Adnane
Ezouhri

01351331

Machine learning
HW3

Problem 1

a) $T(t) = \theta_0 + \theta_1 t$

$$X = \begin{bmatrix} 1 & 1{,}850 \\ 1 & 1{,}900 \\ 1 & 1{,}950 \\ 1 & 2{,}021 \end{bmatrix} ; \quad y = \begin{bmatrix} 75 \\ 75 \\ 77 \\ 90 \end{bmatrix}$$

$$\theta = \begin{bmatrix} \theta_0 \\ \theta_1 \end{bmatrix}$$

b)

$$\widehat{\theta_{LSF}} = \left(X^T X\right)^{-1} X^T y = \begin{bmatrix} 4 & 7{,}721 \\ 7{,}721 & 14{,}919 \end{bmatrix}^{-1} X^T y$$

$$\widehat{\theta}_{LSF} = \begin{bmatrix} 233{,}4 & -120{,}8 \\ -120{,}8 & 62{,}58 \end{bmatrix} \begin{bmatrix} 317 \\ 613{,}3 \end{bmatrix}$$

$$\widehat{\theta}_{LSF} = \begin{bmatrix} -89{,}94 \\ 87{,}65 \end{bmatrix}$$

c) $\epsilon(\theta) = \|W(X\theta - y)\|^2 = (W(X\theta - y))^T (W(X\theta - y))$

$\epsilon(\theta) = \theta^T X^T W^T W X\theta - \theta^T X^T W^T W y - y^T W^T W X\theta + y^T W^T W y$

Since W is a diagonal matrix: $W^T = W$. Hence:

$\epsilon(\theta) = \theta^T X^T W^2 X\theta - \theta^T X^T W^2 y - y^T W^2 X\theta + y^T W^2 y$

We then set $Q = X^T W^2 X$ and $C = X^T W^2 y$ to use: • $\nabla_\theta(\theta^T Q \theta) = 2Q\theta$

• $\nabla_\theta(\theta^T c) = c$

$\epsilon(\theta) = \theta^T Q\theta - \theta^T c - y^T W^2 X\theta + y^T W^2 y$

$\frac{\partial(\epsilon(\theta))}{\partial\theta} = 2X^T W^2 X\theta_{WLS} - 2X^T W^2 y = 0 \Leftrightarrow \boxed{\theta_{WLS} = (X^T W^2 X)^{-1} X^T W^2 y}$

d) $\hat{\theta}_{WLS} = (X^T W^2 X)^{-1} X^T W^2 y$

$\hat{\theta}_{WLS} = \begin{bmatrix} 101,1 & 204,2 \\ 204,2 & 412,6 \end{bmatrix}^{-1} \times \begin{bmatrix} 9085 \\ 1,835 \times 10^4 \end{bmatrix}$

$\boxed{\hat{\theta}_{WLS} = \begin{bmatrix} -232,8 \\ 159,7 \end{bmatrix}}$

e) Temperature per year using LSF:

$\hat{y}_{LSF} = X \times \hat{\theta}_{LSF} = \begin{bmatrix} 1 & 1,850 \\ 1 & 1,900 \\ 1 & 1,950 \\ 1 & 2,021 \end{bmatrix} \times \begin{bmatrix} -89,94 \\ 87,65 \end{bmatrix}$

$\boxed{\hat{y}_{LSF} = \begin{bmatrix} 72,21 \\ 76,6 \\ 80,98 \\ 87,2 \end{bmatrix}}$

Now using the WLS:

$\hat{y}_{WLS} = X \times \hat{\theta}_{WLS} = \begin{bmatrix} 1 & 1,850 \\ 1 & 1,900 \\ 1 & 1,950 \\ 1 & 2,021 \end{bmatrix} \times \begin{bmatrix} -232,8 \\ 159,7 \end{bmatrix}$

$\boxed{\hat{y}_{WLS} = \begin{bmatrix} 62,65 \\ 70,63 \\ 78,62 \\ 89,95 \end{bmatrix}}$

f) To find temperature in 2050, set X such that:

$$X = \begin{bmatrix} 1 & 2{,}050 \end{bmatrix}$$

Using LSF:

$$\hat{y}_{2050} = X \times \hat{\theta}_{LSF} = \begin{bmatrix} 1 & 2{,}050 \end{bmatrix} \times \begin{bmatrix} -89{,}94 \\ 87{,}65 \end{bmatrix} \boxed{= 89{,}74}$$

