

Exercise 1: Risk Minimizers for Generalized L2-Loss

Consider the regression learning setting, i.e., $\mathcal{Y} = \mathbb{R}$, and assume that your loss function of interest is $L(y, f(\mathbf{x})) = (m(y) - m(f(\mathbf{x})))^2$, where $m : \mathbb{R} \rightarrow \mathbb{R}$ is a continuous strictly monotone function.

Disclaimer: In the following we always assume that $\text{Var}(m(Y))$ exists.

- (a) Consider the hypothesis space of constant models $\mathcal{H} = \{f : \mathcal{X} \rightarrow \mathbb{R} \mid f(\mathbf{x}) = \theta \ \forall \mathbf{x} \in \mathcal{X}\}$, where \mathcal{X} is the feature space. Show that

$$\hat{f}(\mathbf{x}) = m^{-1} \left(\frac{1}{n} \sum_{i=1}^n m(y^{(i)}) \right)$$

is the optimal constant model for the loss function above, where m^{-1} is the inverse function of m .

Hint: We can obtain several different notions of a mean value by using a specific function m , e.g., the arithmetic mean by $m(x) = x$, the harmonic mean by $m(x) = 1/x$ (if $x > 0$) or the geometric mean by $m(x) = \log(x)$ (if $x > 0$).

- (b) Verify that the risk of the optimal constant model is $\mathcal{R}_L(\hat{f}) = (1 + \frac{1}{n}) \text{Var}(m(y))$.
- (c) Derive that the risk minimizer (Bayes optimal model) f^* is given by $f^*(\mathbf{x}) = m^{-1}(\mathbb{E}_{y|\mathbf{x}}[m(y) \mid \mathbf{x}])$.
- (d) What is the optimal constant model in terms of the (theoretical) risk for the loss above and what is its risk?
- (e) Recall the decomposition of the Bayes regret into the estimation and the approximation error. Show that the former is $\frac{1}{n} \text{Var}(m(y))$, while the latter is $\text{Var}(\mathbb{E}_{y|\mathbf{x}}[m(y) \mid \mathbf{x}])$ for the optimal constant model $\hat{f}(\mathbf{x})$ if the hypothesis space \mathcal{H} consists of the constant models.

Hint: Use the law of total variance, which states that $\text{Var}(Y) = \mathbb{E}_X[\text{Var}(Y \mid X)] + \text{Var}(\mathbb{E}_{Y|X}[Y \mid X])$, where the conditional variance is defined as $\text{Var}(Y \mid X) = \mathbb{E}_X \left[(Y - \mathbb{E}_{Y|X}(Y \mid X))^2 \mid X \right]$.