CZ4041: Tutorial Week 11&12

Due on April 8, 2021 at 8:30am

Assoc Prof Pan, Sinno Jialin - CS4

Pang Yu Shao
U1721680D

08/04/2021

Problem 1

Suppose a dataset of four 3-dimensional instances is shown in Table 1. Estimate the sample mean and covariance matrix (unbiased).

Table 1: Data set for Question 1.

ID	X_1	X_2	X_3
P1	3	5	-1
P2	-1	8	3
P3	2	-4	-4
P4	0	-1	-6

Solution

Calculate sample mean (unbiased):

$$\begin{aligned}\hat{\mu} &= \frac{1}{N}\sum_{i=1}^{N} x_i \\ &= \frac{1}{4}\begin{bmatrix}(3-1+2) & (5+8-4-1) & (3-1-4-6)\end{bmatrix} \\ &= \begin{bmatrix}1 & 2 & -2\end{bmatrix}\end{aligned}$$

Therefore, centered data matrix:

$$\tilde{X} = \begin{bmatrix}2 & 3 & 1 \\ -2 & 6 & 5 \\ 1 & -6 & -2 \\ -1 & -3 & -4\end{bmatrix}$$

Calculate sample covariance (unbiased):

$$\begin{aligned}\tilde{\Sigma} &= \frac{1}{N-1}\sum_{i=1}^{N}(x_i-\hat{\mu})(x_i-\hat{\mu})^T \\ &= \frac{1}{3}\tilde{X}^T\tilde{X} \\ &= \frac{1}{3}\begin{bmatrix}2 & -2 & 1 & -1 \\ 3 & 6 & -6 & -3 \\ 1 & 5 & -2 & -4\end{bmatrix}\begin{bmatrix}2 & 3 & 1 \\ -2 & 6 & 5 \\ 1 & -6 & -2 \\ -1 & -3 & -4\end{bmatrix} \\ &= \frac{1}{3}\begin{bmatrix}10 & -9 & -6 \\ -9 & 90 & 57 \\ -6 & 57 & 46\end{bmatrix} \\ &= \begin{bmatrix}3.33 & -3 & -2 \\ -3 & 30 & 19 \\ -2 & 19 & 15.33\end{bmatrix}\end{aligned}$$

Problem 2

Suppose a dataset of 5 1-dimensional instances is shown in Table 2. Use histogram estimator with an origin of 0 and a width of 3, naive estimator with a width of 3, and 3-NN estimator to estimate the density function $\hat{p}(x)$ and compute the value of $\hat{p}(2.6)$ at 2.6, respectively.

Table 2: Data set for Question 2.

P1	P2	P3	P4	P5
1.2	2	10	-6	3.5

Solution
For **histogram estimator**:
Window: $0 \leq x_i < 3$

$$\begin{aligned}\hat{p}(2.6) &= \frac{2}{5*3} \\ &= 0.133\end{aligned}$$

For **naive estimator**:
Window: $1.1 \leq x_i < 4.1$

$$\begin{aligned}\hat{p}(2.6) &= \frac{3}{5*3} \\ &= 0.2\end{aligned}$$

For **K-NN estimator**:
Distance from x=2.6:

P2: 0.6
P5: 0.9
P1: 1.4 (3rd nearest neighbour)
P3: 7.4
P4: 8.6

$$\begin{aligned}\hat{p}(2.6) &= \frac{3}{5*(2*1.4)} \\ &= 0.214\end{aligned}$$