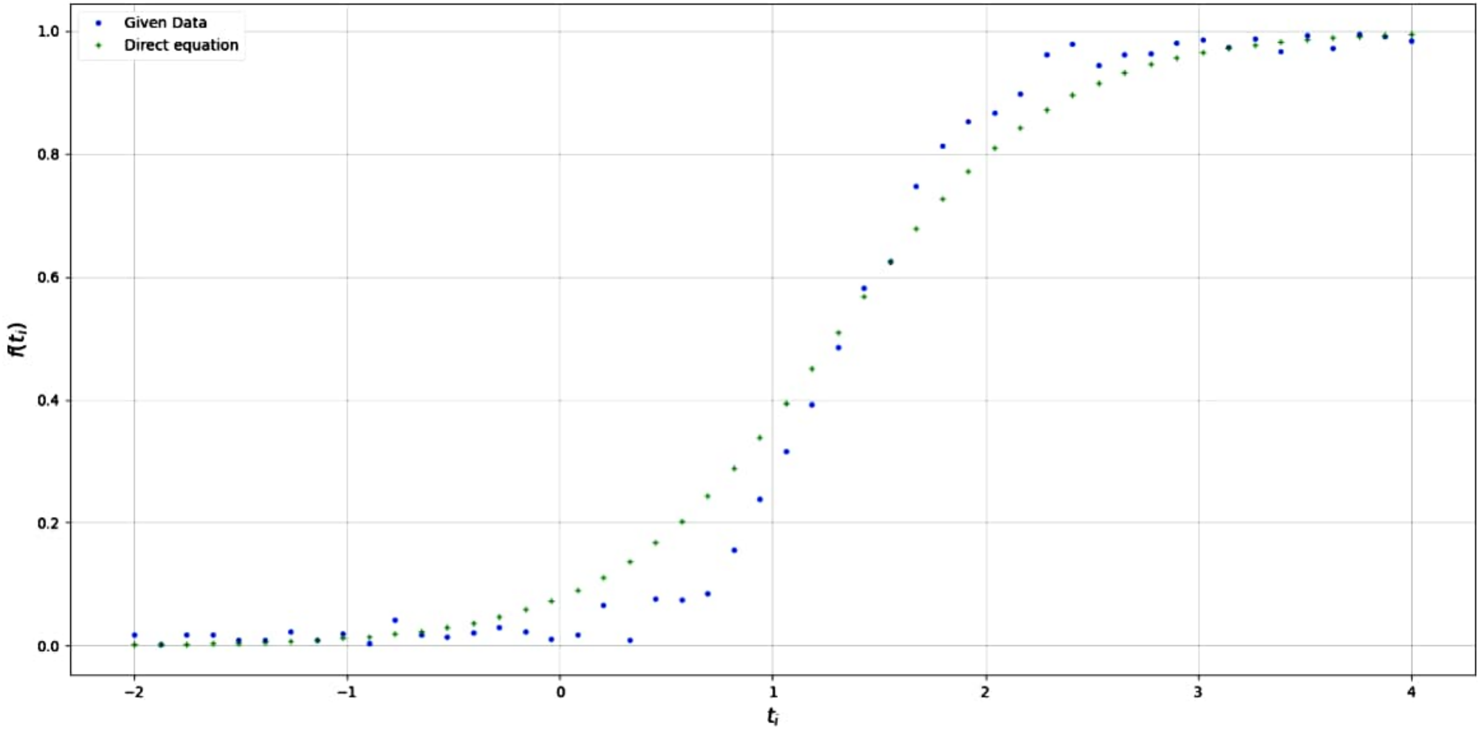
where $C_i = \ln\left(y_i/1-y_i\right)$

```
PS C:\Users\admin\SC_A2> python -u "c:\Users\admin\SC_A2\Q5.py"
Solution from normal equation
x =  [[-2.46632514]
 [ 1.91883939]]
Euclidean norm of residual =  6.980329177231976
------------------------------
Solution from direct equation
x =
 [[-2.46632514]
 [ 1.91883939]]
Euclidean norm of residual calculated from lstsq function =  6.980329177231
976
Euclidean norm of residual calculated from norm function  =  6.98032917723197
------------------------------
Euclidean residual is less in case of normal equations by(using norm function) : 8.881784197001252e-16
```


