Abstract

本期我们主要学习高斯分布的一些性质

Assumption

现在我们有一堆数据：

$$X = (x_1, x_2, \dots, x_N)^T \tag{18}$$
$$x_i \in R^p \tag{19}$$

首先给出我们的模型：高斯线形模型。

这里我们为了简化起见，将 p 设为 1， 因此

$$x \backsim N(\mu, \sigma^2) \tag{20}$$
$$\theta = (\mu, \sigma) \tag{21}$$

接下来我们根据这堆数据，通过极大似然估计(MLE)得出其期望与方差

下面我们给出似然函数：

$$\begin{aligned} p(X|\theta) &= log(\prod_{i=1}^{N} \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)) \\ &= \sum_{i=1}^{N} log(\frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{(x_i-\mu)^2}{2\sigma^2}\right)) \\ &= \sum_{i=1}^{N} log(\frac{1}{\sqrt{2\pi}}) - log(\sigma) - \frac{(x_i-\mu)^2}{2\sigma^2} \end{aligned} \tag{22}$$

Expectation

下面我们首先使用极大似然估计得出期望 μ 的估计值

$$\begin{aligned} \mu_{MLE} &= argmax(p(X|\theta)) \\ &= argmin(\sum_{i=1}^{N}(x_i-\mu)^2) \end{aligned} \tag{23}$$

对式子求导得到：

$$\sum_{i=1}^{N} 2(x_i - \mu) = 0 \tag{24}$$

$$\sum_{i=1}^{N} x_i - N\mu = 0 \tag{25}$$

$$\mu_{MLE} = \frac{1}{N}\sum_{i=1}^{N} x_i \tag{26}$$

Variance

同样的，我们使用极大似然估计得出方差 σ 的估计值

$$\begin{aligned}\sigma_{MLE} &= argmax(p(X|\theta)) \\ &= argmin(\sum_{i=1}^{N} log(\sigma) + \frac{(x_i-\mu)^2}{2\sigma^2})\end{aligned} \tag{27}$$

同样的，我们对式子求导得到：

$$\sum_{i=1}^{N}[\frac{1}{\sigma} - \frac{(x_i-\mu)^2}{\sigma^3}] = 0 \tag{28}$$

最后，我们得到估计值:

$$\sigma_{MLE}^2 = \Sigma_{MLE} = \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu)^2 \tag{29}$$

Biased Estimation

要验证一个估计值是有偏估计还是无偏估计，我们只需计算该估计值的期望即可。

μ

$$\begin{aligned}E[\mu_{MLE}] &= E[\frac{1}{N}\sum_{i=1}^{N}x_i] \\ &= \frac{1}{N}\sum_{i=1}^{N}E[x_i] \\ &= \mu\end{aligned} \tag{30}$$

因此 μ_{MLE} 为无偏估计

σ

首先我们对 σ 的估计值进行变形

$$
\begin{aligned}
\sigma_{MLE}^2 &= \frac{1}{N}\sum_{i=1}^{N}(x_i-\mu_{MLE})^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}(x_i^2-2x_i\mu_{MLE}+\mu_{MLE}^2) \\
&= \frac{1}{N}\sum_{i=1}^{N}x_i^2-2(\frac{1}{N}\sum_{i=1}^{N}x_i)\mu_{MLE}+\mu_{MLE}^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}x_i^2-2\mu_{MLE}^2+\mu_{MLE}^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}x_i^2-\mu_{MLE}^2 \\
&= (\frac{1}{N}\sum_{i=1}^{N}x_i^2-\mu^2)-(\mu_{MLE}^2-\mu^2)
\end{aligned} \tag{31}
$$

令 $f_1=(\frac{1}{N}\sum_{i=1}^{N}x_i^2-\mu^2)$, $f_2=(\mu_{MLE}^2-\mu^2)$

所以:

$$
\begin{aligned}
E[f_1] &= E[\frac{1}{N}\sum_{i=1}^{N}x_i^2-\mu^2] \\
&= E[\frac{1}{N}\sum_{i=1}^{N}(x_i^2-\mu^2)] \\
&= \frac{1}{N}\sum_{i=1}^{N}E[x_i^2]-E[\mu^2] \\
&= \frac{1}{N}\sum_{i=1}^{N}E[x_i^2]-\mu^2 \\
&= \frac{1}{N}\sum_{i=1}^{N}E[x_i^2]-(E[x_i])^2 \\
&= \sigma^2
\end{aligned} \tag{32}
$$

类似的:

$$
\begin{aligned}
E[f_2] &= E[\mu_{MLE}^2-\mu^2] \\
&= E[\mu_{MLE}^2-(E[\mu_{MLE}])^2] \\
&= Var[\mu_{MLE}] \\
&= Var[\frac{1}{N}\sum_{i=1}^{N}x_i] \\
&= \frac{1}{N^2}\sum_{i=1}^{N}Var[x_i] \\
&= \frac{1}{N}\sigma^2
\end{aligned} \tag{33}
$$

最后，将 f_1 与 f_2 相加，得到:

$$E[\sigma^2_{MLE}] = \frac{N-1}{N}\sigma^2 \tag{34}$$

因此我们通过极大似然估计得到的 σ 的估计值比真实值略小，所以为有偏估计

而 σ^2 的无偏估计为 $\frac{1}{N-1}\sum_{i=1}^{N}(x_i-\mu_{MLE})^2$