DS4400 HW2

Xin Guan

1. **Linear Regression:** Consider the modified linear regression problem

$$\hat{\theta} = \arg\min_{\theta} \sum_{i=1}^{N} (\theta^T \phi(x_i) - y_i)^2 + \lambda \|\theta - \mathbf{a}\|_2^2$$

where a is a known and given vector of the same dimension as that of θ. Derive the closed-form solution. Provide all steps of the derivation.

Solution:

$f(\theta) = \sum_{i=1}^{N} (\theta^T \phi(x_i) - y_i)^2 + \lambda \|\theta - \mathbf{a}\|_2^2$

$\frac{\partial f(\theta)}{\partial \theta} = \frac{\partial \sum_{i=1}^{N} (\theta^T \phi(x_i) - y_i)^2}{\partial \theta} + \frac{\partial \lambda \|\theta - a\|_2^2}{\partial \theta}$

$= \sum_{i=1}^{N} [2(\theta^T \phi(x_i) - y_i) \frac{\partial (\theta^T \phi(x_i) - y_i)}{\partial \theta}] + \lambda \frac{\partial \|\theta - a\|_2^2}{\partial \theta}$

$= \sum_{i=1}^{N} [2(\theta^T \phi(x_i) - y_i) \phi(x_i)] + \lambda \frac{\partial (\theta - a)^2}{\partial \theta}$

$\lambda \frac{\partial (\theta - a)^2}{\partial \theta} = \lambda \frac{\partial (\theta^2 - 2\theta a + a^2)}{\partial \theta} = \lambda (2\theta - 2a)$

Therefore, $\frac{\partial f(\theta)}{\partial \theta} = \sum_{i=1}^{N} [2(\theta^T \phi(x_i) - y_i) \phi(x_i)] + 2\lambda(\theta - a)$

Write all data $\phi(x_1), \phi(x_2) \ldots \phi(x_N)$ as a matrix:

$\Phi = \begin{bmatrix} \phi(x_1)^T \\ \phi(x_2)^T \\ \cdots \\ \phi(x_N)^T \end{bmatrix}$ the dimension is $N \times d$

$Y = \begin{bmatrix} y_1 \\ y_2 \\ \cdots \\ y_n \end{bmatrix}$ the dimension is $N \times d$

Then $\frac{\partial f(\theta)}{\partial \theta} = 2(\Phi^T \Phi \theta - \Phi^T Y) - 2\lambda(\theta - a)$

Let $\frac{\partial f(\theta)}{\partial \theta} = 0$

$$\Phi^T \Phi \theta - \Phi^T Y = \lambda(\theta - a)$$

$$\Phi^T \Phi \theta - \Phi^T Y = \lambda I_d \theta - \lambda I_d a$$

$$\Phi^T \Phi \theta - \lambda I_d \theta = \Phi^T Y - \lambda I_d a$$

$$(\Phi^T \Phi - \lambda I_d)\theta = \Phi^T Y - \lambda I_d a$$

$$\theta = (\Phi^T \Phi - \lambda I_d)^{-1} (\Phi^T Y - \lambda I_d a)$$

Therefore,