

Solution 1: Risk Minimizers for the L2-Loss

- (a) As seen in L2-loss slide 2, the risk minimizer f^* is the conditional mean:

$$f^*(\mathbf{x}) = \mathbb{E}_{y|x} [y|x]$$

The distribution of y given x is known, and can be plugged-in:

$$f^*(\mathbf{x}) = \mathbb{E} [\mathcal{N}(a + bx, 1)] = a + bx$$

- (b) The resulting risk can be calculated by using the definition:

$$\begin{aligned} \mathcal{R}_L(f^*) &= \mathbb{E}_{xy} [(y - f^*(\mathbf{x}))^2] \\ &= \mathbb{E}_{xy} [(y - \mathbb{E}_{y|x} [y|x])^2] \\ &= \mathbb{E}_x [\text{Var}_{y|x}(y|x)] \\ &= \mathbb{E}_x [\text{Var}(\mathcal{N}(a + bx, 1))] \\ &= \mathbb{E}_x [1] \\ &= 1 \end{aligned}$$

- (c) The risk minimizer for the L2 loss is the conditional mean. Considering that the hypothesis space is now restricted to constant models, $f(x)$ is a constant for any x . The optimal constant model in terms of the theoretical risk for the L2 loss is the expected value over y .

$$\bar{f}(x) = \mathbb{E}_{y|x} [y|x] = \mathbb{E}_y [y]$$

The Law of total expectation can be used in this case:

$$\begin{aligned} \mathbb{E}_y [y] &= \mathbb{E}_x [\mathbb{E}_{y|x} [y|x]] \\ &= \mathbb{E}_x [a + bx] \\ &= a + b \cdot \mathbb{E}_x [x] \\ &= a + b \cdot 0 \\ &= a \end{aligned}$$

- (d) To calculate the risk of the optimal constant model, the definition can be used:

$$\begin{aligned} \mathcal{R}_L(\bar{f}) &= \mathbb{E}_{xy} [(y - \bar{f}(x))^2] \\ &= \mathbb{E}_{xy} [(y - a)^2] \\ &= \mathbb{E}_y [y^2] - 2a \underbrace{\mathbb{E}_y [y]}_a + a^2 \\ &= \mathbb{E}_y [y^2] - a^2 \end{aligned}$$

As $\mathbb{E}_y[y^2]$ is yet unknown, the calculation is detailed below:

$$\begin{aligned}
 \mathbb{E}_y[y^2] &= \mathbb{E}_x \left[\mathbb{E}_{y|x} \left[y^2 \right] \right] \\
 &= \mathbb{E}_x \left[\text{Var}_{y|x}[y|x] + \mathbb{E}_{y|x} [y|x] \right] \\
 &= \mathbb{E}_x \left[1 + (a + bx)^2 \right] \\
 &= \mathbb{E}_x[1 + a^2 + 2abx + b^2x^2] \\
 &= 1 + a^2 + 2ab \underbrace{\mathbb{E}_x[x]}_{=0} + b^2 \underbrace{\mathbb{E}_x[x^2]}_{=\frac{100}{3}} \\
 &= 1 + a^2 + b^2 \frac{100}{3}
 \end{aligned}$$

Using this result, the risk can be calculated.

$$\mathcal{R}_L(\bar{f}) = 1 + b^2 \frac{100}{3}$$