Boston University

Department of Electrical and Computer Engineering

EC505 STOCHASTIC PROCESSES

Problem Set No. 2 Solutions

Fall 2016

Issued: Wednesday, Sept. 14, 2016 **Due:** Monday, Sept. 26, 2016

Problem 2.1

Consider the random variables X and Y whose joint density function is given by $p_{X,Y}(x,y)$. For each of the possible choices of the joint density function given below determine i) whether X and Y are uncorrelated, ii) whether X and Y are independent, and iii) specify the corresponding covariance matrix:

$$\Sigma_{XY} = \begin{bmatrix} \sigma_{XX} & \sigma_{XY} \\ \sigma_{XY} & \sigma_{YY} \end{bmatrix}.$$

$$p_{X,Y}(x,y) = \begin{cases} 2 & \begin{cases} x, y \geq 0 \ \ \& \\ x+y \leq 1 \end{cases} \\ 0 & \text{otherwise} \end{cases} \qquad p_{X,Y}(x,y) = \begin{cases} 1/2 & \begin{cases} |x+y| \leq 1 \ \ \& \\ |x-y| \leq 1 \end{cases} \\ 0 & \text{otherwise} \end{cases} \qquad p_{X,Y}(x,y) = \begin{cases} 1 & 0 \leq x, y \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

(a) (b) (c)

Solution:

(a) First we'll find the entries of the covariance matrix. To start, find the marginal densities of X and Y:

$$\begin{aligned} p_X(x) &= \int_{-\infty}^{\infty} p_{X,Y}(x,y)\,dy = \int_{y=0}^{y=1-x} 2\,dy = 2(1-x) \ \text{ for } 0 \leq x \leq 1 \\ p_Y(y) &= \int_{-\infty}^{\infty} p_{X,Y}(x,y)\,dx = \int_{x=0}^{x=1-y} 2\,dx = 2(1-y) \ \text{ for } 0 \leq y \leq 1 \end{aligned}$$

We could have obtained the second result by symmetry. Now we can find the moments:

$$\begin{aligned} E(X) &= \int_{-\infty}^{\infty} x p_X(x)\,dx = \int_0^1 2x(1-x)\,dx = \frac{1}{3} \\ E(X^2) &= \int_{-\infty}^{\infty} x^2 p_X(x)\,dx = \int_0^1 2x^2(1-x)\,dx = \frac{1}{6} \\ \sigma_{XX} &= E(X^2) - [E(X)]^2 = \frac{1}{18} \end{aligned}$$

By symmetry we get:

$$E(Y) = \frac{1}{3},\ E(Y^2) = \frac{1}{6},\ \sigma_{YY} = \frac{1}{18}$$

Finally:

$$\sigma_{XY} = E\left[(X-m_X)(Y-m_Y)\right] = \int_{y=0}^{1}\int_{x=0}^{1-y} (xy)2\,dx\,dy - m_X m_Y = \int_{y=0}^{1} y(1-y)^2\,dy - \frac{1}{9} = -\frac{1}{36}$$

Thus the covariance matrix for density (a) is given by:

$$\Sigma_{XY} = \begin{bmatrix} \frac{1}{18} & -\frac{1}{36} \\ -\frac{1}{36} & \frac{1}{18} \end{bmatrix}.$$

In summary, i) X and Y are clearly correlated, since $\sigma_{XY} \neq 0$, ii) X and Y are clearly *not* independent. This can be seen from the fact that $p_X(x)p_Y(y) \neq p_{X,Y}(x,y)$, from the fact that knowledge of Y affects the remaining uncertainty in X (i.e. if we know $Y = 1/2$ then X cannot be $2/3$), or from the fact that the two random variables are correlated. iii) the covariance matrix is given above.

(b) Once more, the marginal densities of X and Y are given by:

$$\begin{aligned} p_X(x) &= \int_{-\infty}^{\infty} p_{X,Y}(x,y)\,dy = \begin{cases} \int_{y=-1-x}^{y=1+x} 1/2\,dy = 1+x, & -1 \leq X \leq 0 \\ \int_{y=-1+x}^{y=1-x} 1/2\,dy = 1-x, & 0 \leq X \leq 1 \end{cases} \\ p_Y(y) &= \int_{-\infty}^{\infty} p_{X,Y}(x,y)\,dx = \begin{cases} \int_{x=-1-y}^{x=1+y} 1/2\,dx = 1+y, & -1 \leq Y \leq 0 \\ \int_{x=-1+y}^{x=1-y} 1/2\,dx = 1-y, & 0 \leq Y \leq 1 \end{cases} \end{aligned}$$

where again the second expression could have been obtained from symmetry. Now find the moments:

$$\begin{aligned} E(X) &= \int_{-1}^{1} x p_X(x)\,dx = 0 \\ E(X^2) &= 2\int_{0}^{1} x^2(1-x)\,dx = \frac{1}{6} \\ \sigma_{XX} &= E(X^2) - [E(X)]^2 = \frac{1}{6} \end{aligned}$$

By symmetry we get:

$$E(Y) = 0,\;\; E(Y^2) = \frac{1}{6},\;\; \sigma_{YY} = \frac{1}{6}$$

Finally:

$$\begin{aligned} \sigma_{XY} &= E\left[(X - m_X)(Y - m_Y)\right] = \int_{x=-1}^{0}\int_{y=-1-x}^{y=1+x} (xy)1/2\,dy\,dx + \int_{x=0}^{1}\int_{y=-1+x}^{y=1-x} (xy)1/2\,dy\,dx - m_X m_Y \\ &= \int_{x=-1}^{0} x/2 \int_{y=-(1+x)}^{y=1+x} y\,dy\,dx + \int_{x=0}^{1} x/2 \int_{y=-(1-x)}^{y=1-x} y\,dy\,dx = 0 + 0 \\ &= 0 \end{aligned}$$

Thus the covariance matrix for density (b) is given by:

$$\Sigma_{XY} = \begin{bmatrix} \frac{1}{6} & 0 \\ 0 & \frac{1}{6} \end{bmatrix}.$$

In summary, i) X and Y are clearly uncorrelated, since $\sigma_{XY} = 0$, ii) Again X and Y are clearly *not* independent. This can be seen from the fact that $p_X(x)p_Y(y) \neq p_{X,Y}(x,y)$, from the fact that knowledge of Y affects the remaining uncertainty in X (i.e. if we know $Y = 1/2$ then X cannot be $2/3$). Recall that being independent is stronger than being uncorrelated. iii) the covariance matrix is given above.

