EECE 5698

Homework 2: Due Monday, Jan. 27, 2014

Reading:

- Appendices A.1, A.2, A.3, A.4, A.5
- Notes, Chapter 2.1-2.7

Problems from Chapter 2 of Book:
Undergraduate Students are required to solve only 2, 4c and 10.
Graduate Students need to solve 2, 4c, 10, 11, 13 and 23f.

2. Suppose two equally probable one-dimensional densities are of the form $p(x|\omega_i) \propto e^{-|x-a_i|/b_i}$ for $i = 1, 2$ and $0 < b_i$.

(a) Write an analytic expression for each density, that is, normalize each function for arbitrary a_i and positive b_i.

(b) Calculate the likelihood ratio as a function of your four variables.

(c) Sketch a graph of the likelihood ratio $p(x|\omega_1)/p(x|\omega_2)$ for the case $a_1 = 0$, $b_1 = 1$, $a_2 = 1$ and $b_2 = 2$.

4. Consider the minimax criterion for a two-category classification problem.

(c) Assume we have one-dimensional Gaussian distributions $p(x|\omega_i) \sim N(\mu_i, \sigma_i^2)$, $i = 1, 2$, but completely unknown prior probabilities. Use the minimax criterion to find the optimal decision point x^* in terms of μ_i and σ_i under a zero-one risk.

10. Consider the following decision rule for a two-category one-dimensional problem: Decide ω_1 if $x > \theta$; otherwise decide ω_2.

(a) Show that the probability of error for this rule is given by

$$P(error) = P(\omega_1) \int_{-\infty}^{\theta} p(x|\omega_1)\, dx + P(\omega_2) \int_{\theta}^{\infty} p(x|\omega_2)\, dx.$$

(b) By differentiating, show that a necessary condition to minimize $P(error)$ is that θ satisfies

$$p(\theta|\omega_1)P(\omega_1) = p(\theta|\omega_2)P(\omega_2).$$

(c) Does this equation define θ uniquely?

(d) Give an example where a value of θ satisfying the equation actually *maximizes* the probability of error.

------------ Required only for Graduate Students --

11. Suppose that we replace the deterministic decision function $\alpha(\mathbf{x})$ with a *randomized rule*, namely, one giving the probability $P(\alpha_i|\mathbf{x})$ of taking action α_i upon observing $\mathbf{x}$.

(a) Show that the resulting risk is given by

$$R = \int \left[\sum_{i=1}^{a} R(\alpha_i|\mathbf{x}) P(\alpha_i|\mathbf{x}) \right] p(\mathbf{x})\, d\mathbf{x}.$$

(b) In addition, show that R is minimized by choosing $P(\alpha_i|\mathbf{x}) = 1$ for the action α_i associated with the minimum conditional risk $R(\alpha_i|\mathbf{x})$, thereby showing that no benefit can be gained from randomizing the best decision rule.

(c) Can we benefit from randomizing a suboptimal rule? Explain.

13. In many pattern classification problems one has the option either to assign the pattern to one of c classes, or to *reject* it as being unrecognizable. If the cost for rejects is not too high, rejection may be a desirable action. Let

$$\lambda(\alpha_i|\omega_j) = \begin{cases} 0 & i = j \qquad i, j = 1, \ldots, c \\ \lambda_r & i = c+1 \\ \lambda_s & \text{otherwise,} \end{cases}$$

where λ_r is the loss incurred for choosing the $(c+1)$th action, rejection, and λ_s is the loss incurred for making any substitution error. Show that the minimum risk is obtained if we decide ω_i if $P(\omega_i|\mathbf{x}) \geq P(\omega_j|\mathbf{x})$ for all j and if $P(\omega_i|\mathbf{x}) \geq 1 - \lambda_r/\lambda_s$, and reject otherwise. What happens if $\lambda_r = 0$? What happens if $\lambda_r > \lambda_s$?

23.

(f) Prove that a general whitening transform $\mathbf{A}_w = \mathbf{\Phi}\mathbf{\Lambda}^{-1/2}$ when applied to a Gaussian distribution ensures that the final distribution has covariance proportional to the identity matrix $\mathbf{I}$.