Using WLS:

$$\hat{y}_{2050} = X \times \hat{\theta}_{WLS} = \begin{bmatrix} 1 & 2{,}050 \end{bmatrix} \times \begin{bmatrix} -232{,}8 \\ 159{,}7 \end{bmatrix} \boxed{= 94{,}59}$$

g)

$$T(t) = \theta_0 + \theta_1 t + \theta_2 t^2$$

$$X = \begin{bmatrix} 1 & 1{,}850 & 1{,}850^2 \\ 1 & 1{,}900 & 1{,}900^2 \\ 1 & 1{,}950 & 1{,}950^2 \\ 1 & 2021 & 2021^2 \end{bmatrix} \qquad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix}$$

$$\hat{\theta}_{LSF} = (X^T X)^{-1} X^T y$$

$$\hat{\theta}_{LSF} = \begin{bmatrix} 4 & 7{,}721 & 14{,}92 \\ 7{,}721 & 14{,}92 & 28{,}86 \\ 14{,}92 & 28{,}86 & 55{,}89 \end{bmatrix}^{-1} \times \begin{bmatrix} 317 \\ 613{,}3 \\ 1188 \end{bmatrix}$$

$$\boxed{\hat{\theta}_{LSF} = \begin{bmatrix} 3166 \\ -3278 \\ 868{,}8 \end{bmatrix}}$$

$$\hat{\theta}_{WLS} = (X^T W^2 X)^{-1} X^T W^2 y$$

$$\hat{\theta}_{WLS} = \begin{bmatrix} 101{,}1 & 204{,}2 & 412{,}6 \\ 204{,}2 & 412{,}6 & 833{,}6 \\ 412{,}6 & 833{,}6 & 1684 \end{bmatrix}^{-1} \times \begin{bmatrix} 9085 \\ 1{,}835 \times 10^4 \\ 3{,}708 \times 10^4 \end{bmatrix}$$

$$\boxed{\hat{\theta}_{WLS} = \begin{bmatrix} 3590 \\ -3715 \\ 981{,}5 \end{bmatrix}}$$

$$y_{2050} = X \times \hat{\theta}_{LSF} = \begin{bmatrix} 1 & 2{,}050 & 2{,}050^2 \end{bmatrix} \times \begin{bmatrix} 3166 \\ -3278 \\ 868{,}8 \end{bmatrix}$$

$$\boxed{y_{2050} = 97{,}23}$$

Problem 3

a)

$$y(t) = \theta_0 + \theta_1 \cos(2\pi t) + \theta_2 \sin(2\pi t)$$

$$X = \begin{bmatrix} 1 & \cos(2\pi \times 0{,}115) & \sin(2\pi \times 0{,}115) \\ 1 & \cos(2\pi \times 0{,}116) & \sin(2\pi \times 0{,}116) \\ 1 & \cos(2\pi \times 0{,}625) & \sin(2\pi \times 0{,}625) \end{bmatrix}$$

$$y = \begin{bmatrix} -1 \\ -1 \\ 1 \end{bmatrix} \qquad \theta = \begin{bmatrix} \theta_0 \\ \theta_1 \\ \theta_2 \end{bmatrix}$$

$$\theta_{LS} = (X^T X)^{-1} X^T y$$

$$\theta_{LS} = \begin{bmatrix} 3 & 2{,}997 & 0{,}0938 \\ 2{,}997 & 2{,}995 & 0{,}0937 \\ 0{,}0938 & 0{,}0937 & 0{,}005 \end{bmatrix}^{-1} \times \begin{bmatrix} -1 \\ -1{,}002 \\ 0{,}0432 \end{bmatrix}$$

$$\boxed{\theta_{LS} = \begin{bmatrix} 1322 \\ -1323 \\ -17{,}92 \end{bmatrix}} \quad y_{LS} = X \propto \theta_{LS} = \begin{bmatrix} -1{,}011 \\ -1{,}011 \\ 0{,}9892 \end{bmatrix}$$

$$\|y_{LS} - y\|^2 = (-1{,}011+1)^2 + (-1{,}011+1)^2 + (0{,}9892-1)^2$$

$$\boxed{\|y_{LS} - y\|^2 = 3{,}59 \times 10^{-4} \cong 0}$$

b) Ridge Regression:

$$\theta_\lambda = (X^T X + \lambda I)^{-1} \propto X^T y$$

$$\theta_\lambda = \begin{bmatrix} 3{,}1 & 2{,}997 & 0{,}0938 \\ 2{,}997 & 3{,}095 & 0{,}0937 \\ 0{,}0938 & 0{,}0937 & 0{,}105 \end{bmatrix}^{-1} \times \begin{bmatrix} -1 \\ -1{,}002 \\ 0{,}0432 \end{bmatrix}$$

$$\boxed{\theta_\lambda = \begin{bmatrix} -0{,}1606 \\ -0{,}1901 \\ 0{,}724 \end{bmatrix}}$$