ECE3093 Optimization, Estimation and Numerical Methods. Stochastic Part

Assignment 2 (2013) Part B

This is an individual and not group assignment. This assignment will be released in three separate parts (through Moodle ECE3093 folder Assignments Part2 2013):

Part A Fantasia heart beat data analysis

Part B Gaussian vector simulation

Part C ARMA time and frequency domain simulation

Your work on Part A, Part B and Part C is to be written up and submitted as three separate reports; one for each part. It expected that you will write a MATLAB m-file for each part of this assignment. Advice on how to write these is given to make this as straight forward as possible. If you require further assistance, please seek help as soon as possible.

At Level 3 you are expected to be mature enough to understand that your assignments must be your own work. Accordingly, it is expected that your MATLAB scripts and their output will be your own, original work. All students who submit identical MATLAB scripts or identical output (eg plots of simulations) will all received a mark of zero for these parts of the assessment.

Your reports are to be submitted through the assignment box for Mr Alan Couchman (ground floor, Building 28) at Clayton Campus (Clayton students) or to Dr Nader Kamrani at Sunway Campus (Sunway students).

Deadlines for reports are:

Part A WEEK9, before 6pm Wednesday 8th May 2013

Part B WEEK11, before 6pm Monday 20th May 2013

Part C WEEK12, before 6pm Friday 31st May 2013

NB Include a printout of your m-files for parts A, B and C in the body or as an appendix to your reports.

Part B Gaussian vector simulation

In this assignment 2 Part B, you need to obtain your personal values for U, V, Z and A from the appropriate Clayton or Sunway version of the ECE3093 Student ID based numbers Z U V A a b DFA file in the Assignments 2013 Part 2 folder.

Let $\mathbf{W} = (W_1, \ldots, W_4) \sim N_4(\mathbf{m}, \mathbf{\Sigma})$ be a Gaussian random vector where $\mathbf{\Sigma}$ is the matrix,

$$\mathbf{\Sigma} = Cov(W_1, W_2, W_3, W_4) = \begin{bmatrix} 9.25 & -1.25 & 0.64 & 0.51 \\ -1.25 & 4.75 & -0.84 & 0.81 \\ 0.64 & -0.84 & Z+1 & -0.75 \\ 0.51 & 0.81 & -0.75 & Z+2 \end{bmatrix}$$

where $\Sigma_{3,3}$ is $\mathrm{Cov}(W_3, W_3) = Z + 1$ and $\Sigma_{4,4}$ is $\mathrm{Cov}(W_4, W_4) = Z + 2$.
$\mathbf{m}$ is the vector,

$$\mathbf{m} = \begin{bmatrix} U+3V & 2U+V & U & V \end{bmatrix}^{\mathbf{T}}$$

In question (1) part (i) and (ii), (A,B,C,D) is a cyclic permutation of the digits 1,2,3,4, starting at the position given by your A. Your (A,B,C,D) is either (2,3,4,1) or (3,4,1,2) or (4,1,2,3). Note: You can convert A to int 16 A by cast it using int16(A).

Write a sequentially organised MATLAB script of your own that does the following calculations and generates the necessary output to answer the questions asked in the following three parts (1), (2) and (3). Include answers to the questions in your report.

(1) (i) Careful construct the $\mathbf{C} = \mathrm{Cov}(W_A, W_B, W_C, W_D)$ for your (A,B,C,D) from $\mathbf{\Sigma}$.

$$\mathbf{C} = \begin{bmatrix} \mathrm{cov}(W_A, W_A) & \mathrm{cov}(W_A, W_B) & \mathrm{cov}(W_A, W_C) & \mathrm{cov}(W_A, W_D) \\ \mathrm{cov}(W_B, W_A) & \mathrm{cov}(W_B, W_B) & \mathrm{cov}(W_B, W_C) & \mathrm{cov}(W_B, W_D) \\ \mathrm{cov}(W_C, W_A) & \mathrm{cov}(W_C, W_B) & \mathrm{cov}(W_C, W_C) & \mathrm{cov}(W_C, W_D) \\ \mathrm{cov}(W_D, W_A) & \mathrm{cov}(W_D, W_B) & \mathrm{cov}(W_D, W_C) & \mathrm{cov}(W_D, W_D) \end{bmatrix}$$

Check that your $\mathbf{C}$ is (a) symmetric, and using MATLAB's eig() function to (b) show that your $\mathbf{C}$ is positive definite.

(ii) Careful construct the $\mathbf{b} = E(W_A, W_B, W_C, W_D)^T$ for your (A,B,C,D) from $\mathbf{m}$.

(2) Let $X = W_A \in \mathbb{R}$ and $\mathbf{Y} = (W_B, W_C, W_D)^T \in \mathbb{R}^3$. $E\mathbf{X} \in \mathbb{R}^1$ and $E\mathbf{Y} \in \mathbb{R}^3$

(i) Use MATLAB to compute the matrix $M := \mathrm{Cov}(X, \mathbf{Y})(\mathrm{Cov}(\mathbf{Y}))^{-1} \in \mathbb{R}^{1\times 3}$ and the variance of the innovation (prediction error) $\mathrm{var}(X - P(X|\mathbf{Y}))$.

(ii) State the distribution of the innovation, $X - P(X|\mathbf{Y}) \sim N(\mu, \sigma^2)$. Calculate $\mu = E[X - P(X|\mathbf{Y})]$ and $\sigma^2 = \text{var}(X - P(X|\mathbf{Y}))$.

