4

(1) $f(x) = \beta_0 + \beta_1 \sin(x) + \beta_2 \cos(x)$

$$X = \begin{bmatrix} 1 & \sin(x_1) & \cos(x_1) \\ 1 & \sin(x_2) & \cos(x_2) \\ & \vdots & \\ 1 & \sin(x_n) & \cos(x_n) \end{bmatrix}$$

$\sin(x)$ and $\cos(x)$ are linearly independent. Therefore, we have design matrix with linearly independent columns. Hence, we can use least-squares method.

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix} \qquad Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}$$

$$\hat{y} = X\beta$$

$$\beta = (X^T X)^{-1} X^T Y$$

(2)

$$g(x) = \beta_0 + \sin(\beta_1 x) + \cos(\beta_2 x)$$

$$\beta = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \beta_2 \end{bmatrix}, \quad RSS = \sum_{i=1}^{n} \left(g(x^{(i)}) - y^{(i)}\right)^2$$

$$\beta^{(t+1)} = \beta^{(t)} - \eta \left(\frac{\partial (RSS)}{\partial \beta} \right)_{(t)}$$

learning rate

$$\frac{\partial (RSS)}{\partial \beta} = \begin{bmatrix} \dfrac{\partial (RSS)}{\partial \beta_0} \\ \dfrac{\partial (RSS)}{\partial \beta_1} \\ \dfrac{\partial (RSS)}{\partial \beta_2} \end{bmatrix}$$

$$\frac{\partial}{\partial \beta_j}(RSS) = \sum_{i=1}^{n} 2\left(g(x^{(i)}) - y^{(i)}\right) \times \frac{\partial}{\partial \beta_j}\left(g(x^{(i)})\right)$$

$$\frac{\partial}{\partial \beta_0}\left(g(x^{(i)})\right) = 1$$

$$\frac{\partial}{\partial \beta_1}\left(g(x^{(i)})\right) = x^{(i)} \cos\left(\beta_1 x^{(i)}\right)$$

$$\frac{\partial}{\partial \beta_2}\left(g(x^{(i)})\right) = -x^{(i)} \sin\left(\beta_2 x^{(i)}\right)$$

Substitute the 3 terms in above equation to get derivatives with β_j $\forall\ j = 0, 1, 2$