高斯分布

hyliu

2020 年 4 月 2 日

目录

1 高斯分布 **1**
1.1 参数估计 1
1.1.1 问题定义 1
1.1.2 高斯分布公式 1
1.1.3 Maximum likelihood estimation (MLE) 2
1.1.4 公式推导 2
1.2 Linear Gaussian Model (线性高斯模型) 4
1.3 有偏误差在贝叶斯网络中的体现 4

1 高斯分布

1.1 参数估计

1.1.1 问题定义

$$Data : X = (x_1, x_2, ..., x_N)^T = \begin{pmatrix} x_1^T \\ x_2^T \\ ... \\ x_N^T \end{pmatrix}_{N*P}$$

$$x_i \in \mathbb{R}^P$$

$$x_i \sim \mathcal{N}(\mu, \Sigma)$$

$$\theta = (\mu, \Sigma)$$

1.1.2 高斯分布公式

1. 一维高斯分布公式

$$P(x) = \frac{1}{2\pi\sigma} \exp(-\frac{(x-\mu)^2}{2\sigma^2})$$

1. 多维高斯分布公式

$$P(x) = \frac{1}{(2\pi)^{\frac{1}{P}} |\Sigma|^{\frac{1}{2}}} \exp(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)) \tag{1}$$

Σ 为协方差矩阵

1.1.3 Maximum likelihood estimation (MLE)

$$\theta_{MLE} = \arg\max_{\theta} P(X|\theta)$$

当 $P=1$, $\theta = (\mu, \sigma^2)$

1.1.4 公式推导

$$\begin{aligned}
\log P(X|\theta) &= log \sum_{i=1}^{N} P(x_i|\theta) = \sum_{i=1}^{N} \log P(x_i|\theta) \\
&= \sum_{i=1}^{N} \log \frac{1}{2\pi\sigma} \exp(-\frac{(x_i-\mu)}{2\sigma^2}) \\
&= \sum_{i=1}^{N} \left[\log \frac{1}{\sqrt{2\pi}} + log \frac{1}{\sigma} - \frac{(x_i-\mu)^2}{2\sigma^2} \right]
\end{aligned} \tag{2}$$

μ_{MLE} 是无偏估计, σ_{MLE} 是有偏估计。

1. μ_{MLE} 推导

$$\begin{aligned}\mu_{MLE} &= \arg\max_{\mu} \log P(X|\theta) \\ &= \arg\max_{\mu} \sum_{i=1}^{N} -\frac{(x_i-\mu)^2}{2\sigma^2} \\ &= \arg\min_{\mu} \sum_{i=1}^{N} (x_i-\mu)^2\end{aligned} \tag{3}$$

$$\begin{aligned}\frac{\partial}{\partial\mu}\sum (x_i-\mu)^2 = \sum_{i=1}^{N} 2*(x_i-\mu)*(-1) = 0 \\ \sum_{i=1}^{N}(x_i-\mu) = 0 \\ \sum_{i=1}^{N} x_i - \sum_{i=1}^{N}\mu = 0 \\ N*\mu = \sum_{i=1}^{N} x_i \\ \mu_{MLE} = \frac{1}{N}\sum_{i=1}^{N} x_i\end{aligned} \tag{4}$$

$$E\left(\mu_{MLE}\right) = \frac{1}{N}\sum_{i=1}^{N} E[x_i] = \frac{1}{N}\sum_{i=1}^{N}\mu = \mu$$

2. σ_{MLE} 推导

$$\begin{aligned}\sigma^2_{MLE} &= \arg\max_{\sigma} P(X|\theta) \\ &= \arg\max_{\sigma} \sum_{i=1}^{N}(-\log\sigma - \frac{(x_i-\mu_i)^2}{2\sigma^2})\end{aligned} \tag{5}$$

$$
\begin{aligned}
&\mathcal{L}(\sigma) = -\log\sigma - \frac{(x_i-\mu_i)^2}{2\sigma^2} \\
&\frac{\partial\mathcal{L}}{\partial\sigma} = \sum_{i=1}^{N}\left[-\frac{1}{\sigma} + \sigma^{-3}\left(x_i-\mu\right)^2\right] \\
&\sum_{i=1}^{N}\left[-\sigma^2 + \left(x_i-\mu\right)^2\right] = 0 \\
&-N\sigma^2 + \sum_{i=1}^{N}\left(x_i-\mu\right)^2 = 0 \\
&\sigma^2_{MLE} = \frac{1}{N}\sum_{i=1}^{N}\left(x_i-\mu_{MLE}\right)^2
\end{aligned}
\tag{6}
$$

$$
\sigma^2_{MLE} = \frac{1}{N}\sum_{i=1}^{N}\left(x_i-\mu_{MLE}\right)^2 = \frac{1}{N}\sum_{i=1}^{N}\left(x_i^2 - 2x_i\mu_{MLE} + \mu_{MLE}\right) = \frac{1}{N}\sum_{i=1}^{N}x_i^2 - \frac{1}{N}\sum_{i=1}^{N}2x_i\mu_{MLE} +
$$

$$
Var(\mu_{MLE}) = Var(\frac{1}{N}\sum_{i=1}^{N}x_i) = \frac{1}{N^2}\sum_{i=1}^{N}Var(x_i) = \frac{1}{N}Var(x_i) = \frac{1}{N}\sigma^2
$$

$$
\begin{aligned}
E[\sigma^2_{MLE}] &= E[\frac{1}{N}\sum_{i=1}^{N}x_i^2 - \mu^2_{MLE}] = E[(\frac{1}{N}\sum_{i=1}^{N}x_i^2 - \mu^2) - \left(\mu^2_{MLE} - \mu^2\right)] \\
&= E[\frac{1}{N}\sum_{i=1}^{N}x_i^2 - \mu^2] - E(\mu^2_{MLE} - \mu^2) \\
&= [\frac{1}{N}\sum_{i=1}^{N}E(x_i^2 - \mu^2)] - [E(\mu^2_{MLE}) - E(\mu^2)] \\
&= [\frac{1}{N}\sum_{i=1}^{N}(E(x_i^2) - \mu^2)] - [E(\mu^2_{MLE}) - \mu^2] \\
&= [\frac{1}{N}\sum_{i=1}^{N}(Var(x_i))] - [E(\mu^2_{MLE}) - E(\mu^2_{MLE})^2] \\
&= [\sigma^2] - [Var(\mu_{MLE})] \\
&= [\sigma^2] - [\frac{1}{N}\sigma^2] \\
&= \frac{N-1}{N}\sigma^2
\end{aligned}
\tag{7}
$$

$$E(\sigma_{MLE}) = \frac{N-1}{N}\sigma^2$$

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(x_i - \mu_{MLE})^2$$

1.2 Linear Gaussian Model (线性高斯模型)

1.3 有偏误差在贝叶斯网络中的体现