Homework #3

Monday, March 1, 2021 10:49 PM

Problem 3.1

a) $$C = X^T X + Q$$

$$C_i = \bar{X}^T \bar{X} - x_i x_i^T + Q$$

$$C_i = (\bar{X}^T \bar{X} + Q) - x_i \; x_i^T$$

$$C_i = C - x_i \; x_i^T$$

$$d = X^T y$$

$$d_i = X^T y - x_i y_i$$

$$d_i = d - x_i \; y_i^T$$

b) $$C_i^{-1} = \left(\left(C^T - (x_i \; x_i^T) \right) \right)^{-1}$$

$$C_i = C^{-1} - \frac{-C^{-1} x_i x_i^T C^{-1}}{1 - x_i^T C^{-1} x_i}$$

c) $$C \bar{w} = X^T y$$

$$C \bar{w} = d$$

$$C_i w_i = d_i$$

$$w_i = C_i^{-1} d_i$$

$$w_i = \left(C^{-1} - \frac{C^{-1} x_i x_i^T C^{-1}}{1 - x_i^T C^{-1} x_i} \right) \left(X^T y - x_i y_i \right)$$

$$w_i = \left(C^{-1} - \frac{C^{-1} x_i x_i^T C^{-1}}{1 - x_i^T C^{-1} x_i}\right)\left(X^T y - x_i y_i\right)$$

$$\bar{w}_i = \bar{w} + C^{-1} x_i y_i - \frac{C^{-1} x_i x_i^T C^{-1}}{1 - x_i^T C^{-1} x_i} X^T y - \frac{C^{-1} x_i x_i^{-1}}{1 - x_i^T C^{-1} x_i} x_i y_i$$

$$\bar{w}_i = \bar{w} + C^{-1} x_i \left(\frac{x_i^T w - y_i}{1 - x_i^T C^{-1} x_i}\right)$$

d) $l = \bar{w}_i x_i - y_i$

$$l = x_i^T \left(\bar{w} + C^{-1} x_i \left(\frac{x_i w - y_i}{1 - x_i^T C^{-1} x_i}\right)\right) - y_i$$

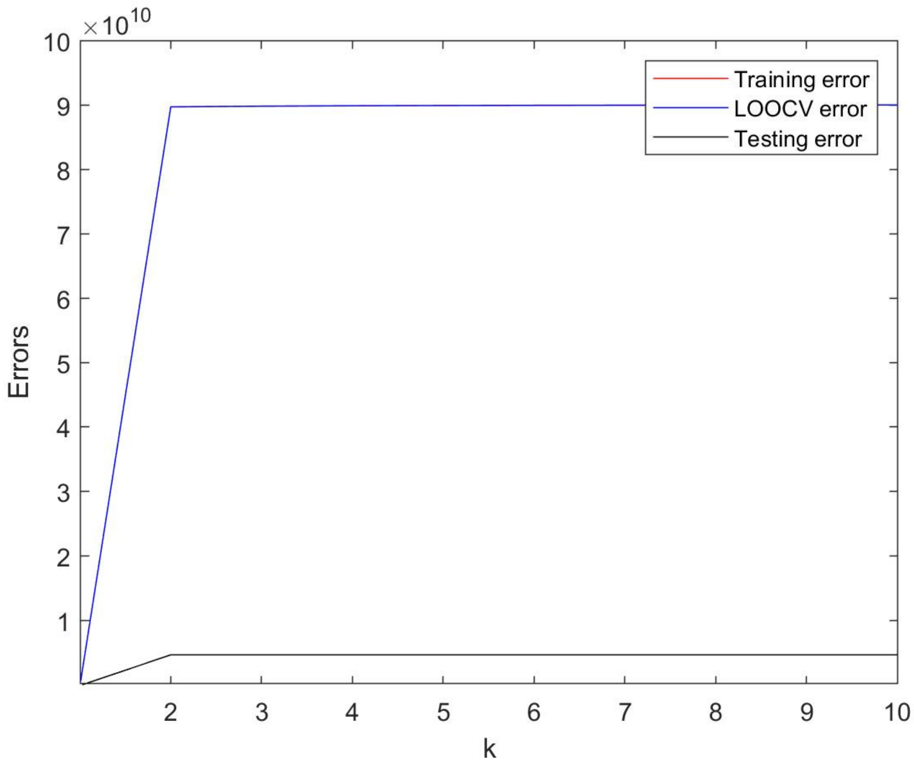
$$l = x_i^T \bar{w} - y_i + \bar{C} x_i x_i^T \left(\frac{x_i w - y_i}{1 - x_i^T C^{-1} x_i}\right)$$

$$w_t^T x_t - y = \frac{w^T x_i - y}{1 - x_i^T C^{-1} x_i}$$

e) This is significantly less computationally complicated than the naive LOOCV since instead of estimating the model m times with d dimensions it is estimated at the end

Problem 3.2

b) This makes sense
the data is underfitted
when we have too many
parameters, we really need
to decrease the number of
parameters to decrease
the error or increase the
number of samples



$$\|w_{t+1}\|_2^2 \le (\|w_i\|^2 + R)\eta$$

after m updates

$$\|w_{T+1}\|_2^2 \le mR^2\eta$$

$$\sqrt{m}$$

$$\gamma \cancel{m} \|w^*\|_2 \le \|w^*\| R \cancel{\sqrt{m}}\, \eta$$

$$m \le \frac{R^2\eta^2}{\gamma}$$

c) The average error slowly decrease but the last solution error is much more volatile