(iii) Suppose that $\mathbf{y} = (U + V + A, U - 2V, UV + A)^T$ is observed (to convert A to double A use double(A)). Use the results of (i) and MATLAB to give the best linear prediction of X. State the distribution of $X \sim N_1(b_X, C_X)$.

(3) Let $\mathbf{X} = (W_A, W_B)^T \in \mathbb{R}^2$ and $\mathbf{Y} = (W_C, W_D)^T \in \mathbb{R}^2$. $E\mathbf{X} \in \mathbb{R}^2$ and $E\mathbf{Y} \in \mathbb{R}^2$

(i) Explain, using the results of the affine transport theorem, how the affine transport, $\mathbf{X} = \mathbf{a} + Q\mathbf{Y}$, where $C = QDQ^T$, maps $X \sim N_2(\mathbf{a}, C)$ into $Y \sim N_2(\mathbf{0}, D)$, where $C = QDQ^T$ is the orthogonal decomposition of covariance matrix C.

(ii) Generate a scatter plot, (W_A, W_B), by simulating N=3000 points of $\mathbf{X} \sim N_2(E\mathbf{X}, \text{Cov}(\mathbf{X}))$.

[Advice: Use $\mathbf{X} = E\mathbf{X} + Q\sqrt{D}\mathbf{Z}$, where $\mathbf{Z}$ is a standard normal vector and $\text{Cov}(\mathbf{X}) = QDQ^T$; use $[Q, D] = \textbf{eig}(\text{Cov}(\mathbf{X}))$ to obtain the decomposition of $\text{Cov}(\mathbf{X})$ and use `randn(2,N)` to simulate components of $\mathbf{Z}$. To calculate $\mathbf{X} = E\mathbf{X}+Q\sqrt{D}\mathbf{Z}$ use X=**diag**(b(1:2))***ones**(2,N)+Q***sqrt**(D)***randn**(2,N)].

(iii) Use MATLABs `mvnrnd()` to generate a plot to confirm your results to part (ii). In what way do the results to (ii) and (iii) differ?

(iv) Generate a contour plot of the density of $\mathbf{X}$ using `contour()` and `mvnpdf()`. How is the covariance matrix responsible for the shape and orientation of the contour plot? In your answer consider the role of the eigenvalues and eigenvectors $C = QDQ^T$.

[Advice: Use [x,y]=**meshgrid**(xmin:xstep:xmax,ymin:ystep:ymax);[nx,ny]=**size**(x); mu=b(1:2)';z=**mvnpdf**([x(:) y(:)],mu,Cov($\mathbf{X}$));z=**reshape**(z,nx,ny);**contour**(x,y,z).]

(v) Use MATLAB to compute the matrix $\mathbf{M} = \text{Cov}(\mathbf{X}, \mathbf{Y})(\text{Cov}(\mathbf{Y}))^{-1} \in \mathbb{R}^{2\times 2}$.

(vi) Suppose that $\mathbf{y} = (2V - U, U + V)^T$ is observed. Use $\mathbf{M}$ and MATLAB to give the best linear prediction $P(\mathbf{X}|\mathbf{Y} = \mathbf{y})$. Comment on your result.

(vii) Use $\mathbf{M}$ to calculate the prediction error, $\text{Cov}(X - P(X|\mathbf{Y})) \in \mathbb{R}^{2\times 2}$, hence state the distribution of the innovation, $X - P(X|\mathbf{Y}) \sim N(\mathbf{a}, \mathbf{K})$, where, $\mathbf{a} = E[\mathbf{X} - P(\mathbf{X}|\mathbf{Y})]$ and $\mathbf{K} = \text{Cov}(\mathbf{X} - P(\mathbf{X}|\mathbf{Y}))$.

(viii) Simulate the random vector, $P(\mathbf{X}|\mathbf{Y})$, by simulating N=3000 points of $\mathbf{Y} = (\mathbf{W_C}, \mathbf{W_D})$ as input to $P(\mathbf{X}|\mathbf{Y})$, to calculate, $P(\mathbf{X}|\mathbf{Y}) = \mathbf{EX} + \text{Cov}(\underline{f}\mathbf{X}, \mathbf{Y})(\text{Cov}(\mathbf{Y}))^{-\mathbf{1}}(\mathbf{Y} - \mathbf{EY})$.

[Advice: Use $\mathbf{Y} - E\mathbf{Y} = Q\sqrt{D}\mathbf{Z}$, where $\mathbf{Z}$ is a standard normal vector and $\text{Cov}(\mathbf{Y}) = QDQ^T$; use $[Q, D] = \textbf{eig}(\text{Cov}(\mathbf{Y}))$ to obtain the decomposition of $\text{Cov}(\mathbf{Y})$ and use `randn(2,N)` to simulate components of $\mathbf{Z}$. To calculate

$\mathbf{Y} - E\mathbf{Y} = Q\sqrt{D}\mathbf{Z}$ use $\mathbf{Y} - E\mathbf{Y}$=Q***sqrt**(D)***randn**(2,N)) and to calculate $E\mathbf{X}$ use **diag**(b(1:2))***ones**(2,N),and then calculate, $P(\mathbf{X}|\mathbf{Y})$].

(ix) Overlay scatterplot of (viii) $\mathbf{Y}$ and (viii) $P(\mathbf{X}|\mathbf{Y})$ on your contour plot from part (iv), and comment on this overlay. Include a legend.

END OF PART B