(c) Again, find the marginal densities of X and Y:

$$\begin{aligned} p_X(x) &= \int_{-\infty}^{\infty} p_{X,Y}(x,y)\,dy = \int_{y=0}^{y=1} 1\,dy = 1 \quad 0 \leq X \leq 1 \\ p_Y(y) &= \int_{-\infty}^{\infty} p_{X,Y}(x,y)\,dx = \int_{x=0}^{x=1} 1\,dx = 1 \quad 0 \leq Y \leq 1 \end{aligned}$$

Find the moments:

$$
\begin{aligned}
E(X) &= \int_0^1 x p_X(x)\, dx = \frac{1}{2} \\
E(X^2) &= \int_0^1 x^2 1\, dx = \frac{1}{3} \\
\sigma_{XX} &= E(X^2) - [E(X)]^2 = \frac{1}{12}
\end{aligned}
$$

By symmetry we get:

$$
E(Y) = \frac{1}{2},\ E(Y^2) = \frac{1}{3},\ \sigma_{YY} = \frac{1}{12}
$$

Finally:

$$
\sigma_{XY} = E\left[(X - m_X)(Y - m_Y)\right] = \int_{y=0}^1 \int_{x=0}^1 (xy)1\, dx\, dy - m_X m_Y = \left(\int_{y=0}^1 y\, dy\right)\left(\int_{x=0}^1 x\, dx\right) - m_X m_Y = 0
$$

Thus the covariance matrix is given by:

$$
\Sigma_{XY} = \begin{bmatrix} \frac{1}{12} & 0 \\ 0 & \frac{1}{12} \end{bmatrix}.
$$

In summary, i) X and Y are clearly uncorrelated, since $\sigma_{XY} = 0$, ii) X and Y are independent in this case. This can be seen from the fact that $p_X(x)p_Y(y) = p_{X,Y}(x,y)$ or from the equivalent fact (by Bayes rule) that $p_{X|Y}(x \mid y) = p_X(x)$ or $p_{Y|X}(y \mid x) = p_Y(y)$. iii) the covariance matrix is given above.

Problem 2.2

Given two random variables X and Y, knowledge of Y generally gives us information about the random variable X (and vice versa). Suppose we want to estimate X based on knowledge of Y (a topic we will study in much greater detail later in the course). In particular, suppose we want to estimate X as an affine function of Y:

$$
\widehat{X} = \widehat{X}(Y) = aY + b,
$$

where a and b are constants.

(a) Find expressions for a and b in terms of σ_{XX}, σ_{XY}, σ_{YY}, m_X, and m_Y so that the expected value of the square error between X and its estimate $\widehat{X}$ is minimized:

$$
\min_{a,b} E\left[(\widehat{X} - X)^2\right]
$$

(b) For each of the joint densities specified in Problem 2.1, what are the corresponding expressions for the values of a and b and the resulting estimator? In each case, graph the function $\widehat{X}(Y)$ on the plot of the joint density. What observations can you make? When does knowledge of Y *not* affect the affine estimate of X?

Solution:

(a) We are given that the form of the estimator is $\widehat{X} = aY + b$. Note that $\widehat{X}$ can be viewed as a random variable since a and b are constants and Y is a random variable. What we want is to:

$$
\begin{aligned}
\min E\left[(\widehat{X} - X)^2\right] &= \min_{a,b} E\left[(aY + b - X)^2\right] = \min_{a,b} E\left[(a^2Y^2 + b^2 + X^2 + 2abY - 2aXY - 2bX)\right] \\
&= \min_{a,b} \left(a^2 E\left[Y^2\right] + b^2 + E\left[X^2\right] + 2abE\left[Y\right] - 2aE\left[XY\right] - 2bE\left[X\right]\right)
\end{aligned}
$$

which is an unconstrained minimization, so we just take derivatives with respect to the unknowns a and b and set equal to zero. Note that a,b are constants with respect the expectation operator, so we may pull them outside of it, and the expectations themselves are just constants with respect to the differentiation operator.

$$\begin{aligned}
\frac{\partial}{\partial a}\left[a^2 E\left[Y^2\right] + b^2 + E\left[X^2\right] + 2abE\left[Y\right] - 2aE\left[XY\right] - 2bE\left[X\right]\right] &= \\
&= 2aE\left[Y^2\right] + 2bE\left[Y\right] - 2E\left[XY\right] = 0 \qquad (1)
\end{aligned}$$

$$\begin{aligned}
\frac{\partial}{\partial b}\left[a^2 E\left[Y^2\right] + b^2 + E\left[X^2\right] + 2abE\left[Y\right] - 2aE\left[XY\right] - 2bE\left[X\right]\right] &= \\
&= 2b + 2aE[Y] - 2E[X] = 0 \qquad (2)
\end{aligned}$$

Equations (1) and (2) give 2 equations in 2 unknowns (a, and b). Note that (2) gives that $b = E(X) - aE(Y)$. Substituting into (1) and solving for a and b we get:

$$a = \frac{\sigma_{XY}}{\sigma_{YY}} \qquad b = E(X) - \frac{\sigma_{XY}}{\sigma_{YY}} E(Y)$$

and the general form of the best mean square linear estimator is given by:

$$\widehat{x} = \frac{\sigma_{XY}}{\sigma_{YY}} Y + E(X) - \frac{\sigma_{XY}}{\sigma_{YY}} E(Y) = E(X) + \frac{\sigma_{XY}}{\sigma_{YY}} \left(y - E(Y)\right)$$

Note, we should really check that the solution we found is a minimum and not a maximum by checking the second derivative – I'll leave that to you. Note that the formulas for a and b are quite general and yield an intuitively pleasing result. The optimal estimator starts with the mean $E(X)$ and and modifies it only if the observed value of y is different from its mean. The amount of this modification is proportional to cross-covariance normalized by the covariance.

(b) (i) The coefficients and the estimator are given by:

$$a = -\frac{1}{2}, \quad b = \frac{1}{2}, \quad \widehat{x}(y) = \frac{1}{2}(1 - y)$$

(ii) The coefficients and the estimator are given by:

$$a = 0, \quad b = 0, \quad \widehat{x}(y) = 0y = 0$$

(iii) The coefficients and the estimator are given by:

$$a = 0, \quad b = \frac{1}{2}, \quad \widehat{x}(y) = 0y + 1/2 = 1/2$$

Note that when the two random variables are uncorrelated then the estimate is a vertical line, and knowledge of Y does not affect our estimate of X. Note also that the estimate functions $\widehat{x}(y)$ seem to follow the conditional mean (i.e. the expected value of X for the given value of Y). Thus, when the conditional mean happens to be a linear function (it need not always be, right?), then it appears that the conditional mean is the minimum mean square estimate of X given Y. This is in fact true in general, not just when the conditional mean is linear.

