Quiz 1, ECE7252, February 27, 2008

1. For a bivariate normal random vector, $\mathbf{u} = [x, y]^t$, with $\mu = [\mu_x, \mu_y]^t$ and $\Sigma = [\mathbf{v}_1, \mathbf{v}_2]$, with $\mathbf{v}_1 = [\sigma_x^2, \rho\sigma_x\sigma_y]^t$ and $\mathbf{v}_2 = [\rho\sigma_x\sigma_y, \sigma_y^2]^t$. With the constraints that σ_x and σ_y being positive, and $|\rho| < 1$, then the joint probability density function, $f(x, y)$, is simply:

$$f(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}} \exp\left[-\frac{1}{2(1-\rho^2)}\left\{\left(\frac{x-\mu_x}{\sigma_x}\right)^2 + 2\rho\left(\frac{x-\mu_x}{\sigma_x}\right)\left(\frac{y-\mu_y}{\sigma_y}\right) + \left(\frac{y-\mu_y}{\sigma_y}\right)^2\right\}\right], \tag{1}$$

(a) find the marginal probability density function, $f(x)$, of random variable, X; (b) find the conditional probability density function, $f(x|y)$, of X given $Y = y$; (c) derive the conditional mean, $\mu_{x|y}$, and variance, $\sigma_{x|y}^2$, in (b) above; (d) compare $\sigma_{x|y}^2$ and σ_x^2, what can you conclude? Why?

2. Given a Gaussian random variable x which has an unknown mean μ and known variance σ^2, it is clearly that the maximum likelihood estimate of μ given a set of n iid sample $X = \{x_1, ..., x_i, ..., x_n\}$ is the sample mean (in HW2). If we assume that the prior density of μ is also Gaussian with a known mean ν and a known variance τ^2, do the following:

(a) derive the posterior probability density, $f(\mu|X)$ of μ, conditioing on observing X; (b) the specific μ that maximizes $f(\mu|X)$ is called the maximum a posterior (MAP) estimate (not ML). derive the MAP estimate $\hat{\mu}_{MAP}$, and show that it is a weighted sum of the sample mean $\bar{x}$ and the prior mean ν [Note: you can use the same technique when solving for the ML estimate, $\mu_{ML} = \bar{x}$]; (c) consider some special cases when n is large or $n = 0$ how does μ_{MAP} relate to $\bar{x}$ or ν? Are the relations intuitively sound? Sate your reasons. This estimation framework is called Bayesian which is widely used in practice.

3. Consider a linear regression model, $y = X_1\beta_1 + X_2\beta_2 + \epsilon$, in which β_1 and β_2 can be vectors. A researcher is interested in estimating β_1. She suggests the following three-step procedure: (1) estimating b_2 in the model: $y = X_2b_2 + e_1$; (2) calculating $z = y - X_2\hat{b}_2$; (3) estimating the least square estimate for b_1 in the model: $z = X_1b_1 + e_2$ [Note: $X_1^TX_2$ is well-defined]. Do the following: (a) Express the least square estimate of β_1, $\hat{b}_1$, in the

original model in terms of X_1 and X_2; (b) Under what condition is $\hat{b}_1$ an unbiased estimate of β_1, i.e. $E[\hat{b}_1] = \beta_1$ using the above 3-step procedure?

4. A causal linear time-invariant (LTI) system has a system function (with an all-pole LPC filter of order $p = 3$):

$$H(z) = \frac{1}{1 - 0.875z^{-1} + 0.75z^{-2} - 0.25z^{-3}} = \frac{1}{A(z)}. \tag{2}$$

   Use backward iteration of the Durbin (Levinson) recursion (Slide 18 of Lecture 9) to determine whether or not the system is stable by solving for the three PARCOR coefficients, k_3, first, k_2, next, and k_1, at last. Show your work [Note: a filter is stable if $|k_i| < 1$, for all $k = 1, ..., p$. Solutions based on calculator root solvers will not count for your grade].

5. Consider Fisher discriminant analysis in which we are interested in using the within-class and between class scatter matrices, S_W and S_B, to maximize the following ratio:

$$J_F(w) = \frac{w^T S_B w}{w^T S_W w}. \tag{3}$$

   (a) Show that $J_F(w)$ is maximized by w^* such that $S_W^{-1} S_B w^* = J(w^*) w^*$, i.e. w^* is an eigenvector of S_B, and $J(w^*)$ is a corresponding eigenvalue; (b) If we formulate the above problem by constrained optimization with a Lagrange formulation to solve for a weight vector a so as to maximize the following ratio (as in Exercise 4.1):

$$J_L(a) = \frac{a^T S_B a}{a^T S_W a}, \tag{4}$$

   under the constraint that $a^T S_W a = 1$. Do we arrive at the same generalized eigenvalue problem as above in (a)? If so, what are the corresponding eigenvalue and eigenvector here. How are the two sets of eigenvalues and eigenvectors related?