

SIMG-713

Homework 5

Spring 2002

1. Photons strike a detector at an average rate of λ photons per second. The detector produces an output with probability β whenever it is struck by a photon. Compute the DQE of the detector.
2. This problem examines the correlation and covariance of two random variables \mathbf{X} and \mathbf{Y} . The covariance function has been defined as

$$\text{cov}(X, Y) = E [(\mathbf{X} - \boldsymbol{\mu}_x)(\mathbf{Y} - \boldsymbol{\mu}_y)]$$

where $\mu_x = E[\mathbf{X}]$ and $\mu_y = E[\mathbf{Y}]$. The correlation coefficient is defined to be the ratio

$$\rho_{xy} = \frac{\text{cov}(\mathbf{X}, \mathbf{Y})}{\sigma_x \sigma_y}$$

where, respectively, σ_x^2 and σ_y^2 are the variances of \mathbf{X} and \mathbf{Y} .

- (a) This is a result that will be needed in some other parts of this problem. Show that the function $v = au^2 + bu + c$ is positive for all values of u provided $c > 0$ and $b^2 < 4ac$. The quantity $d = b^2 - 4ac$ is called the *discriminant*. The condition that $v > 0$ for all values of u is equivalent to $d < 0$.
 - (b) Show that $|\text{cov}(\mathbf{X}, \mathbf{Y})| \leq \sigma_x \sigma_y$. Hint: Look at $E \left[((\mathbf{X} - \mu_x)u + (\mathbf{Y} - \mu_y))^2 \right]$ as a function of u . Make use of the discriminant.
 - (c) Show that $|\rho_{xy}| \leq 1$.
 - (d) Show that for any two random variables \mathbf{X} and \mathbf{Y} , $E^2[\mathbf{XY}] \leq E[\mathbf{X}^2] E[\mathbf{Y}^2]$.
 - (e) Show that if $E[\mathbf{X}^2] = E[\mathbf{Y}^2]$ then $\mathbf{S} = \mathbf{X} + \mathbf{Y}$ and $\mathbf{D} = \mathbf{X} - \mathbf{Y}$ are orthogonal random variables.
 - (f) Show that if random variables \mathbf{X} and \mathbf{Y} are uncorrelated then $\sigma_s^2 = \sigma_x^2 + \sigma_y^2$, where $\mathbf{S} = \mathbf{X} + \mathbf{Y}$. Extend this to the sum of any number of uncorrelated random variables. Note that we once proved this for statistically independent random variables, but that this requirement is new because it is weaker.
3. If \mathbf{X} and \mathbf{Y} are jointly normal random variables, then so are random variables $\mathbf{S} = a\mathbf{X} + b\mathbf{Y}$ and $\mathbf{D} = c\mathbf{X} + d\mathbf{Y}$ for any real coefficients (a, b, c, d) .
 - (a) Given $\mu_x = 10$, $\mu_y = 0$, $\sigma_x = 2$, $\sigma_y = 1$, $\rho_{xy} = 0.5$ construct a contour plot of for $f_{XY}(x, y)$ showing the locus of constant probability points
 - (b) Find the probability density function $f_{SD}(x, y)$ for the random variables $\mathbf{S} = \mathbf{X} + \mathbf{Y}$ and $\mathbf{D} = \mathbf{X} - \mathbf{Y}$.

- (c) Construct a contour plot of for $f_{SD}(x, y)$ showing the locus of constant probability points.
4. Let X and Y be random variables with a joint probability density function $f_{XY}(x, y)$. Let $\hat{Y} = g(X)$ be a predictor of Y .

(a) Show that the mean-squared prediction error can be expressed as

$$E \left[(Y - \hat{Y})^2 \right] = E [Y^2] - 2E [g_0(X)g(X)] + E [g^2(X)]$$

where $g_0(X) = E[Y|X]$.

(b) Let X and Y be random variables with a joint probability density function $f_{XY}(x, y)$. Let $\hat{Y}_0 = g_0(X) = E[Y|X]$ be a predictor of Y . Show that the mean-squared prediction error can be expressed as

$$E \left[(Y - \hat{Y}_0)^2 \right] = E [Y^2] - E [g_0^2(X)]$$

(c) Show that the previous two problems establish that

$$E \left[(Y - \hat{Y})^2 \right] \geq E \left[(Y - \hat{Y}_0)^2 \right]$$

for any prediction function $\hat{Y} = g(X)$. That is, the conditional expectation of Y given X gives the least-mean-square prediction of Y . This shows that $g_0(X) = E[Y|X]$ is the “optimum” predictor function and provides a tool to find the optimum predictor.