Problem 2.3

Consider the following 3×3 matrices:

$$A_1 = \begin{bmatrix} 10 & 3 & 1 \\ 2 & 5 & 0 \\ 1 & 0 & 2 \end{bmatrix} \quad A_2 = \begin{bmatrix} 10 & 5 & 2 \\ 5 & 3 & 3 \\ 2 & 3 & 2 \end{bmatrix} \quad A_3 = \begin{bmatrix} 10 & 5 & 2 \\ 5 & -3 & 3 \\ 2 & 3 & 2 \end{bmatrix} \quad A_4 = \begin{bmatrix} 10 & 5 & 2 \\ -5 & 3 & 1 \\ -2 & -1 & 2 \end{bmatrix}$$

$$A_5 = \begin{bmatrix} 10 & -5 & 2 \\ -5 & 3 & -1 \\ 2 & -1 & 2 \end{bmatrix} \quad A_6 = \begin{bmatrix} 10 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 2 \end{bmatrix} \quad A_7 = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix}$$

(a) Which of the above matrices could be the covariance matrix of some random vector?

(b) Which of the above matrices could be the cross-covariance matrix of two random vectors?

(c) Which of the above matrices could be the covariance matrix of a random vector in which one component is a linear combination of the other two components?

(d) Which of the above matrices *could* be the covariance matrix of a vector with statistically independent components? Must a random vector with such a covariance matrix have statistically independent components?

Solution:

(a) A_1 and A_4 are not valid covariances since they are not symmetric. A_3 is not a valid covariance since $(A_3)_{22} < 0$. A_2 is not a valid covariance since it is not positive semi-definite. In particular, it has a negative eigenvalue. Also note that if $u = [0, 1, -1]^T$ then $u^T A_2 u = -1 < 0$, so its not positive semi-definite. Finally, $\det(A_2) = -32$, but $\det(A_2) = \Pi(\text{eigenvals})$ and so must be non-negative for a positive semi-definite matrix. The rest are valid covariance matrices (i.e., A_5, A_6, and A_7). In particular they are all symmetric, positive semi-definite matrices.

(b) Since there are no structural restrictions on cross-covariance matrices, they could all be cross-covariance matrices. Recall that a cross-covariance matrix need not even be square.

(c) If one component of the random vector is a linear combination of the other two, then as we saw in class the corresponding covariance matrix must be singular. Of the 3 valid covariance matrices (A_5, A_6, and A_7), A_5 and A_6 are non-singular (and thus positive definite) – i.e. their determinant is positive. The determinant of $A_7 = 0$ however, so it is indeed singular and thus only positive *semi*-definite. Since it does not have full rank one of the components of X must be a linear combination of the others plus a constant.

(d) If a random vector has statistically independent components, the we know that it also has uncorrelated components. This means that the cross-covariances will be zero. Equivalently, this means that the covariance matrix must be diagonal. A_6 is the only diagonal matrix (and is also a valid covariance matrix).

A vector with such a covariance necessarily is composed of uncorrelated components (by definition), but this does not necessarily mean that the components are independent. Uncorrelatedness is only a second-order property.

Problem 2.4

Let $\underline{X} = \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ denote a Gaussian random vector with mean $\underline{m}_X = \begin{bmatrix} 0 \\ a \end{bmatrix}$ and covariance $\Sigma_X = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix}$ for some real a.

(a) Verify that Σ_X is a valid covariance matrix.

(b) What is the marginal probability density function $p_{X_1}(x_1)$ for X_1?

(c) Find the conditional density $p_{X_1|X_2}(x_1|x_2)$.

(d) Find a linear transformation T defining two new variables $\begin{bmatrix} Y_1 \\ Y_2 \end{bmatrix} = T \begin{bmatrix} X_1 \\ X_2 \end{bmatrix}$ such that Y_1 and Y_2 are uncorrelated and such that $TT^T = I$.

(e) Are Y_1 and Y_2 also statistically independent?

Solution:

(a) Σ_X is symmetric and is positive definite. We can see the latter by finding its eigenvalues or by applying "Sylvester's test" for positive definiteness, which says that the determinants of all principal minors of Σ_X must be positive:

$$4 > 0 \qquad \begin{vmatrix} 4 & 2 \\ 2 & 2 \end{vmatrix} = 4 > 0$$

(b) Since $\underline{X}$ is a Gaussian random vector both X_1 and X_2 are individually Gaussian. In particular, X_1 is Gaussian with mean $(\underline{m})_1 = 0$ and variance $(\Sigma_X)_{11} = 4$. Thus $X_1 \sim N(0, 4)$.

(c) We know that since X_1 and X_2 are jointly Gaussian, then $p_{X_1|X_2}(x_1|x_2)$ will also be Gaussian with parameters:

$$\begin{aligned} m_{X_1|X_2} &= m_{X_1} + \frac{\sigma_{X_1X_2}}{\sigma_{X_2X_2}}(x_2 - m_{X_2}) = 0 + \frac{2}{2}(x_2 - a) \\ \sigma_{X_1|X_2} &= \sigma_{X_1X_1} - \frac{\sigma^2_{X_1X_2}}{\sigma_{X_2X_2}} = 4 - \frac{4}{2} = 2 \end{aligned}$$

thus $p_{X_1|X_2}(x_1|x_2) = N(x_1; x_2 - a, 2)$

(d) We know $\underline{Y} = T\underline{X}$, $T^T = T^{-1}$, and $\Sigma_Y = T\Sigma_X T^T$. We want $\underline{Y}$ uncorrelated, which means its covariance matrix Σ_Y is diagonal. If we choose $T = U^T$ where U is the normalized eigenvector matrix of Σ_X then Σ_Y will be diagonal and the components of $\underline{Y}$ will be uncorrelated. First find the eigenvalues of Σ_X:

$$\Sigma_X = \begin{bmatrix} 4 & 2 \\ 2 & 2 \end{bmatrix} \Longrightarrow |\Sigma_X - \lambda I| = \left| \begin{bmatrix} 4-\lambda & 2 \\ 2 & 2-\lambda \end{bmatrix} \right| = (4-\lambda)(2-\lambda) - 4 = \lambda^2 - 6\lambda + 4 = 0$$

