Statistics 2 - Problem Sheet 1

Dom Hutchinson

January 28, 2020

Question - 1.

Question 1 a)
If $\mathbf{Y}$ and $\mathbf{X}$ are random vectors st $\mathbf{Y} = C\mathbf{X}$ where C is a matrix of fixed coefficients, show that if Σ_X and Σ_Y are the covariance matrices for $\mathbf{X}$ and $\mathbf{Y}$ respecitvely then

$$\Sigma_y = C\Sigma_X C^T$$

Answer 1 a)

$$\begin{array}{rcl}
\mu_y & = & \mathbb{E}(Y) \\
& = & \mathbb{E}(CX) \\
& = & C\mathbb{E}(X) \\
& = & C\mu_X \\
\Sigma_y & = & \mathbb{E}[(Y-\mu_Y)(Y-\mu_Y)^T] \\
& = & \mathbb{E}[(CX - C\mu_X)(CX - C\mu_X)^T] \\
& = & \mathbb{E}[C(X-\mu_X)(X-\mu_X)^T C^T] \\
& = & C\mathbb{E}[(X-\mu_X)(X-\mu_X)^T]C^T \\
& = & C\Sigma_X C^T
\end{array}$$

Question 1 b)
Consider a multivariate normal random vector $\mathbf{X} \sim \text{Normal}(\boldsymbol{\mu}_X, \Sigma_X)$ and suppose that the covariance matrix can be decomposed $\Sigma_X = CC^T$ (THis can always be done for a full rank covariance matrix using a Choleski decomposition). Show that $\Sigma_X^{-1} = C^{-T}C^{-1}$ and that $\mathbf{Y} = C^{-1}(\mathbf{X} - \boldsymbol{\mu}_X) \sim \text{Normal}(\mathbf{0}, I)$.

Answer 1 b)

$$\begin{array}{rrcl}
& \Sigma_X\Sigma_X^{-1} & = & I \\
\implies & CC^T\Sigma_X^{-1} & = & 1 \\
\implies & C^T V_X^{-1} & = & C^{-1} \\
\implies & V_X^{-1} & = & C^{-T}C^{-1}
\end{array}$$

$$\begin{array}{rcl}
Y & = & C^{-1}(X-\mu_X) \\
\mathbb{P}(Y=x) & = & \mathbb{P}(C^{-1}(X-\mu_X) = x) \\
& & TODO
\end{array}$$

Question 1 c)
Assuming that $\mathbf{X} \sim \text{Normal}(\boldsymbol{\mu}_X, \Sigma_X)$ show that

$$(\mathbf{X} - \boldsymbol{\mu}_X)^T\Sigma_X^{-1}(\mathbf{X} - \boldsymbol{\mu}_X) = \mathbf{Y}^T\mathbf{Y} \text{ where } \mathbf{Y} \sim \text{Normal}(\mathbf{0}, I)$$

Answer 1 c)
TODO

Question 1 d)
If $Z_i \overset{\text{iid}}{\sim} \text{Normal}(0,1)$ random variables then

$$\sum_{i=1}^{n} Z_i^2 \sim \chi_n^2$$

What is the distribution of

$$(\mathbf{X} - \boldsymbol{\mu}_X)^T \Sigma_X^{-1} (\mathbf{X} - \boldsymbol{\mu}_X)$$

if $\mathbf{X} \sim \text{Normal}(\boldsymbol{\mu}_X, \Sigma_X)$?

Answer 1 d)
TODO

Question - 2.

