

线性回归推导

2020年4月23日 星期四 上午11:17

线性回归

3.1 最小二乘法 (超拟合, 欠拟合) \rightarrow 拟合问题的误差最小

损失函数: $L(w) = \sum_{i=1}^N ||w^T x_i - y_i||^2$

正则化: $L_1 \rightarrow Lasso$, $L_2 \rightarrow Ridge$

数据表示: $D = \{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$

矩阵表示: $X = \begin{bmatrix} x_1^T \\ \vdots \\ x_N^T \end{bmatrix}$, $y = \begin{bmatrix} y_1 \\ \vdots \\ y_N \end{bmatrix}$

最小二乘估计: $\hat{w} = (X^T X)^{-1} X^T y$

3.2 线性回归的概率分布

假设: $x_i \in \mathbb{R}^p$, $y_i \in \mathbb{R}$

模型: $y_i = w^T x_i + \epsilon_i$

噪声分布: $\epsilon_i \sim \mathcal{N}(0, \sigma^2)$

似然函数: $L(w) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(y_i - w^T x_i)^2}{2\sigma^2})$

最大似然估计: $\hat{w} = \arg \max_w L(w)$

正则化: L_1 (Lasso), L_2 (Ridge)

3.3 线性回归的正则化

损失函数: $L(w) = \sum_{i=1}^N ||w^T x_i - y_i||^2$

正则化项: $\lambda ||w||^2$

正则化损失: $J(w) = L(w) + \lambda ||w||^2$

正则化估计: $\hat{w} = (X^T X + \lambda I)^{-1} X^T y$

3.4 正则化框架

正则化框架: $J(w) = L(w) + \lambda P(w)$

正则化项: $P(w)$

正则化估计: $\hat{w} = \arg \min_w J(w)$

正则化估计: $\hat{w} = \arg \min_w (L(w) + \lambda ||w||^2)$