Given Data
Normal equation
Direct equation
1.0
0.8
0.6
0.4
0.2
0.0
f(t_i)
−2
−1
0
1
2
3
4
t_i



Given Data
Normal equation
1.0
0.8
0.6
0.4
0.2
0.0
−2
−1
0
1
2
3
4
$f(t_i)$
t_i



Given Data
Direct equation
1.0
0.8
0.6
0.4
0.2
0.0
$f(t_i)$
−2
−1
0
1
2
3
4
t_i

Page No.
Date :

Q6 $A = \begin{bmatrix} 1 & 1 \\ 10^{-k} & 0 \\ 0 & 10^{-k} \end{bmatrix}$ $b = \begin{bmatrix} -10^{-k} \\ 1 + 10^{-k} \\ 1 - 10^{-k} \end{bmatrix}$

(a) Writing normal equation.

$A^T A x = A^T b$.

- $A^T A = \begin{bmatrix} 1 & 10^{-k} & 0 \\ 1 & 0 & 10^{-k} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 10^{-k} & 0 \\ 0 & 10^{-k} \end{bmatrix} = \begin{bmatrix} 1+10^{-2k} & 1 \\ 1 & 1+10^{-2k} \end{bmatrix}$

- $A^T b = \begin{bmatrix} 1 & 10^{-k} & 0 \\ 1 & 0 & 10^{-k} \end{bmatrix} \begin{bmatrix} -10^{-k} \\ 1+10^{-k} \\ 1-10^{-k} \end{bmatrix} = \begin{bmatrix} -10^{-k}+10^{-k}+10^{-2k} \\ -10^{-k}+10^{-k}-10^{-2k} \end{bmatrix}$

$$= \begin{bmatrix} 10^{-2k} \\ -10^{-2k} \end{bmatrix} = 10^{-2k} \begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

DATE
PAGE

$$(A^TA)^{-1} = \frac{1}{(1+10^{-2k})^2-1}\begin{bmatrix} 1+10^{-2k} & -1 \\ -1 & 1+10^{-2k}\end{bmatrix}$$

$$= \frac{1}{(10^{-2k})(2+10^{-2k})}\begin{bmatrix} 1+10^{-2k} & -1 \\ -1 & 1+10^{-2k}\end{bmatrix}$$

In Double precision FP ; $1+10^{-2k} = 1$, $k \geq 8$
i.e the above Matrix is singular for $k \geq 8$. (as $(1+10^{-2k}) = 1$) (10^{-16} is smaller than ε_M)

$$x = (A^TA)^{-1}A^Tb$$

$$= \frac{1}{(10^{-2k})(2+10^{-2k})}\begin{bmatrix} 1+10^{-2k} & -1 \\ -1 & 1+10^{-2k}\end{bmatrix} \times (10^{-2k})\begin{bmatrix} 1 \\ -1\end{bmatrix}$$

$$= \frac{1}{(2+10^{-2k})} \begin{bmatrix} 2+10^{-2k} \\ -2-10^{-2k}\end{bmatrix} = \begin{bmatrix} 1 \\ -1\end{bmatrix}$$

c) ~~Since~~ Using normal equation for $k \geq 8$, would result in A^TA being a Singular matrix, therefore we ~~get no solutions~~ cannot compute soln for $k \geq 8$.
for $k = 6$ and 7, 7 is more closer to the analytical solution than $k = 6$.

DATE
PAGE

for QR decomposition, for $k = 6, 7$ we get Solution which is exactly equal to Analytical solution. (less than (c) part where some error exists)

However as $k\uparrow$, error in Computation increases and our solution deviates from the Analytical soln. by larger Ammount. (~~error~~ is

for $k = 15$, $X = \begin{bmatrix} 0.8635 \\ -0.8635 \end{bmatrix}$

~~ie here error is the di~~

Here, error $= \begin{bmatrix} |1 - X\text{calculated}[0][0]| \\ |-1 - X\text{calculated}[1][0]| \end{bmatrix}$

Q7) ~~Ass~~ Observation : As value of λ increases curve becomes smoother and x (estimated signal) deviates by larger ammount from x_{noisy}.

