Machine Learning HW#6

Evgeny Marshakov

Problem 1

We know that Y^0 and y have the same distribution, so $\mathbb{E}_{Y^0} Y_i^0 = \mathbb{E}_y y_i$ and $\mathbb{E}_{Y^0}(Y_i^0)^2 = \mathbb{E}_y y_i^2$

$$\mathbb{E}\text{ op} = \mathbb{E}\left[\sum_{i=1}^{N} \mathbb{E}_{Y^0}(Y_i^0 - \hat{f}(x_i))^2 - (y_i - \hat{f}(x_i))^2\right] = \frac{1}{N}\sum_{i=1}^{N} \mathbb{E}_y \left[\mathbb{E}_{Y^0}(Y_i^0)^2 - 2\hat{y}_i \mathbb{E}_{Y^0} Y_i^0 - y_i^2 + 2y_i\hat{y}_i\right] = \quad (1)$$

$$= \frac{1}{N}\sum_{i=1}^{N} \left[\mathbb{E}_{Y^0}(Y_i^0)^2 - 2\mathbb{E}_y \hat{y}_i \mathbb{E}_{Y^0} Y_i^0 - \mathbb{E}_y y_i^2 + 2\mathbb{E}_y y_i\hat{y}_i\right] = \frac{2}{N}\sum_{i=1}^{N} Cov(\hat{y}_i, y_i) \quad (2)$$

Problem 2

$$\sum_{i=1}^{N} Cov(\hat{y}_i, y_i) = \sum_{i=1}^{N} Cov\left(\sum_{j=1}^{N} S_{ji} y_j, y_i\right) = \sum_{i=1}^{N} S_{ii}\sigma_\epsilon^2 = \text{trace}(S)\sigma_\epsilon^2 \quad (3)$$

because we know that y_i and y_j are linearly independent if $i \neq j$.