統計學習初論（105-2）

Self-Assessment Homework 1

國立台灣大學資管系

Question 1

Let $x = [x_1\ x_2 \cdots x_D]$ be jointly Gaussian, that is, $x \sim MN(\mu,\ \Sigma) = \frac{1}{2\pi^{D/2}} \frac{1}{|\Sigma|^{1/2}} exp\left\{-\frac{1}{2}(x - \mu)^T \Sigma^{-1}(x-\mu)\right\}$, where μ is a $D \times 1$ vector and Σ is a $D \times D$ symmetric square matrix. We partitioned x into two groups $x = \begin{bmatrix} x_a \\ x_b \end{bmatrix}$, where $x_a = [x_1\ x_2\ \ldots x_M]^T$, and $x_b = [x_{M+1}\ \ldots x_D]^T$. As a result, we can partition the mean and covariance matrix accordingly: $\mu = \begin{bmatrix} \mu_a \\ \mu_b \end{bmatrix}, \Sigma = \begin{bmatrix} \Sigma_{aa} & \Sigma_{ab} \\ \Sigma_{ba} & \Sigma_{bb} \end{bmatrix}$.

Show that the marginal distribution $f(x_a) = \int MN(x|\mu, \Sigma) dx_b = MN(x_a|\mu_a, \Sigma_{aa})$.

Question 2

- We have derived the maximum likelihood estimator of multivariate Gaussian. To briefly recap the result, Given identically and independently distributed data $X = (x_1, x_2,\ \ldots,\ x_N)^T$. Assuming that each data point x_i follows a multivariate Gaussian distribution $MN(\mu, \Sigma)$. The maximum likelihood estimator (MLE) of μ is $\mu_{ML} = \frac{1}{N}\sum_{i=1}^{N} x_i$, and the MLE of Σ is $\Sigma_{ML} = \frac{1}{N}\sum_{i=1}^{N}(x_i - \mu_{ML})\ (x_i - \mu_{ML})^T$.

(a) Show that $E[\mu_{ML}] = \mu$.

(b) Show that $E[\Sigma_{ML}] = \frac{N-1}{N}\Sigma$.

Question 3: Exercise 1.9

Question 4: Exercise 1.10

Question 5: Exercise 2.26