In the problem, we are trying to ~~min~~

$$\min \; \|x - x_{noisy}\|_2^2 + \lambda \sum_{i=1}^{999} (x_{i+1} - x_i)^2$$

$\min \|x - x_{noisy}\|_2^2$ is used to minimize deviation between x and x_{noisy}.

$\min (x_{i+1} - x_i)^2$ is used to bring successive points closer to each other, that is making the curve smooth.

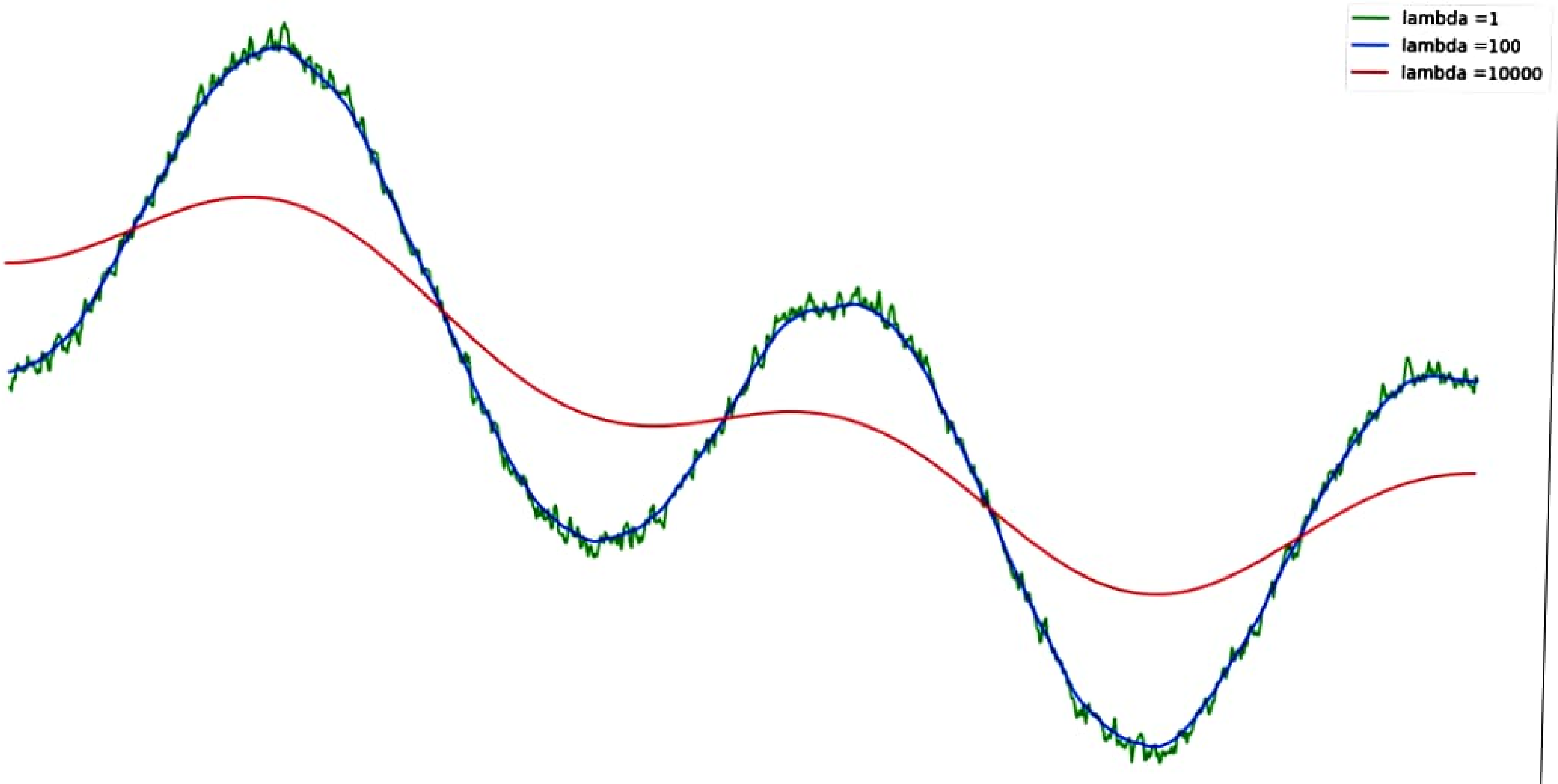
$\lambda \leftarrow$ gives weight to the term $\sum_{i=1}^{999} (x_{i+1} - x_i)^2$, i.e as $\lambda\uparrow$ we give higher priority to make curve smooth and deviation b/w x and $x_{noisy}\uparrow$.