Thus the eigenvalues are $\lambda_1 = 3 + \sqrt{5}$ and $\lambda_2 = 3 - \sqrt{5}$. The eigenvector u_i corresponding to the eigenvalue λ_i is found by solving:

$$\begin{bmatrix} 4-\lambda & 2 \\ 2 & 2-\lambda \end{bmatrix} u_i = 0$$

This only defines them to a constant. We then normalize them to obtain:

$$\lambda_1 = 3 + \sqrt{5} \Longleftrightarrow u_1 = \begin{bmatrix} \frac{2}{\sqrt{10-2\sqrt{5}}} \\ \frac{\sqrt{5}-1}{\sqrt{10-2\sqrt{5}}} \end{bmatrix}$$

$$\lambda_2 = 3 - \sqrt{5} \Longleftrightarrow u_2 = \begin{bmatrix} \frac{2}{\sqrt{10+2\sqrt{5}}} \\ -\frac{\sqrt{5}+1}{\sqrt{10+2\sqrt{5}}} \end{bmatrix}$$

Thus:

$$T = U^T = \begin{bmatrix} u_1^T \\ u_2^T \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{10-2\sqrt{5}}} & \frac{\sqrt{5}-1}{\sqrt{10-2\sqrt{5}}} \\ \frac{2}{\sqrt{10+2\sqrt{5}}} & -\frac{\sqrt{5}+1}{\sqrt{10+2\sqrt{5}}} \end{bmatrix} = \begin{bmatrix} 0.8507 & 0.5257 \\ 0.5257 & -0.8507 \end{bmatrix}$$

(e) Yes, since $\underline{Y}$ is obtained as a linear combination of Gaussian R.V.s it is also Gaussian. Since it is Gaussian and its elements are uncorrelated, they are also statistically independent. For two Gaussian R.V.s uncorrelated is equivalent to statistically independent (though this is not the case in general).

Problem 2.5

The joint density function for two two-dimensional random vectors $\underline{X}$ and $\underline{Y}$ is

$$p_{\underline{X},\underline{Y}}(\underline{x},\underline{y}) = \begin{cases} x_1x_2 + 3y_1y_2 & 0 \le x_1, x_2, y_1, y_2 \le 1 \\ 0 & \text{otherwise} \end{cases}$$

Are $\underline{X}$ and $\underline{Y}$ statistically independent? Explain.

Solution:

No – If they are statistically independent then $p_{\underline{X},\underline{Y}}(\underline{x},\underline{y}) = p_{\underline{X}}(\underline{x})p_{\underline{Y}}(\underline{y})$ Find the marginals:

$$p_{\underline{X}}(\underline{x}) = \int_0^1 \int_0^1 (x_1x_2 + 3y_1y_2)\, dy_1\, dy_2 = x_1x_2 + \frac{3}{4} \qquad 0 \le x_1, x_2 \le 1$$

$$p_{\underline{Y}}(\underline{y}) = \int_0^1 \int_0^1 (x_1x_2 + 3y_1y_2)\, dx_1\, dx_2 = \frac{1}{4} + 3y_1y_2 \qquad 0 \le y_1, y_2 \le 1$$

Now note that $p_{\underline{X},\underline{Y}}(\underline{x},\underline{y}) \neq p_{\underline{X}}(\underline{x})p_{\underline{Y}}(\underline{y})$.

Another, perhaps simpler, way to argue this is that if they are independent then the joint density has to have the characteristics of a product of two densities (one for each coordinate) – that is, it must be "separable." Since the given joint density is the *sum* of two terms, each only involving x_i or y_i and having no cross terms, it can't come from such a product – that is, the joint density can't be factored in the necessary way.

Problem 2.6

Define a random process $X(t)$ based on the outcome k of tossing a fair die independently at each time t, where t is discrete-valued. The definition of X is as follows:

$$X(t) = \begin{cases} -2 & \text{if } k = 1, \\ -1 & \text{if } k = 2, \\ 1 & \text{if } k = 3, \\ 2 & \text{if } k = 4, \\ t & \text{if } k = 5, \\ -t & \text{if } k = 6 \end{cases}$$

(a) Find the joint probability density function of $X(0), X(2)$.

(b) Find the marginal probability density functions of $X(0), X(2)$.

(c) Find $E[X(0)], E[X(2)], E[X(0)X(2)]$.

Solution:

(a) Since they are independent, $p_X(a, b; 0, 2) = p_X(a; 0)p_X(b; 2)$. Now for any t:

$$p_X(a,t) = \frac{1}{6}\left(\delta(a+2) + \delta(a+1) + \delta(a-1) + \delta(a-2) + \delta(a-t) + \delta(a+t)\right)$$

$$\text{For } t=0 \qquad p_X(a,0) = \frac{1}{6}\left[\delta(a+2)+\delta(a+1)+\delta(a-1)+\delta(a-2)+2\delta(a)\right]$$

$$\text{For } t=2 \qquad p_X(b,2) = \frac{1}{6}\left[2\delta(b+2)+\delta(b+1)+\delta(b-1)+2\delta(b-2)\right]$$

Thus:

$$p_X(a,b;0,2) = \frac{1}{36}\left[\delta(a+2)+\delta(a+1)+\delta(a-1)+\delta(a-2)+2\delta(a)\right]\left[2\delta(b+2)+\delta(b+1)+\delta(b-1)+2\delta(b-2)\right]$$

(b) We have already done this above.

(c)

$$E[X(0)] = \int_{a=-\infty}^{\infty} a p_X(a,0)\,da = \int_{a=-\infty}^{\infty} a\frac{1}{6}\left[\delta(a+2)+\delta(a+1)+\delta(a-1)+\delta(a-2)+2\delta(a)\right]\,da = 0$$

And similarly:

$$E[X(2)] = \int_{b=-\infty}^{\infty} b p_X(b,2)\,db = \int_{b=-\infty}^{\infty} b\frac{1}{6}\left[2\delta(b+2)+\delta(b+1)+\delta(b-1)+2\delta(b-2)\right]\,db = 0$$

Finally,

$$E[X(0)X(2)] = E[X(0)]E[X(2)] = 0$$

since they are independent.

Problem 2.7 (Old Exam Problem) Let α and β be two statistically independent, identically distributed Gaussian random variables with means $E[\alpha] = E[\beta] = 0$ and variances $\sigma_\alpha^2 = \sigma_\beta^2 = 1$. Define the stochastic process $X(t) = \alpha\cos(t) + \beta\sin(t)$.

(a) Find the mean, $m_X(t)$, and autocorrelation, $R_{XX}(t_1,t_2)$, of the process $X(t)$.

(b) Is the process $X(t)$ wide-sense stationary? Explain.

(c) Is the process $X(t)$ a Gaussian random process? Explain.