5. Suppose that $Y = \alpha X + \beta + Z$ where Z is a random variable statistically independent of X with mean $E[Z] = \mu_z$. What is the optimum predictor function $\hat{Y} = g(X)$, based on the above results?
6. In this problem we will construct a formulation of the probability density function for the bivariate normal distribution based on the covariance matrix and mean values. This approach extends to any number of dimensions and is very useful in constructing algorithms. We begin by assuming that $\mathbf{X} = [X_1, X_2]^T$ is a column vector whose elements are statistically independent normal random variables.

(a) Show that

$$f_{\mathbf{X}}(\mathbf{x}) = \frac{1}{2\pi \det(\Gamma)^{1/2}} e^{-(\mathbf{x} - \mathbf{m}_x)^T \Gamma^{-1} (\mathbf{x} - \mathbf{m}_x)}$$

where $\mathbf{m} = [E[X_1], E[X_2]]^T$ is a column vector of the mean values and Γ is the covariance matrix

$$\Gamma = \begin{bmatrix} \text{var}(X_1) & \text{cov}(X_1, X_2) \\ \text{cov}(X_1, X_2) & \text{var}(X_2) \end{bmatrix} = \begin{bmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_2^2 \end{bmatrix}$$

In reality, this is just a compact way to express the equation

$$f_{X_1 X_2}(x_1, x_2) = \frac{1}{2\pi\sigma_1\sigma_2} \exp -\left[\left(\frac{x_1-m_1}{s_1}\right)^2 + \left(\frac{x_2-m_2}{s_2}\right)^2\right]$$

- (b) Let \mathbf{G} be a square matrix with $\det(\mathbf{G}) \neq 0$ and let $\mathbf{Y} = \mathbf{G}\mathbf{X}$. That is, \mathbf{Y} is a vector of random variables formed by a linear combination of elements of \mathbf{X} . The only restriction we are making is that the transformation should have an inverse. Show that $\mathbf{m}_y = \mathbf{G}\mathbf{m}_x$. This means that $\mathbf{Y} - \mathbf{m}_y = \mathbf{G}(\mathbf{X} - \mathbf{m}_x)$.
- (c) Show that the covariance matrix for \mathbf{Y} is

$$\Lambda = \mathbf{G}\Gamma\mathbf{G}^T$$

- (d) One can make a change of variables in n dimensions by

$$f_{\mathbf{Y}}(\mathbf{y}) = f_{\mathbf{X}}(\mathbf{G}^{-1}\mathbf{y}) |\det(\mathbf{G}^{-1})|$$

The exponent is transformed by

$$\begin{aligned} (\mathbf{x} - \mathbf{m}_x)^T \Gamma^{-1} (\mathbf{x} - \mathbf{m}_x) &= [\mathbf{G}^{-1}(\mathbf{y} - \mathbf{m}_y)]^T (\mathbf{G}^{-1}\Lambda\mathbf{G}^{-T})^{-1} [\mathbf{G}^{-1}(\mathbf{y} - \mathbf{m}_y)] \\ &= (\mathbf{y} - \mathbf{m}_y)^T \mathbf{G}^{-T} \mathbf{G}^T \Lambda^{-1} \mathbf{G} \mathbf{G}^{-1} (\mathbf{y} - \mathbf{m}_y) \\ &= (\mathbf{y} - \mathbf{m}_y)^T \Lambda^{-1} (\mathbf{y} - \mathbf{m}_y) \end{aligned}$$

Also, $\det(\Gamma) = \det(\mathbf{G}^{-1}\Lambda\mathbf{G}^{-T}) = \det(\Lambda) \det^2(\mathbf{G}^{-1})$. When all this is substituted back we find

$$f_{\mathbf{Y}}(\mathbf{y}) = \frac{1}{2\pi \det(\Lambda)^{1/2}} e^{-(\mathbf{y}-\mathbf{m}_y)^T \Lambda^{-1} (\mathbf{y}-\mathbf{m}_y)}$$

This is exactly the same form, but now it accommodates random variables that are not uncorrelated. This is a demonstration that a linear transformation of normal random variables produces another set of normal random variables. Assume that we are working in 2D and that X_1 and X_2 are statistically independent normal random variables. Find expressions for m_{y1} , m_{y2} , σ_{y1} , σ_{y2} , and ρ in terms of m_{x1} , m_{x2} , σ_{x1} , σ_{x2} and the elements of the transformation matrix

$$\mathbf{G} = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix}$$

Assume $\det \mathbf{G} \neq 0$. Write an expression for $f_{\mathbf{Y}}(\mathbf{y})$ in terms of m_{y1} , m_{y2} , σ_{y1} , σ_{y2} , and ρ .