NATIONAL UNIVERSITY OF SCIENCE AND TECHNOLOGY

SMA 4252

FACULTY OF APPLIED SCIENCES

DEPARTMENT OF APPLIED MATHEMATICS

SMA 4252 MULTIVARIATE ANALYSIS

DECEMBER 2002

Time : 3 hours

Candidates should attempt **ALL** questions from Section A and **THREE** questions from Sections B.

SECTION A: Answer ALL questions in this section [55].

A1. Let the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 2 \\ 2 & 1 & 2 \\ 2 & 2 & 1 \end{bmatrix}$$

(a) Find the eigenvalues and the corresponding normalised eigenvectors for the matrix **A**. [Hint: $(x-2)$ is a factor of the polynomial $x^3 - 12x + 16$. **A** is not positive definite.] [10]

(b) Show that the determinant of **A** is equal to $\prod_{i=1}^{3} \lambda_i$ and that $Trace(\mathbf{A}) = \sum_{i=1}^{3} \lambda_i$. [4]

(c) Represent **A** by its spectral decomposition. [4]

(d) Represent $\mathbf{A}^{-1}$ by its spectral decomposition. [5]

A2. Let $\mathbf{X} = [X_1, X_2, X_3, X_4]'$ be $N_4(\mu, \Sigma)$ with $\mu' = (1, 2, -2, -1)$ and $\mathbf{\Sigma} = \begin{bmatrix} 1 & 1 & 2 & 0 \\ 1 & 3 & 1 & 0 \\ 2 & 1 & 4 & 1 \\ 0 & 0 & 1 & 2 \end{bmatrix}$

(a) Find the distribution of $Z_1 = X_1$ [1]

(b) Find the distribution of $Z_2 = -2X_1 - X_2 + X_3 + 2X_4$ [3]

(c) Find the distribution of $\mathbf{Z} = \begin{bmatrix} Z_1 \\ Z_2 \end{bmatrix}$ [3]

(d) Find a 3×1 vector $\mathbf{a}$ such that X_2 and $X_2 - \mathbf{a}\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$ are independent. [3]

A3. The following measurements are obtained on the variables X_1, X_2 and X_3

Variable	Observation				
	1	2	3	4	5
X_1	9	2	6	5	8
X_2	12	8	6	4	10
X_3	3	4	0	2	1

(a) Calculate the sample variance-covariance matrix $\mathbf{S}$ for the vector $\mathbf{X} = \begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix}$ [5]

(b) Calculate the correlation matrix using $\rho = \mathbf{D^{-1}SD^{-1}}$. [4]

A4. Let $\mathbf{X_1}, \mathbf{X_2}, \ldots, \mathbf{X_{75}}$ be a random sample of a multivariate normal population, with $\mathbf{X} \sim N(\mu_\mathbf{p}, \Sigma_\mathbf{pxp})$.

What is the approximate distribution of

(a) $\bar{\mathbf{X}}$, [2]

(b) $n(\bar{\mathbf{X}} - \mu)^\mathbf{T}\mathbf{S}^{-1}(\bar{\mathbf{X}} - \mu)$. [3]

A5. Observations for two variables are made on five individuals as follows;

Individual	Variable	
	1	2
1	1	2
2	1	1
3	6	3
4	8	2
5	8	0

(a) Construct a dissimilarity matrix using Euclidean distance. [5]

(b) Construct a dendogram using the nearest neighbour method on the dissimilarity matrix. [3]

SECTION B: Answer ANY THREE questions in this section [45].

B6. Suppose that the random variables $X_1, X_2,$ and X_3 have covariance matrix

$$\Sigma = \begin{bmatrix} 6 & 0 & -1 \\ 0 & 5 & 0 \\ -1 & 0 & 2 \end{bmatrix}$$

The eigenvalue-eigenvector pairs are:

$$\begin{aligned} \lambda_1 &= 6.2361 & \mathbf{e}_1' &= [0.9732,\ 0.0000,\ -0.2298] \\ \lambda_2 &= 5.0 & \mathbf{e}_1' &= [0.0000,\ 1.0000,\ 0.0000] \\ \lambda_3 &= 0.7596 & \mathbf{e}_1' &= [0.2298,\ 0.0000,\ 0.9732] \end{aligned}$$

(a) Obtain the principal components (Y_1, Y_2, Y_3) of Σ. [6]

(b) Verify that $Var(Y_1) = \lambda_1$ and that $Cov(Y_1, Y_2) = 0$. [2]

(c) Calculate the proportion of the total variation accounted for by each principal component. [3]

(d) Calculate the correlations between (Y_1 and X_1), (Y_1 and X_2), (Y_2 and X_1) and (Y_2 and X_2). [4]

B7. Observations on two responses are collected for three treatments. The observation vectors $\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$ are

Treatment 1 $\begin{bmatrix} 6 \\ 7 \end{bmatrix}, \begin{bmatrix} 5 \\ 9 \end{bmatrix}, \begin{bmatrix} 8 \\ 6 \end{bmatrix}, \begin{bmatrix} 4 \\ 9 \end{bmatrix}, \begin{bmatrix} 7 \\ 9 \end{bmatrix}$

Treatment 2 $\begin{bmatrix} 3 \\ 3 \end{bmatrix}, \begin{bmatrix} 1 \\ 6 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

Treatment 3 $\begin{bmatrix} 2 \\ 3 \end{bmatrix}, \begin{bmatrix} 5 \\ 1 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \end{bmatrix}$

The data above can be presented by the model $X_{ij} = \mu + \tau_i + e_{ij}$ where
i = the treatment number and $X_{ij} = j^{th}$vector observation under treatment i,
μ = overall mean, $\tau_i = i^{th}$treatment effect and e_{ij} is random error

(a) Construct a MANOVA table for the data. [10]

(b) Evaluate Wilks lambda from the MANOVA table. [5]

B8. Construct a 95% confidence ellipsoid for the population mean vector $\begin{bmatrix} \mu_1 \\ \mu_2 \end{bmatrix}$ if the following data were obtained for the random variables X_1 and X_2:

x_1 :	2	8	6	8
x_2 :	12	9	9	10

[15]

B9. The random vector **X** has covariance matrix

$$\Sigma = \begin{bmatrix} 3 & 1 & 1 \\ 1 & 3 & 1 \\ 1 & 1 & 5 \end{bmatrix}$$

(a) Find the unrotated factor loadings for the factor analysis model

$$X_i = \lambda_{i1} F_1 + \lambda_{i2} F_2 + e_i$$

where the F_j's are the communalities, the λ_{ij}'s are the factor loadings and the e_i's are the specific factors. [12]

(b) What are the assumptions you make to come up with a factor analysis model. [3]

END OF QUESTION PAPER