Solution:

(a)

$$\begin{aligned}
m_X(t) &= E[X(t)] = E[\alpha\cos(t)+\beta\sin(t)] = E[\alpha]\cos(t)+E[\beta]\sin(t) = 0\cos(t)+0\sin(t) = 0 \\
R_{XX}(t_1,t_2) &= E[X(t_1)X(t_2)] = E\left[(\alpha\cos(t_1)+\beta\sin(t_1))(\alpha\cos(t_2)+\beta\sin(t_2))\right] \\
&= E\left[\alpha^2\cos(t_1)\cos(t_2)+\alpha\beta\cos(t_1)\sin(t_2)+\alpha\beta\cos(t_2)\sin(t_1)+\beta^2\sin(t_1)\sin(t_2)\right] \\
&= E[\alpha^2]\cos(t_1)\cos(t_2)+E[\alpha\beta]\cos(t_1)\sin(t_2)+E[\alpha\beta]\cos(t_2)\sin(t_1)+E[\beta^2]\sin(t_1)\sin(t_2) \\
&= 1\cos(t_1)\cos(t_2)+E[\alpha]E[\beta]\cos(t_1)\sin(t_2)+E[\alpha]E[\beta]\cos(t_2)\sin(t_1)+1\sin(t_1)\sin(t_2) \\
&= \cos(t_1)\cos(t_2)+\sin(t_1)\sin(t_2) = \cos(t_1-t_2) = \cos(t_2-t_1)
\end{aligned}$$

(b) Yes, $X(t)$ is wide-sense stationary because $m_X(t) = 0$, which is constant, and $R_{XX}(t_1,t_2)$ is only a function of time difference (t_1-t_2).

(c) Yes the process is Gaussian. Consider an arbitrary linear combination of $x(t)$ at an arbitrary set of times t_i:

$$\sum_i c_i X(t_i) = \alpha\left(\sum_i c_i\cos(t_i)\right) + \beta\left(\sum_i c_i\sin(t_i)\right)$$

This is always just a linear combination of two IID Gaussian random variables, which is Gaussian. Thus the process is Gaussian.

Problem 2.8 (Old Exam Problem)
Consider the random process

$$X(n) = \begin{cases} Z & n \text{ even} \\ -Z & n \text{ odd} \end{cases}$$

where $Z \sim N(m, 1)$. Be sure your answers are valid for all possible values of m and to give explanations for your answers.

(a) Find the mean and autocovariance functions of $X(n)$.

(b) Is $X(n)$ wide-sense stationary?

(c) Is $X(n)$ a Markov process?

(d) Is $X(n)$ strict-sense stationary?

(e) Is $X(n)$ a Gaussian random process?

(f) Is $X(n)$ an independent increments process?

Solution:

(a)

$$m_X(n) = E[X(n)] = \begin{cases} E[Z] & n \text{ even} \\ -E[Z] & n \text{ odd} \end{cases} = \begin{cases} m & n \text{ even} \\ -m & n \text{ odd} \end{cases} = m(-1)^n$$

$$\begin{aligned} K_{XX}(n,k) &= \text{cov}[X(n), X(k)] = \begin{cases} \text{cov}(Z,Z) & (n,k \text{ even}) \text{ or } (n,k \text{ odd}) \\ \text{cov}(Z,-Z) & (n \text{ even}, k \text{ odd}) \text{ or } (n \text{ odd}, k \text{ even}) \end{cases} \\ &= \begin{cases} 1 & n-k \text{ even} \\ -1 & n-k \text{ odd} \end{cases} = (-1)^{n-k} \end{aligned}$$

(b) A process is WSS if $m_X(n)$ is constant and $K_{XX}(n,k)$ is only a function of $n-k$.

Case 1: $m \neq 0$ Notice that $m_X(n)$ is not constant and thus $X(n)$ is clearly not WSS.

Case 2: $m = 0$ Now $m_X(n) = 0$ so it is possible it is WSS. We need to examine $K_{XX}(n,k)$. But $K_{XX}(n,k) = (-1)^{n-k}$ is indeed just a function of $n-k$ so in this case $X(n)$ is WSS.

(c) Yes it is a Markov process. The process is Markov if $p_{X(n_N)|X(n_{N-1}),X(n_{N-2}),\cdots}(x_{n_N}|x_{n_{N-1}}, x_{n_{N-2}}, \cdots) = p_{X(n_N)|X(n_{N-1})}(x_{n_N}|x_{n_{N-1}})$. Now $X(n_N) = \pm X(n_{N-1})$ (with the sign depending on whether $n_N - n_{N-1}$ is even or odd) and thus $X(n_N)$ is a *deterministic function* of $X(n_{N-1})$. This argument holds for any set of times we choose, so clearly the Markov condition holds.

(d) To be strict sense stationary (SSS) the process must be at least wide-sense stationary (WSS), and to be WSS the mean must be constant. Now $m_X(n) = m(-1)^n$, so it is clearly not SSS if $m \neq 0$. If $m = 0$ then we need to check if all the n-th order marginals are stationary. But note that if $m = 0$ then the density is symmetric around zero and since it alternates symmetrically we can't really distinguish a shift of the time origin. More precisely, note that for any set of times $n_1, \ldots, n_N$:

$$\begin{aligned} & p_{X(n_1),X(n_2),\cdots,X(n_N)}(x_1, x_2, \cdots, x_N) \\ &= p_{X(n_2),\cdots,X(n_N)|X(n_1)}(x_2, \cdots, x_N|x_1) p_{X(n_1)}(x_1) \\ &= \delta\left[x_2 - (-1)^{n_2-n_1}x_1\right] \delta\left[x_3 - (-1)^{n_3-n_1}x_1\right] \cdots \delta\left[x_N - (-1)^{n_N-n_1}x_1\right] p_{X(n_1)}(x_1) \end{aligned}$$

Now suppose we shift the original set of times by an arbitrary amount n_0 we get:

$$\begin{aligned}
& p_{X(n_1+n_0),X(n_2+n_0),\cdots,X(n_N+n_0)}(x_1,x_2,\cdots,x_N) \\
&= p_{X(n_2+n_0),\cdots,X(n_N+n_0)|X(n_1+n_0)}(x_2,\cdots,x_N|x_1)p_{X(n_1+n_0)}(x_1) \\
&= \delta\left[x_2-(-1)^{n_2+n_0-n_1-n_0}x_1\right]\delta\left[x_3-(-1)^{n_3+n_0-n_1-n_0}x_1\right]\cdots\delta\left[x_N-(-1)^{n_N+n_0-n_1-n_0}x_1\right]p_{X(n_1+n_0)}(x_1) \\
&= \delta\left[x_2-(-1)^{n_2-n_1}x_1\right]\delta\left[x_3-(-1)^{n_3-n_1}x_1\right]\cdots\delta\left[x_N-(-1)^{n_N-n_1}x_1\right]p_{X(n_1)}(x_1)
\end{aligned}$$

since $p_{X(n_1)}(x_1) \sim N(0,1)$ is the same for every n_1. Thus we see that the n-th order densities are shift invariant, and thus the process is strict sense stationary (SSS). We could have seen this (almost) above as the joint densities can be seen to depend only on the *time differences*–the only subtlety here is what happens with the marginal term $p_{X(n_1)}(x_1)$.

