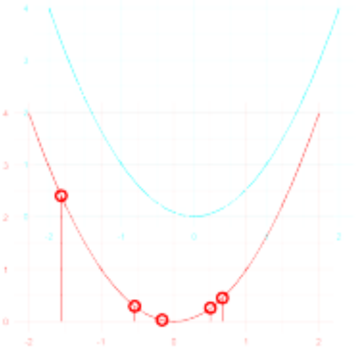
Introduction to Machine Learning

Advanced Risk Minimization

Regression Losses: L2 and L1 loss

Regression Losses: L2 and L1 loss



Learning goals

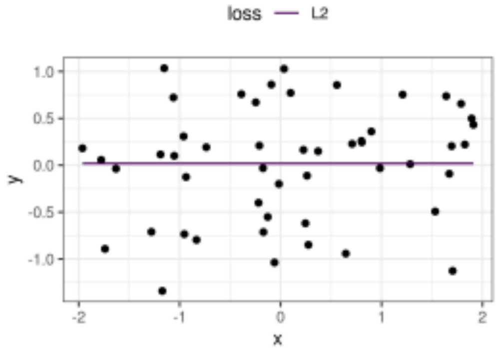
- Derive the risk minimizer of the L2-loss
- Derive the optimal constant model for the L2-loss
- Know risk minimizer and optimal constant model for L1-loss

Learning goals

- Derive the risk minimizer of the L2-loss
- Derive the optimal constant model for the L2-loss
- Know risk minimizer and optimal constant model for L1-loss

L2-LOSS: OPTIMAL CONSTANT MODEL / 2

The optimizer $\hat{f}_c$ of the empirical risk is $\bar{y}$ (the empirical mean over $y^{(i)}$), which is the empirical estimate for $\mathbb{E}_y[y]$.



L2-LOSS: OPTIMAL CONSTANT MODEL / 3

Proof:

For the optimal constant model f_c^* for the L2-loss $L\left(y, f\right) = \left(y - f\right)^2$ we solve the optimization problem

$$\arg\min_{f \in \mathcal{H}} \mathcal{R}_{\text{emp}}(f) = \arg\min_{\theta \in \mathbb{R}} \sum_{i=1}^{n} (y^{(i)} - \theta)^2.$$

We calculate the first derivative of $\mathcal{R}_{\text{emp}}$ w.r.t. θ and set it to 0:

$$\begin{aligned}
\frac{\partial \mathcal{R}_{\text{emp}}(\theta)}{\partial \theta} = -2 \sum_{i=1}^{n} \left(y^{(i)} - \theta\right) &\overset{!}{=} 0 \\
\sum_{i=1}^{n} y^{(i)} - n\theta &= 0 \\
\hat{\theta} &= \frac{1}{n} \sum_{i=1}^{n} y^{(i)} =: \bar{y}.
\end{aligned}$$