Question 2 a)
Define $\mathbf{y} = \begin{pmatrix} 1 \\ -3 \end{pmatrix}$ and $B = \begin{pmatrix} -1 & 2 & -1 \\ 2 & -3 & 0 \end{pmatrix}$.
Find $B^T\mathbf{y}$.

$$\begin{pmatrix} -1 & 2 & -1 \\ 2 & -3 & 0 \end{pmatrix}^T \begin{pmatrix} 1 \\ -3 \end{pmatrix} = \begin{pmatrix} -1 & 2 \\ 2 & -3 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 1 \\ -3 \end{pmatrix} = \begin{pmatrix} -1-7 \\ 2+9 \\ -1+0 \end{pmatrix} = \begin{pmatrix} -7 \\ 11 \\ -1 \end{pmatrix}$$

Answer 2 a)

Question 2 b)
Let A be a full rank 3×3 matrix and B be a full rank 5×3 matrix.
State the dimensions of the following, if they exist. For those that do not exist, state why in a single sentence.

i) $A^{-1}B^T$
$A^{-1} \in \mathbb{R}(3 \times 3),\ B^T \in \mathbb{R}(3 \times 5) \implies A^{-1}B^T \in \mathbb{R}(3 \times 5)$

ii) $A^{-1}B$
$A^{-1} \in \mathbb{R}(3 \times 3),\ B \in \mathbb{R}(5 \times 3) \implies$ these matrices are incompatible for multiplication.

iii) $B^{-1}A$
$B^{-1} \in \mathbb{R}(3 \times 5),\ A \in \mathbb{R}(3 \times 3) \implies$ these matrices are incompatible for multiplication.

iv) BA
$B \in \mathbb{R}(5 \times 3),\ A \in \mathbb{R}(3 \times 3) \implies BA \in \mathbb{R}(5 \times 3)$

v) $B^{-1}A^T$
$B^{-1} \in \mathbb{R}(3 \times 5),\ A^T \in \mathbb{R}(3 \times 3) \implies$ these matrices are incompatible for multiplication.

vi) BA^{-1}
$B \in \mathbb{R}(5 \times 3),\ A^{-1} \in \mathbb{R}(3 \times 3) \implies BA^{-1} \in \mathbb{R}(5 \times 3)$.

vii) $(BA)^{-1}$
We know that $BA \in \mathbb{R}(5 \times 3) \implies (BA)^{-1} \in \mathbb{R}(3 \times 5)$.

viii) B^TA
$B^T \in \mathbb{R}(3 \times 5),\ A \in \mathbb{R}(3 \times 3) \implies$ these matrices are incompatible for multiplication.

ix) $B+A$
$B\in\mathbb{R}(3\times 5),\ A\in\mathbb{R}(3\times 3)$. Matrices must have the exact same dimensions in order to be added together, thus this in an illegal equation.

x) $B+A^T$
$B\in\mathbb{R}(3\times 5),\ A\in\mathbb{R}(3\times 3)$. Matrices must have the exact same dimensions in order to be added together, thus this in an illegal equation.

Question - 3.

The *Exponential Distribution* is often a reasonable model of the times between random events. Suppose then, that $x_1,\dots,x_n$ are observations of times between hardware faults on a computer network, and it is reasonable to treat the faults as independent. To plan for fault tolerance the network managers need a reasonable model for the fault occurence rate. The *pdf* of an *Exponential Distribution* is

$$f(x)=\mathbb{1}\{x\geq 0\}\lambda e^{-\lambda x}$$

where λ is a positive parameter.
The variance of an exponential random variable is λ^{-2}.

Question 3 a) - Let $X\sim\text{Exponential}(\lambda)$. Find $\mathbb{E}(X)$.

Answer 3 a)

$$\begin{array}{rcl}\mathbb{E}(X) & = & \displaystyle\int_{-\infty}^{\infty} f_X(t)dt\\ & = & \displaystyle\int_{-\infty}^{\infty} \mathbb{1}\{t\geq 0\}\lambda e^{-\lambda t}dt\\ & = & \displaystyle\int_{0}^{\infty} \lambda e^{-\lambda t}dt\\ & = & \left[-\dfrac{1}{\lambda}e^{-\lambda t}\right]_0^{\infty}\\ & = & [0]-\left[-\dfrac{1}{\lambda}\right]\\ & = & \dfrac{1}{\lambda}\end{array}$$

Question 3 b) - Hence, suggest an estimator, $\hat{\lambda}$, for λ.