Another line of argument is the following. The process is Gaussian (see part (e)) and is also wide-sense stationary (WSS) when $m=0$, thus for this case it is also strict sense stationary.

(e) Yes it is Gaussian. We need to check that $\sum_{i=1}^N a(n_i)X(n_i)$ is Gaussian for all $a(n_i)$ and N. Note that $X(n)=(-1)^n Z$ so that:

$$\sum_{i=1}^N a(n_i)X(n_i) = \left(\sum_{i=1}^N a(n_i)(-1)^{n_i}\right)Z = ZK \sim N(Km,K^2)$$

where K is some constant, so it is always Gaussian.

(f) No, it is not an independent increments process. Note that $X(3)-X(2)$ is *equal* to $X(1)-X(0)$ so they can't be independent. Another way to see it is that if it is an independent increments process then it should certainly be an uncorrelated increments process (a condition implied by independence of the increments) – i.e. we should have $\text{cov}(X(n)-X(m),X(k)-X(l))=0$ for all $n>m\geq k>l$. Now consider e.g. $\text{cov}(X(2)-X(1),X(1)-X(0)) = \text{cov}(2Z,-2Z) = -4$ which is not zero, so the increments are not uncorrelated, so they certainly won't be independent.

Problem 2.9 (Old exam question) A random process $X(t)$ is defined as follows:

$$X(t) = \begin{cases} A & t<\Theta \\ B & t\geq\Theta \end{cases}$$

where A, B, and Θ are statistically independent unit-variance Gaussian random variables with means -1, 1, and 0 respectively, i.e., $A\sim N(-1,1)$, $B\sim N(1,1)$, $\Theta\sim N(0,1)$.

(a) Sketch a typical sample function of $X(t)$.

(b) Find the first order density $p_{X(t)}(x)$, and the conditional density $p_{X(t_2)|X(t_1)}(x_2|x_1)$ for $t_2>t_1$. Is $X(t)$ a Gaussian random process? Explain.

(c) Is $X(t)$ strict sense stationary? Explain.

(d) Is $X(t)$ a Markov process? Explain. (Hint: Consider $p_{X(t_2)|X(t_1),X(t_0)}(x_2|x_1,x_0)$ for different choices of $X(t_1),X(t_0)$).

(e) Is $X(t)$ an independent increments process? Explain.

If appropriate, you may express your answers to some parts of this problem in terms of the "Q" function defined below.

$$Q(t) = \frac{1}{\sqrt{2\pi}}\int_t^\infty e^{-\alpha^2/2}\,d\alpha$$

Solution:

(a) In the sample path, a is probably near -1 and b is probably near 1, and the crossover occurs at $t = \theta$, which is probably near zero.

(b) To find the first order density we can find the joint density for $X(t)$ and Θ using Bayes rule to write it in terms of a conditional and then sum out Θ.

$$p_{X(t_0)}(x) = \int p_{X(t_0),\Theta}(x, \theta)\, d\theta = \int p_{X(t_0)|\Theta}(x|\theta)\, p_\Theta(\theta)\, d\theta$$

Now note that:

$$\begin{array}{lll} \Theta \leq t_0 & \Longrightarrow & p_{X(t_0)|\Theta}(x|\theta) \sim N(+1, 1) \\ \Theta > t_0 & \Longrightarrow & p_{X(t_0)|\Theta}(x|\theta) \sim N(-1, 1) \end{array}$$

Thus,

$$\begin{aligned} p_{X(t_0)}(x) &= \int p_{X(t_0)|\Theta}(x|\theta)\, p_\Theta(\theta)\, d\theta = \int_{\Theta \leq t_0} N(x; +1, 1) p_\Theta(\theta)\, d\theta + \int_{\Theta > t_0} N(x; -1, 1) p_\Theta(\theta)\, d\theta \\ &= N(x; +1, 1) \int_{\Theta \leq t_0} p_\Theta(\theta)\, d\theta + N(x; -1, 1) \int_{\Theta > t_0} p_\Theta(\theta)\, d\theta \\ &= p_{X(t_0)|\Theta \leq t_0}(x|\theta \leq t_0) \Pr(\Theta \leq t_0) + p_{X(t_0)|\Theta > t_0}(x|\theta > t_0) \Pr(\Theta > t_0) \\ &= N(x; +1, 1)\, [1 - Q(t_0)] + N(x; -1, 1) Q(t_0) \end{aligned}$$

Note, this is a mixture of Gaussians, as shown below:

Actually, the drawing is exaggerated to make the point – the Gaussians are close enough together given their variance that they would look like a single large lump.

Now to find the conditional density, note that $X(t_1)$ and $X(t_2)$ are independent Gaussian random variables if $t_1 < \theta \leq t_2$ and $X(t_1) = X(t_2)$ otherwise. These are the only two possibilities. Thus:

$$\begin{aligned} p_{X(t_2)|X(t_1)}(x_2|x_1) &= \delta(x_2 - x_1) \Pr\left[\Theta \leq t_1 \text{ or } t_2 < \Theta\right] + N(x_2; +1, 1) \Pr\left[t_1 < \Theta \leq t_2\right] \\ &= \delta(x_2 - x_1)\, [1 - Q(t_1) + Q(t_2)] + N(x_2; +1, 1)\, [Q(t_1) - Q(t_2)] \\ &= \delta(x_2 - x_1)\, [Q(-t_1) + Q(t_2)] + N(x_2; +1, 1)\, [Q(t_1) - Q(t_2)] \end{aligned}$$

Where we have used the fact that $1 - Q(t_1) = Q(-t_1)$, by the symmetry of the Gaussian. Note that this distribution looks as in the figure, and thus is <u>not</u> Gaussian. Clearly the process is not even first order Gaussian, so it can't be Gaussian random process. Another way to see this is that the conditional density isn't Gaussian.

