Optimal Estimation University of Florida
Mechanical and Aerospace Engineering

HW 4
Issued: September 18, 2009, Due: September 25, 2009 (in class)

Note: Your work for the questions with "[0 pt]" are not to be turned in.

Problem 1. [5 pt]
Show that $var(aX) = a^2\, var(X)$ (a is a deterministic parameter).

Problem 2. [10 + 10 = 20 pt]
Prove the CBS (Cauchy-Bunyakovskii-Schwarz) Inequality for random variables. That is, show that for two random variables X and Y,

$$Covar(X,Y) \leq \sqrt{Var(X)Var(Y)}$$

and similarly

$$|\mathrm{E}[XY]|^2 \leq \mathrm{E}[X^2]\,\mathrm{E}[Y^2]$$

Problem 3. [10 + 10 = 20 pt]
Show the following triangle inequality for random variables, that for two random variables X and Y,

$$\sqrt{\mathrm{E}[(X+Y)^2]} \leq \sqrt{\mathrm{E}[X^2]} + \sqrt{\mathrm{E}[Y^2]}$$

and

$$\sqrt{Var(X+Y)} \leq \sqrt{Var(X)} + \sqrt{Var(Y)}$$

Problem 4 (Sample variance as an estimate of the variance)**.** [5+10+0 = 15 pt]
Let $X_1, \dots, X_n$ be n independent random variables that have the same mean $\mu = \mathrm{E}[X_i]$ and variance $\sigma^2 = \mathrm{E}[(X-\mu)^2]$, and $\mu < \infty, \sigma^2 < \infty$. Now consider the *sample mean* $\hat{\mu}$ and the *sample variance* $\hat{\sigma^2}$:

$$\hat{\mu} = \frac{X_1 + X_2 + \dots X_N}{N},$$
$$\hat{\sigma^2} = \frac{(X_1-\hat{\mu})^2 + (X_2-\hat{\mu})^2 + \cdots + (X_N-\hat{\mu})^2}{N-1},$$

1. What is the variance of the sample mean?
2. Show that the sample variance $\hat{\sigma^2}$ is an unbiased estimate of the variance σ^2. (Hint: use $Var(X) = \mathrm{E}[X^2] - (\bar{X})^2$ to keep the calculations for becoming too complicated)

3. Read about the variance of the error in the variance estimate (e.g., in Papoulis 3rd Ed., pp. 188,189, and pp. 239)

Problem 5 (Properties of covariance matrices). [5 + 10 + 10 + 5 =20 pt]

Let $R \triangleq Cov(\mathbf{X}, \mathbf{X})$ be the covariance matrix of the random vector $\mathbf{X} = [X_1, \ldots, X_n]^T$ and let $T \triangleq \mathrm{E}[\mathbf{X}\mathbf{X}^T]$ be its correlation matrix. Show that

1. the covariance matrix R is symmetric.
2. the covariance matrix R is positive semidefinite. (Hint: $Var(\alpha^T \mathbf{X}) \geq 0$ for a real vector α.)
3. the correlation matrix T is positive semidefinite. (Hint: this can be proved using part 1 and a relationship you proved in HW 3.)
4. if $X_1, \ldots, X_n$ are independent and $Var(X_i) = \sigma_i^2$, where $\sigma_i^2 > 0$ for every i, then the covariance matrix R is positive definite.

Problem 6. [0 pt]

First some discussion. If x is a vector in $\mathbb{R}^n$, then the $n \times n$ matrix xx^T has rank 1. Therefore, for $n > 1$, xx^T is not positive definite, no matter how you choose x. Consider the random vector $\mathbf{X}$ consisting of n random variables and the vector $\mathbf{X}(\omega)$, which is the value of the random vector $\mathbf{X}$ for a particular outcome ω of the underlying experiment. Since $\mathbf{X}(\omega)$ is a vector in $\mathbb{R}^n$, for each ω, the matrix $\mathbf{X}(\omega)\mathbf{X}^T(\omega)$ is not positive definite. The previous problem shows, however, that the *expected value* of the matrix $\mathbf{X}\mathbf{X}^T$ can be positive definite. In fact, it can become positive definite even without the random variables $X_1, \ldots, X_n$ being independent.

Let's do some numerical calculations to see such a phenomenon. Generate a n-dimensional random vector in Matlab, e.g., by using `x = randn(n,1)` to generate n pseudo-random numbers that are Normally distributed with mean 0 and variance 1. Repeat it for a few iterations and estimate a covariance matrix by taking a sample mean. You can estimate the individual entries of the covariance matrix or average the matrix xx^T directly. Compute the rank of this sample mean at every iteration. Since rank computation is usually quite sensitive to small errors, you can do a SVD decomposition to evaluate the "numerical" rank. That is, you can declare a square matrix as having full rank if all of its singular values are greater than some specified, small, positive number. Check how many "sample vectors" you need to average before the estimated covariance matrix becomes full rank.

Problem 7 (Two Jointly Normal Random Variables). [5 pt]

Two random variables X and Y are said to be *jointly normal* if their joint probability density function is given by

$$f_{XY}(x,y) = \frac{1}{2\pi\sigma_X\sigma_Y\sqrt{1-\rho^2}} \times \exp\left(\frac{-1}{2(1-\rho^2)}\left\{\left(\frac{x-\bar{X}}{\sigma_X}\right)^2 - 2\rho\frac{(x-\bar{X})(y-\bar{Y})}{\sigma_X\sigma_Y} + \left(\frac{y-\bar{Y}}{\sigma_Y}\right)^2\right\}\right)$$

Note that five parameters, $\bar{X}, \bar{Y}, \sigma_x, \sigma_Y$, and ρ are needed to specify this density.

Let

$$R_2 = \begin{bmatrix} Var(X) & Covar(X,Y) \\ Covar(Y,X) & Var(Y,Y) \end{bmatrix}$$

be the covariance matrix of the random vector $\mathbf{X} \triangleq [X, Y]^T$. Then the expression above for the density function can be compactly represented as:

$$f_{\mathbf{X}}(x, y) = \frac{1}{(2\pi)[\det(R_2)]^{1/2}} \exp\left(-\frac{1}{2}(x - \bar{\mathbf{X}})^T R_2^{-1}(x - \bar{\mathbf{X}})\right)$$

where $\det(\cdot)$ represents determinant and $x \triangleq [x, y]^T$. In general, $X_1, \dots, X_n$ are called jointly Normal if their joint p.d.f is given by

$$f_{\mathbf{X}}(x) = \frac{1}{(2\pi)^{n/2}[\det(R)]^{1/2}} \exp\left(-\frac{1}{2}(x - \bar{\mathbf{X}})^T R^{-1}(x - \bar{\mathbf{X}})\right)$$

where $\mathbf{X} \triangleq [X_1, \dots, X_n]$, $x \triangleq [x_1, \dots, x_n]^T$ and $R \triangleq Cov(\mathbf{X}, \mathbf{X})$ is its covariance matrix.

Now, for the two-dimensional case, choose some values of the three parameters (the σ's have to be positive and $-1 \leq \rho \leq 1$) and plot this density as a 2-D surface (f as a function of x and y).