Homework 2 – Analytical Derivation of LOOCV MSE — PHAM Nam

Objective

Show that the mean squared error for leave-one-out cross-validation (LOOCV) of ℓ_2-regularized regression using a linear model

$$f_\theta(x) = \sum_{j=1}^{b} \theta_j \phi_j(x)$$

can be analytically calculated as:

$$\text{LOOCV MSE} = \frac{1}{n} \left\| \tilde{H}^{-1} H y \right\|^2$$

where:

$$\begin{aligned} H &= \Phi(\Phi^\top \Phi + \lambda I)^{-1} \Phi^\top \\ \tilde{H} &= \operatorname{diag}(H_{11}, H_{22}, \ldots, H_{nn}) \end{aligned}$$

1. Linear Regression Setup

Let $\Phi \in \mathbb{R}^{n \times b}$ be the design matrix, with rows $\phi_i^\top = \Phi_{i,:}$, and $y \in \mathbb{R}^n$ the output vector. The ℓ_2-regularized least squares solution is:

$$\hat{\theta} = (\Phi^\top \Phi + \lambda I)^{-1} \Phi^\top y$$

The predicted output is:

$$\hat{y} = \Phi \hat{\theta} = H y, \quad \text{where } H = \Phi(\Phi^\top \Phi + \lambda I)^{-1} \Phi^\top$$

2. Leave-One-Out Prediction

To compute the prediction $\hat{y}_i^{(-i)}$ excluding the i-th data point, define:

$$U = \Phi^\top \Phi + \lambda I, \quad \phi_i = \Phi_{i,:}^\top$$

Use the Sherman-Morrison-Woodbury identity:

$$(U - \phi_i \phi_i^\top)^{-1} = U^{-1} + \frac{U^{-1} \phi_i \phi_i^\top U^{-1}}{1 - \phi_i^\top U^{-1} \phi_i}$$

This leads to:

$$\hat{y}_i^{(-i)} = \frac{\hat{y}_i - H_{ii} y_i}{1 - H_{ii}} \quad \Rightarrow \quad y_i - \hat{y}_i^{(-i)} = \frac{y_i - \hat{y}_i}{1 - H_{ii}}$$

3. LOOCV MSE

Hence, the LOOCV error becomes:

$$\text{LOOCV MSE} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i^{(-i)}\right)^2 = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - \hat{y}_i}{1 - H_{ii}}\right)^2$$

4. Final Form

This can be expressed using $\tilde{H}^{-1}$ as:

$$\text{LOOCV MSE} = \frac{1}{n}\left\|\tilde{H}^{-1}Hy\right\|^2$$

where $\tilde{H}$ is a diagonal matrix with entries H_{ii}.