(c) No, it is not SSS. We can see that even the first order density found in (b) is not stationary, so again it cannot be SSS. Also, Suppose $t_2 > t_1$. Then

$$\begin{aligned} R_{XX}(t_2,t_1) &= E[X(t_2)X(t_1)] = E\left[E\left[X(t_2)X(t_1)|\Theta\right]\right] \\ &= E\left[X(t_2)X(t_1)|\Theta \le t_1 \text{ or } t_2 < \Theta\right]\Pr\left(\Theta \le t_1 \text{ or } t_2 < \Theta\right) + \\ &\quad E\left[X(t_2)X(t_1)|t_1 < \Theta \le t_2\right]\Pr\left(t_1 < \Theta \le t_2\right) \end{aligned}$$

Now note that when $\Theta \le t_1$ or $t_2 < \Theta$ then $X(t_2) = X(t_1)$, and when $t_1 < \Theta \le t_2$, then $X(t_2)$ is independent of $X(t_1)$. Also, $E[X^2(t_1)] = 1 + 1^2 = 2$. Thus we have:

$$\begin{aligned} R_{XX}(t_2,t_1) &= E\left[X(t_2)X(t_1)|\Theta \le t_1 \text{ or } t_2 < \Theta\right]\Pr\left(\Theta \le t_1 \text{ or } t_2 < \Theta\right) + \\ &\quad E\left[X(t_2)X(t_1)|t_1 < \Theta \le t_2\right]\Pr\left(t_1 < \Theta \le t_2\right) \\ &= E\left[X^2(t_1)\right]\Pr\left(\Theta \le t_1 \text{ or } t_2 < \Theta\right) + E[X(t_2)]E[X(t_1)]\Pr\left(t_1 < \Theta \le t_2\right) \\ &= 2\cdot\Pr\left(\Theta \le t_1 \text{ or } t_2 < \Theta\right) + (+1)\cdot(-1)\Pr\left(t_1 < \Theta \le t_2\right) \\ &= 2\Pr\left(\Theta \notin (t_1,t_2]\right) - \Pr\left(\Theta \in (t_1,t_2]\right) \\ &= 2 - 3\left[Q(t_1) - Q(t_2)\right] \end{aligned}$$

By symmetry, we get in general that $R_{XX}(t_2,t_1) = 2 - 3\left[Q(\min(t_1,t_2)) - Q(\max(t_1,t_2))\right]$ so the autocorrelation depends on absolute t, and not just $|t_1 - t_2|$ so it is not even WSS. Also, you can see it is not WSS by noting that the mean is not constant, but is a function of t. At one extreme the process mean tends to the mean of A as $t \to -\infty$ and to the mean of B as $t \to +\infty$.

(d) No. It is not Markov. For Markovianity,

$$\Pr[X(t_N) = X_N|X(t_{N-1}) = X_{N-1}, X(t_{N-2}) = X_{N-2}, ..., X(t_1) = X_1] = \Pr[X(t_N) = X_N|X(t_{N-1}) = X_{N-1}],$$

But suppose $X_{N-1} \neq X_{N-2}$. Then $\Pr(X(t_N) = X_{N-1}) = 1$ (i.e. we know the value of $X(t)$ at all future times with certainty) because the crossover at $t = \Theta$ must have already occurred with absolute certainty before $t = t_{N-1}$, whereas there is some uncertainty of having already had a crossover given only one previous observation of $X(t)$ at $t = t_{N-1}$. In other words,

$$\Pr[X(t_N) = X_{N-1}|X(t_{N-1}) = X_{N-1}, X(t_{N-2}) = X_{N-2}, X_{N-1} \neq X_{N-2}] = 1$$

but

$$\Pr[X(t_N) = X_{N-1}|X(t_{N-1}) = X_{N-1}] < 1$$

so it can't be Markov.

(e) The process is not IIP. Since all IIP processes are Markov and our process is not Markov, it can't be. Another, more direct argument would be as follows. Consider the process at three times $t_1 < t_2 < t_3$. Consider the increment $X(t_3) - X(t_2)$ and its first order density $p_{X(t_3)-X(t_2)}(x_3 - x_2)$. Clearly, knowing nothing else, there is uncertainty in the value of this increment. Now consider $p_{X(t_3)-X(t_2)|X(t_2)-X(t_1)}(x_3 - x_2|x_2 - x_1)$. If the increments are independent then knowledge of $x_2 - x_1$ *should not affect the density* – i.e. for independent processes the conditioning event has no effect.

Now suppose that we know that $x_2 - x_1 \neq 0$ (a particular possible conditioning event). Given this information note that we *know* that $x_3 - x_2 = 0$ since we know the "step" of the process has occured – i.e. we know that $p_{X(t_3)-X(t_2)|X(t_2)-X(t_1)}(x_3 - x_2|x_2 - x_1 \neq 0) = \delta(x_3 - x_2)$. Clearly the conditioning event has affected the density! Thus it can't be independent increments.

Problem 2.10 Let $X(t)$ be a Gaussian random process with mean $m_X(t)$, autocorrelation $R_X(t_1, t_2)$ and let $t_1 < t_2$. Find an expression for $E[X(t_2)|X(t_1)]$ in terms of m_X, R_X.
Solution: From the Gaussian conditional expectation formula we have,

$$E[X(t_2) \mid X(t_1)] = m_X(t_2) + \frac{(R_X(t_1, t_2) - m_X(t_2)m_X(t_1))}{(R_X(t_1, t_1) - m_X(t_1)m_X(t_1))}(X(t_1) - m_X(t_1))$$

Computer Problems

Problem 2.11 Gaussian Random Vectors

The purpose of this problem is to learn how to generate samples of Gaussian random vectors with a given covariance structure and, conversely, given data, how to estimate the covariance matrix of a random vector. Random vectors also have a close relationship with random processes, particularly when we attempt to work with random processes in a computer. Thus we will also learn how to manipulate such vectors in MATLAB.

(a) Suppose $\underline{Z} = [Z_1, Z_2]^T$ is a vector of two independent Gaussian random variables Z_i, with $Z_i \sim N(0,1)$ and L is a 2×2 matrix. Define the (Gaussian) random vector $\underline{X} = [X_1, X_2]^T = L\underline{Z}$. What is the covariance matrix R_X of $\underline{X}$ in terms of L? If L is as given below in (3), what is the corresponding R_X?

$$L = \begin{bmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{bmatrix} \tag{3}$$

(b) We will use MATLAB to confirm part (a) experimentally. To this end we want to perform the following experiment:

(i) Generate N independent samples $\underline{z}_i$ of the vector $\underline{Z}$ in MATLAB

(ii) For each vector sample generated in (i), form the corresponding transformed sample $\underline{x}_i = L\underline{z}_i$,

(iii) Estimate the covariance matrix by calculating: $\widehat{R}_x = \frac{1}{N}\sum_{i=1}^{N} \underline{x}_i \underline{x}_i^T - \widehat{m}_x \widehat{m}_x^T$ where $\widehat{m}_x = \frac{1}{N}\sum_{i=1}^{N} \underline{x}_i$.

