

# 高斯噪声

Guangyao Zhao

2022-02-03

## Contents

高斯噪声指的是概率密度函数服从高斯分布的一类噪声： $\epsilon \sim N(0, \sigma^2)$ 。在线性回归中，隐藏的一个假设就是噪声（预测值和真实值之差）服从高斯分布：

$$\begin{aligned}y_i &= \hat{y}_i + \epsilon \\ &= w^T x_i + \epsilon\end{aligned}$$

即： $y_i|x_i; w \sim N(w^T x_i, \sigma^2)$ 。利用最大似然估计：

$$\begin{aligned}L(w|\mathbf{X}) &= \ln p(\mathbf{X}|w) = \ln \prod_{i=1}^n p(y_i|x_i; w) = \sum_{i=1}^n \ln p(y_i|x_i; w) \\ &= \sum_{i=1}^n \left( \ln \frac{1}{\sqrt{2\pi}\sigma} - \frac{(y_i - w^T x_i)^2}{2\sigma^2} \right) \\ &= - \sum_{i=1}^n (y_i - w^T x_i)^2\end{aligned}$$

求得目标函数：

$$\begin{aligned}\hat{w} &= \operatorname{argmax}_w - \sum_{i=1}^n (y_i - w^T x_i)^2 \\ &= \operatorname{argmin}_w \sum_{i=1}^n (y_i - w^T x_i)^2 \\ &= \operatorname{argmin}_w \sum_{i=1}^n (y_i - \hat{y})^2\end{aligned}$$