for lower values of λ, not (less) much priority is given to $(x_{i+1} - x_i)^2$ term, as a result deviation between x and x_{noisy} is less and curve is not smooth, i.e there are rapid changes in ~~x~~ ~~estimated~~ x.

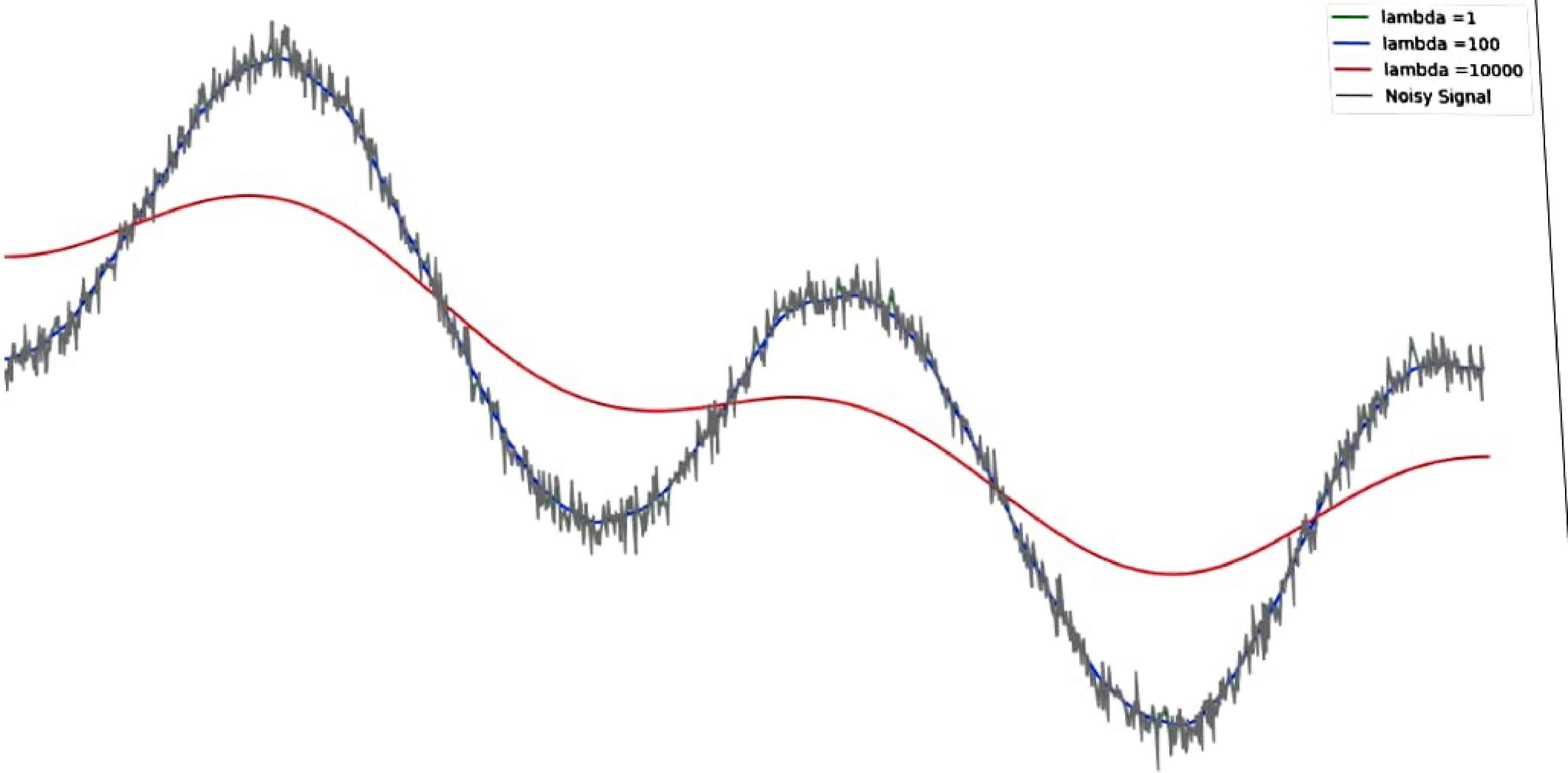
- $\lambda = 1$, There are rapid changes (not smooth curve) but ~~xxx~~ but deviation b/w X and Xnoisy is very small. (They seems to be overlapping)

