

The Pattern Recognition 's homework

Wu Bingzhe

1200010666

The school of mathematical science

October 20, 2015

1 Ex 1

The log-likelihood function is:

$$L(\mu) = \ln p(\mathcal{X}|\mu) = \sum_{i=1}^N \ln p(x_i|\mu) = \frac{1}{2} \sum_{i=1}^N (x_i - \mu)^2 + C \quad (1)$$

The derivation of $L(\mu)$, we could get that :

$$\frac{\partial L(\mu)}{\partial \mu} = \sum_{i=1}^N x_i - N\mu = 0$$

So the maximum likelihood estimation of $\hat{\mu} = \frac{1}{N} \sum_{i=1}^N x_i$

The bayesian estimation :

First we get the posterior:

$$\begin{aligned} p(\mu|\mathcal{X}) &= \frac{p(\mathcal{X}|\mu)p(\mu)}{\int p(\mathcal{X}|\mu)p(\mu)d\mu} \\ &= \alpha \prod_{i=1}^N p(x_i|u)p(u) \\ &= \alpha \prod_{i=1}^N \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x_i - u)^2}{2\sigma^2}\right] \cdot \sqrt{2\pi}\sigma_0 \exp\left[-\frac{(u - u_0)^2}{2\sigma_0^2}\right] \\ &= \alpha' \left[-\frac{1}{2} \left[\left(\frac{N}{\sigma^2} + \frac{1}{\sigma_0^2}\right)\mu^2 - 2\left(\frac{1}{\sigma^2} \sum_{i=1}^N x_i + \frac{\sigma_0}{\sigma_0^2}\right)\mu\right]\right] \end{aligned}$$

According to the formula ,apply the method of undetermined coefficients ,we can get that :

$$\hat{u} = \int \mu p(\mu|\mathcal{X})d\mu = \int \mu \frac{1}{\sqrt{2\pi}\sigma_N} \exp\left[-\frac{1}{2}\left(\frac{u - u_N}{\sigma_N}\right)^2\right]du = u_N$$

Combine with $\sigma = 1, u_0 = 0, \sigma_0 = 1$, we can get the bayesian estimation :

$$\hat{\mu} = \frac{1}{N+1} \sum_{i=1}^N x_i$$

2 Ex 2

According to the theory of maximum likelihood estimation and derivation of EX1,

$$l(\theta) = P(x|\theta) = \prod_{k=1}^n p(x_k|\theta)$$

$$H(\theta) = \ln l(\theta) = \sum_{k=1}^n \ln p(x_k|\theta) = \sum_{k=1}^n \sum_{i=1}^d [x_{ki} \ln(\theta_i) + (1 - x_{ki}) \ln(1 - \theta_i)]$$

In terms of $0 = \frac{\partial H(\theta)}{\partial \theta_i} = \sum_{k=1}^n \left[\frac{x_{ki}}{\theta_i} - \frac{1 - x_{ki}}{1 - \theta_i} \right]$

We get that:

$$\hat{\theta} = \frac{1}{n} \sum_{k=1}^n x_{ki}, i = 1, \dots, d;$$

So, $\hat{\theta} = \frac{1}{n} \sum_{k=1}^n x_k$

3 Ex3

According to the theory of bayes estimation, and the derivation of Ex1, the results of multivariate normal distribution is :

$$\begin{cases} \Sigma_N^{-1} = \Sigma_0^{-1} + N\Sigma^{-1} \\ u_N^T = (\Sigma^{-1} \sum_{i=1}^N x_i^T + \mu_0^T \Sigma_0^{-1}) \Sigma_N \end{cases} \quad (2)$$