Answer 3 b)

$$\hat{\lambda}=1/\bar{x}\text{ where }\bar{x}=\frac{1}{n}\sum_{i=1}^n x_i$$

Question 3 c) - What is the variance of $\hat{\lambda}^{-1}$?
Answer 3 c)

$$\begin{array}{rcl}\text{var}\left(\hat{\lambda}^{-1}\right) & = & \text{var}(\bar{X})\\ & = & \text{var}\left(\dfrac{1}{n}\displaystyle\sum_{i=1}^n X_i\right)\\ & = & \dfrac{n}{n^2}\text{var}(X_1)\\ & = & \dfrac{1}{n\lambda^2}\end{array}$$

Question 3 d)
Let $\bar{x} = \frac{1}{n}\sum x_i$.
Find a first order Taylor expansion of $\hat{\lambda}$ about $\mathbb{E}(\bar{x})$, considering $\hat{\lambda}$ as a function of $\bar{x}$.

Answer 3 d)
TODO

Question 3 e)
Hence find an approximation for the variance of $\hat{\lambda}$¡ in terms of n and $\bar{x}$. This use of Taylor expansions to computer approximate variances via linearisation is known as the Δ-*method* in statistics.

Answer 3 e)
TODO

Question - 4.

Consider again the setup from the previous question, but now taking a Bayesian approach. This means taht we need to augment our model with a prior distribution for the parameter, $\lambda \sim \Gamma(\alpha, \theta)$. So the prior *pdf* of λ is

$$f(\lambda) = \frac{\lambda^{\alpha-1}e^{-\lambda/\theta}}{\theta^\alpha\Gamma(\alpha)}$$

with $\mathbb{E}(\lambda) = \alpha\theta$ and $\text{var}(\lambda) = \alpha\theta^2$.

Question 4 a) - Write down the *pdf* for the joint distribution of the data $x_1, x_2, \dots$ given λ.

Answer 4 a)

$$\begin{aligned} f_n(\mathbf{x};\lambda) &= \prod_{i=1}^n f(x_i;\lambda) \\ &= \prod_{i=1}^n \mathbb{1}\{x_i \geq 0\}\lambda e^{-\lambda x_i} \\ &= \mathbb{1}\{\text{all } x \geq 0\}\lambda^n e^{-\lambda n\bar{x}} \end{aligned}$$

Question 4 b) - By considering the joint distribution of λ and $\mathbf{x}$, indentify the posterior distribution of λ given $\mathbf{x}$.

Answer 4 b)

$$\begin{aligned} f(\lambda;\mathbf{x}) &\propto f(\mathbf{x};\lambda)f(\lambda) \\ &= \lambda^n e^{-\lambda n\bar{x}} \cdot \frac{\lambda^{\alpha-1}e^{-\lambda/\theta}}{\theta^\alpha\Gamma(\alpha)} \\ &\propto \lambda^{n+\alpha-1}e^{-\lambda(n\bar{x}+1/\theta)} \\ &= \lambda^{n+\alpha-1}e^{-\lambda\frac{n\bar{x}\theta+1}{\theta}} \\ &\sim \Gamma\left(n+\alpha, \frac{\theta}{n\bar{x}\theta+1}\right) \end{aligned}$$

Question 4 c) - What are the posterior expection and variance of λ?

Answer 4 c)

$$\mathbb{E}[\lambda; \mathbf{x}] = \frac{\theta(n+\alpha-1)}{n\bar{x}\theta+1} \quad \text{var}(\lambda; \mathbf{x}) = \frac{\theta^2(n+\alpha-1)}{(n\bar{x}\theta+1)^2}$$

Question 4 d)
Consider the situation in which $n \to \infty$.
What happens to the Bayesian and frequentist inferences aboute λ in this case?

Answer 4 d)

$$\mathbb{E}[\lambda; \mathbf{x}] \stackrel{n\to\infty}{\longrightarrow} \frac{1}{\bar{x}} = \hat{\lambda}_{\text{Frequentist}}$$