Assignment Sheet - II

Stylianos Sygletos
s.sygletos@aston.ac.uk

May 17, 2021

Problem 1 [2 points]

Show that the mean and variance of the Gamma probability density function :

$$Gamma(x|\alpha, b) = \frac{b^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-bx}, \qquad \alpha, b, x > 0 \tag{1}$$

are given by

$$\mathrm{E}[x] = \frac{\alpha}{b}, \quad \text{and} \quad \sigma_x^2 = \frac{\alpha}{b^2}$$

Problem 2 [2 points]

Prove the below identity

$$\begin{pmatrix} A & B \\ C & D \end{pmatrix}^{-1} = \begin{pmatrix} M & -MBD^{-1} \\ -D^{-1}CM & D^{-1} + D^{-1}CMBD^{-1} \end{pmatrix} \tag{2}$$

where we have defined $M = (A - BD^{-1}C)^{-1}$.

Exercise 3 [2 points]

Using the above (problem 2) partitioned matrix formula, and the precision matrix R:

$$R = \begin{pmatrix} \Lambda + A^T LA & -A^T L \\ -LA & L \end{pmatrix} \tag{3}$$

show that its inverse, i.e. R^{-1}, is given by the covariance matrix:

$$\mathrm{cov}[z] = R^{-1} = \begin{pmatrix} \Lambda^{-1} & \Lambda^{-1}A^T \\ A\Lambda^{-1} & L^{-1} + A\Lambda^{-1}A^T \end{pmatrix} \tag{4}$$

Exercise 4 [3 points]

Consider two multidimensional random vectors $\boldsymbol{x}$ and $\boldsymbol{z}$ having Gaussian distributions $p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}_x, \boldsymbol{\Sigma}_x)$ and $p(\boldsymbol{z}) = \mathcal{N}(\boldsymbol{z}|\boldsymbol{\mu}_z, \boldsymbol{\Sigma}_z)$ respectively, together with their sum $\boldsymbol{y} = \boldsymbol{x} + \boldsymbol{z}$. Use the equations (2.109), (2.110) from the PRML-Bishop textbook, i.e.:

$$\mathrm{E}[\boldsymbol{y}] = \mathbf{A}\boldsymbol{\mu} + \boldsymbol{b} \tag{5}$$

$$\mathrm{cov}[\boldsymbol{y}] = \mathbf{L}^{-1} + \mathbf{A}\boldsymbol{\Lambda}^{-1}\mathbf{A}^T \tag{6}$$

to find an expression for the marginal distribution $p(\boldsymbol{y})$ by considering the linear-Gaussian model comprising the product of the marginal distribution $p(\boldsymbol{x})$ and the conditional distribution $p(\boldsymbol{y}|\boldsymbol{x})$.

Exercise 5 [3 points]

Consider a joint distribution $p(\boldsymbol{x}, \boldsymbol{y})$ defined by the marginal and conditional distributions given by:

$$p(\boldsymbol{x}) = \mathcal{N}(\boldsymbol{x}|\boldsymbol{\mu}, \boldsymbol{\Lambda}^{-1}) \tag{7}$$

$$p(\boldsymbol{y}|\boldsymbol{x}) = \mathcal{N}(\boldsymbol{y}|\mathbf{A}\boldsymbol{x} + \boldsymbol{b}, \mathbf{L}^{-1}) \tag{8}$$

where $\boldsymbol{\mu}$, $\mathbf{A}$ and $\boldsymbol{b}$ are parameters governing the means, and $\boldsymbol{\Lambda}$ and $\mathbf{L}$ are precision matrices. By examining the quadratic form in the exponent of the joint distribution, and using the technique of "completing the square", find expressions for the mean and covariance of the marginal distribution $p(\boldsymbol{y})$ in which the variable $\boldsymbol{x}$ has been integrated out. Hint: Make use of the Woodbury inversion formula: $(A + BCD)^{-1} = A^{-1} - A^{-1}B(C^{-1} + DA^{-1}B)^{-1}DA^{-1}$

Exercise 6 [3 points]

Using different random seeds generate two data sets, X (training set) and X_1 (test set), each consisting of $N = 1000$ 3-dimensional vectors that stem from three *equiprobable* classes, ω_1, ω_2 and ω_3. The classes are modelled by Gaussian distributions with means $\boldsymbol{m}_1 = [0, 0, 0]^T$ and $\boldsymbol{m}_2 = [1, 2, 2]^T$, and $\boldsymbol{m}_3 = [3, 3, 4]^T$ respectively. Their covariance matrices are:

$$\boldsymbol{S}_1 = \begin{bmatrix} 0.8 & 0.2 & 0.1 \\ 0.2 & 0.8 & 0.2 \\ 0.1 & 0.2 & 0.8 \end{bmatrix}, \quad \boldsymbol{S}_2 = \begin{bmatrix} 0.6 & 0.01 & 0.01 \\ 0.01 & 0.8 & 0.01 \\ 0.01 & 0.01 & 0.6 \end{bmatrix}, \quad \boldsymbol{S}_3 = \begin{bmatrix} 0.6 & 0.1 & 0.1 \\ 0.1 & 0.6 & 0.1 \\ 0.1 & 0.1 & 0.6 \end{bmatrix}$$

i. Using X, compute the maximum likelihood estimates of the mean values and the covariance matrices of the distributions of the three classes.[0.7pt]

ii. Use the Mahalanobis distance classifier to classify the points of X_1 based on the ML estimates computed before. [0.7pt]

iii. Use the Bayesian classifier to classify the points of X_1 based on the ML estimates computed before. [0.6pt]

iv. For each case, compute the error probability and compare the results [0.6pt]