In terms of the equation of the matrix ,we have:

$$\begin{aligned} \Sigma_N &= (\Sigma_0^{-1} + (\frac{1}{N}\Sigma)^{-1})^{-1} \\ &= \Sigma_0(\Sigma_0 + \frac{1}{N}\Sigma)^{-1} \frac{1}{N}\Sigma \\ &= \Sigma_0(N\Sigma_0 + \Sigma)^{-1}\Sigma \\ \mu_N^T &= (\Sigma^{-1} \sum_{i=1}^N x_i^T + u_0^T \Sigma_0^{-1}) \cdot \Sigma_0(N\Sigma_0 + \Sigma)^{-1}\Sigma \\ &= \sum_{i=1}^N x_i^T (N\Sigma_0 + \Sigma)^{-1}\Sigma_0 + \mu_0^T (N\Sigma_0 + \Sigma)^{-1}\Sigma \\ &= \Sigma_0(N\Sigma_0 + \Sigma)^{-1} \sum_{i=1}^N x_i + \Sigma(N\Sigma_0 + \Sigma)^{-1}\mu_0 \end{aligned}$$

4 Ex4

4.1 (1)

The maximum likelihood estimation of the mean of the normal distribution μ is $\hat{\mu} = \frac{1}{N} \sum_{i=1}^N x_i$. So, we have:

$$E[\hat{\mu}] = E\left[\frac{1}{N} \sum_{i=1}^N x_i\right] = \frac{1}{N} \sum_{i=1}^N \mu = \mu$$

So the $\hat{\mu}$ is unbiased estimation.

4.2 (2)

The maximum likelihood estimation of the variance of the normal distribution σ^2 is $\hat{\sigma}^2 = \frac{1}{N} \sum_{i=1}^N (x_i - \hat{\mu})^2$. So:

$$\begin{aligned} E[\hat{\sigma}^2] &= E\left[\frac{1}{N} \sum_{i=1}^N (x_i - \hat{\mu})^2\right] \\ &= E\left[\frac{1}{N} \sum_{i=1}^N x_i^2 - \hat{\mu}^2\right] \\ &= \frac{1}{N} \sum_{i=1}^N (\sigma^2 + \mu^2) - \left(\frac{\sigma^2}{N} + \mu^2\right) \\ &= \frac{N-1}{N} \sigma^2 \neq \sigma^2 \end{aligned}$$

So $\hat{\sigma}^2$ is not a unbiased estimation.

5 Ex5

5.1 (1)

$$\begin{aligned} \lim_{N \rightarrow \infty} E[\hat{p}_N(x)] &= \lim_{N \rightarrow \infty} E\left[\frac{1}{NV_N} \sum_{i=1}^N \phi\left(\frac{x - x_i}{h_N}\right)\right] \\ &= E\left[\lim_{N \rightarrow \infty} \sigma_N(x - x_i)\right] \\ &= \int \delta(x - V) p(V) dv \\ &= p(x) \end{aligned}$$

5.2 (2)

$$\begin{aligned}
\lim_{N \rightarrow \infty} \text{Var}[\hat{p}_N(x)] &= \lim_{N \rightarrow \infty} E\hat{p}_N^2(x) - (E\hat{p}_N(x))^2 \\
&= \lim_{N \rightarrow \infty} E\left[\frac{1}{NV_N} \sum_{i=1}^N \phi\left(\frac{x-x_i}{h_N}\right) \frac{1}{NV_N} \sum_{j=1}^N \phi\left(\frac{x-x_j}{h_N}\right)\right] - \left[\lim_{N \rightarrow \infty} E\hat{p}_N(x)\right]^2 \\
&= \frac{1}{N^2} E\left[\lim_{N \rightarrow \infty} \frac{1}{V_N} \sum_{i=1}^N \phi\left(\frac{x-x_i}{h_N}\right) \lim_{N \rightarrow \infty} \frac{1}{V_N} \sum_{j=1}^N \phi\left(\frac{x-x_j}{h_N}\right)\right] - p^2(x) \\
&= E\left[\lim_{N \rightarrow \infty} \delta_N(x-V_1)\delta(x-V_2)\right] - p^2(x) \\
&= E(\delta(x-V_1)\delta(x-V_2)) - p^2(x) \\
&= \int \delta^2(x-V)^2 p^2(V) dv - p^2(x) \\
&= p^2(x) - p^2(x) \\
&= 0
\end{aligned}$$