$\lambda = 100$, There are less rapid changes and deviation between X and Xnoisy is larger as compared to $\lambda = 1$.

For $\lambda = 10000$, Curve is ~~xxx~~ very smooth (no rapid changes), However the deviation between X and Xnoisy is largest in this case.



lambda =1
lambda =100
lambda =10000


lambda =1
lambda =100
lambda =10000
Noisy Signal

Page No.
Date :

(Q4) Given: $\hat{x}$ is the sol^n of LLS : $\min \|b - Ax\|_2^2$
, $b \notin \operatorname{Im} A$.

$$[A \ b] = QR = [q_1 \ \cdots \ q_{n+1}] \begin{bmatrix} R_{11} & R_{12} & \cdots & R_{1,n+1} \\ & R_{22} & \cdots & R_{2,n+1} \\ & & \ddots & \vdots \\ & & & R_{n+1,n+1} \end{bmatrix}$$

$R_{ij} = \langle a_j, q_i \rangle$

(a) $\|b\|_2$

$$q_{n+1} = \frac{b - \langle b, q_1\rangle q_1 \cdots \langle b, q_n\rangle q_n}{\|b - \langle b,q_1\rangle q_1 - \cdots - \langle b, q_n\rangle q_n\|}$$

$\cancel{\Rightarrow} \ \cancel{b = q_{n+1} R_{n+1} + q_n R_{n+1} + \cdots + q_1 R}$

$b = R_{n+1,n+1} q_{n+1} + R_{n,n+1} q_n \ \cdots \ R_{n+1,1} q_1$

$$b = \sum_{i=1}^{n+1} (R_{i,n+1}) q_i$$

$$\|b\|_2 = \sqrt{\sum_{i=1}^{n+1} \|q_i R_{i,n+1}\|}$$

(Since q_i $i = 1 \cdots n$ are orthogonal to each other)

$$\|b\|_2 = \sqrt{\sum_{i=1}^{n+1} \|q_i\| \|R_{i,n+1}\|} = \sqrt{\sum_{i=1}^{n+1} \|R_{i,n+1}\|}$$

$$= \sqrt{\sum_{i=1}^{n+1} \|R_{i,n+1}\|}$$

$$= \sqrt{\sum_{i=1}^{n+1} R_{i,n+1}^2}$$

Page No.

Date :

(b) $\|A\hat{x}\|_2^2$

$$A\hat{x} = \sum_{i=1}^{n} q_i R_{i,n+1}$$

$$\|A\hat{x}\|_2 = \left(\sum_{i=1}^{n} |q_i R_{i,n+1}|\right)^{1/2}$$

$$= \left(\sum_{i=1}^{n} \|q_i\| \|R_{i,n+1}\|\right)^{1/2} = \left(\sum_{i=1}^{n} \|R_{i,n+1}\|\right)^{1/2}$$

$$= \left(\sum_{i=1}^{n} R_{i,n+1}^2\right)^{1/2}$$

(c) ~~Since Regardless~~

$$\|b - Ax\| = \cancel{\|b\|_2 - \|Ax\|_2} \; |R_{n+1,n+1}|$$

Using (a) & (b)

$$= \cancel{R_{n+1,n+1}}$$

$$\|b - Ax\| = |R_{n+1,n+1}|$$