Verify and implement this procedure for the L specified in (3) and compare your answers to what the theory says R_X should be. Use as large a value of N as is reasonable.

Note/Hint: With a bit of thought you can perform the necessary calculations efficiently in MATLAB with out having to resort to loops. For example, suppose we store each independent sample vector of $\underline{Z}$ as a row in the MATLAB matrix `Z`. Then calculating all the corresponding transformed vectors $\underline{X}$ given the matrix `L` is easily done in MATLAB via the operation: `X = (L*Z')'`, where `Z'` takes a matrix transpose and the matrix `X` now stores the transformed sample vectors row-wise in the same way as `Z` itself. Further, if you read the help page on the MATLAB function `cov.m` you will see you are setup to find the covariance too.

We will use variants of this representational scheme for random vector and process experiments in MATLAB throughout the rest of the lab, so be sure you are comfortable with it.

(c) Use `pdf2d.m` to generate the empirical joint pdfs $p_{Z_1,Z_2}(z_1, z_2)$ and $p_{X_1,X_2}(x_1, x_2)$. What is the effect of applying the linear transformation L to $\underline{Z}$?

(d) You have now solved the "forward" problem of producing samples of a random vector given the *linear transformation* L. Use these insights to suggest a way to generate a Gaussian random vector $\underline{X}$ with a given, *desired covariance* structure R_X. Write a MATLAB function `covgen.m` based on your scheme, which takes as input a $P \times P$ desired covariance `R` and the desired number of sample vectors `N` and produces a $N \times P$ matrix `X`, whose rows are samples of the specified vector. Thus the function call would look like: `X = covgen(N,R)`. You may find the MATLAB functions `sqrtm.m`, which produces a "generic" matrix square root or `chol.m`, which factors a symmetric positive definite matrix, useful. Use your function when R_X is as given in (4) and verify it works as expected. You now have a way to generate vectors with arbitrary covariances!

$$R_X = \begin{bmatrix} 5/2 & 3/2 \\ 3/2 & 5/2 \end{bmatrix} \tag{4}$$

(e) In what ways would the random vector $\underline{X}$ and its covariance change if in the above development the elements of $\underline{Z}$ were independent, identically distributed, zero-mean, unit variance, *uniformly distributed* random variables rather than standard Gaussian random variables (i.e. if `Z = sqrt(12)*(rand(N,2)-.5)` were used instead of `Z = randn(N,2)`)? Try this and calculate both the covariance and the joint pdf.

Solution:

(a)

$$\begin{aligned}
R_X &= E\left[\underline{X}\underline{X}^T\right] - m_x m_x^T = E\left[L\underline{Z}\underline{Z}^T L^T\right] - 0 = LE\left[\underline{Z}\underline{Z}^T\right] L^T \\
&= LR_Z L^T = LL^T \\
R_X &= \begin{bmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 3/2 & 1/2 \\ 1/2 & 3/2 \end{bmatrix} \\
&= \begin{bmatrix} 5/2 & 3/2 \\ 3/2 & 5/2 \end{bmatrix} = \begin{bmatrix} 2.5 & 1.5 \\ 1.5 & 2.5 \end{bmatrix}
\end{aligned}$$

(b) A fragment of MATLAB code to perform the calculations is given below:

```
N = 100000;
L = [3/2 1/2;1/2 3/2];
Z = randn(N,2);
X = (L*Z')';
R = cov(X);
```

Each row of `Z` is an independent sample vector of $\underline{Z}$ (thus we are storing the vectors row-wise). In the second line, we have used the transpose operator to match the dimensions of `Z` to those of `L` so that all the transformations are done simultaneously. The structure of `X` matches that of `Z`, i.e. each row is a different experimental outcome and each column a different component of $\underline{X}$. With data stored this way, `cov.m` calculates the empirical covariance matrix and `mean.m` calculates the empirical mean vector.

The corresponding estimate of R_X obtained by MATLAB is:

$$\widehat{R}_x = \begin{bmatrix} 2.5103 & 1.5084 \\ 1.5084 & 2.4991 \end{bmatrix}$$

(c) The empirical joint pdfs are shown in Figure 1. The effect of applying the linear transformation is to rotate and stretch the pdf.

(d) The key idea is that when $R_Z = I$ then $R_X = LL^T$. Thus we want to choose the transformation L as a *matrix square root* of R_X. A program that does this `covgen.m` is given below:

```
function [X,Re] = covgen(N,R)
% [X,Re] = covgen(N,R)
%
% N : Number of sample vectors to generate
% R : Desired covariance structure
%
% X : Matrix of sample Gaussian vectors, each with the desired covariance
%     structure. Each row is a different sample vector, each column
%     is a different element of the vector.
% Re: Experimental covariance
%
```

Figure 1: Empirical joint pdfs when Z is Gaussian.

```
% Creates series of samples of a Gaussian vector with a given covariance
% structure. The vectors are stored as the rows of the output matrix X.

% W. C. Karl 1/20/97

if max(size(N))>1
  error('N must be a scalar')
end;

if size(R,1)~=size(R,2)
  error('Covariance matrix must be square')
end;

L = sqrtm(R);
Z = randn(N,2);
X = (L*Z')';
Re = cov(X);
```

When this code is used with the given R_X and $N = 500,000$, the associated estimated covariance of X is given below. So it indeed works.

$$\widehat{R}_x = \begin{bmatrix} 2.5043 & 1.5057 \\ 1.5057 & 2.5040 \end{bmatrix}$$

(e) The covariance will stay the same. The pdfs will change to a uniform pdf for $\underline{Z}$ and a rotated and stretched uniform pdf (i.e. a trapezoid) for $\underline{X}$. The covariance calculated by MATLAB for the uniform case is given by:

$$\widehat{R}_x = \begin{bmatrix} 2.5010 & 1.5045 \\ 1.5045 & 2.5065 \end{bmatrix}$$

Essentially the same as before with the Gaussian case. The corresponding empirical joint pdfs for the uniform case are shown in Figure 2. The effect of applying the linear transformation is again to rotate and stretch the pdf. Note that one conclusion is that random vectors can have the same covariance matrices and still have very different pdfs.

Figure 2: Empirical joint pdfs when Z is Uniform.