参数估计的评价准则

3月机器学习在线班　邹博

2015年3月8日

历史遗留问题

- 估计量的优良性准则
  - 无偏性
  - 均方误差准则

无偏性 $E(\hat{\theta}) = \theta$

- 利用已知样本 $X_1, X_2, \cdots X_n$ 能够得到参数的一个估计 $\hat{\theta}$，因此，$\hat{\theta}$ 可以写成 $\hat{\theta}(X_1, X_2, \cdots X_n)$，对于不同的样本，$\hat{\theta}$ 的值一般不同。因此，可以看成是关于样本的随机变量。它是可以求均值的：$E(\hat{\theta})$
- 如果 $E(\hat{\theta})$ 等于总体的实际分布θ，就说这个估计是无偏估计。$E(\hat{\theta}) = \theta$
  - 用 $\hat{\theta}$ 去估计θ，有时偏高，有时偏低，但平均来说，它等于位置参数θ

举例

- 无偏估计

$$\hat{\mu} = X_1$$

$$\hat{\mu} = \frac{X_1 + X_2}{2}$$

$$\hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

- 不是无偏估计

$$\hat{\mu} = 2X_1$$

$$\hat{\mu} = \frac{X_1 + X_2}{3}$$

$$\hat{\mu} = \frac{1}{n-3}\sum_{i=1}^{n} X_i$$

样本均值和方差是总体的无偏估计

- 设总体均值为μ，方差为σ²，$X_1,X_2,\ldots X_n$为来自该总体的样本，即：

$$\overline{X}=\frac{1}{n}\sum_{i=1}^{n}X_i$$

$$S=\frac{1}{n-1}\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2$$

- 则：

$$E\left(\overline{X}\right)=\mu$$

$$E\left(S^2\right)=\sigma^2$$

均值的无偏性

- 因为$X_1,X_2,\ldots X_n$为同分布的，于是$E(X_i)=\mu$，所以：

$$E\left(\overline{X}\right)=\frac{1}{n}\sum_{i=1}^{n}X_i=\frac{1}{n}\cdot n\mu=\mu$$

方差的无偏性

- 首先 $Var\left(\overline{X}\right)=Var\left(\frac{1}{n}\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}Var\left(\sum_{i=1}^{n}X_i\right)=\frac{1}{n^2}\sum_{i=1}^{n}Var\left(X_i\right)=\frac{\sigma^2}{n}$
- 因此

$$E\left(\overline{X}^2\right)=Var\left(\overline{X}\right)+\left[E\left(\overline{X}\right)\right]^2=\frac{\sigma^2}{n}+\mu^2$$

$$E\left(X_i^2\right)=Var\left(X_i\right)+\left[E\left(X_i\right)\right]^2=\sigma^2+\mu^2$$

- 而

$$\sum_{i=1}^{n}\left(X_i-\overline{X}\right)^2=\sum_{i=1}^{n}X_i^2-2\left(\sum_{i=1}^{n}X_i\right)\overline{X}+n\overline{X}^2=\sum_{i=1}^{n}X_i^2-2\left(n\overline{X}\right)\cdot\overline{X}+n\overline{X}^2=\sum_{i=1}^{n}X_i^2-n\overline{X}^2$$

- 所以

$$E\left(S^2\right)=\frac{1}{n-1}\left[E\left(\sum_{i=1}^{n}X_i^2\right)-nE\left(\overline{X}^2\right)\right]=\frac{1}{n-1}\left[\left(n\sigma^2+n\mu^2\right)-n\left(\frac{\sigma^2}{n}+\mu^2\right)\right]=\sigma^2$$

均方误差准则 $MSE(\hat{\theta}) = E\left[(\hat{\theta}-\theta)^2\right]$

- 用估计量 $\hat{\theta}$ 去估计θ，其误差是 $\hat{\theta}-\theta$，该误差显然随样本 $X_1, X_2, \cdots X_n$ 而定，因此，$\hat{\theta}-\theta$ 是随机变量，它的平方的均值，称作均方误差。这个量越小，平均误差越小，估计结果越优。

$$MSE(\hat{\theta}) = E\left[(\hat{\theta}-\theta)^2\right]$$

- 显然，若 $\hat{\theta}$ 是无偏估计，则MSE即方差。

$$MSE(\hat{\theta}) = E\left[(\hat{\theta}-\theta)^2\right] = E\left[\left(\hat{\theta}-E(\hat{\theta})\right)^2\right] = Var(\hat{\theta})$$

感谢大家！

恳请大家批评指正！