EECS659
Final Exam
Due Wed. Dec. 19, 2001

This take home exam is open book and open notes and you can use any sources that you please. However you are required to complete the exam by yourself, you are not allowed to collaborate in any way with anyone. Please answer questions as completely as possible to obtain full credit. Provide copies of your matlab or other routines you wrote in addition to their output. If you feel that additional assumptions need to be made to answer any part of a question state your assumption explicitly. Please make sure that your name, social security number, and signed honor code are on your exam, and if not using a blue book, make sure all of your pages are stapled in the correct order before handing in.

1. [33] **RLS/LMS via MLE**: Assume that the primary signal has the representation $Y_k = W^T X_k + N_k$ in terms of the vector of p weights $W = [W_1, \dots, W_p]^T$, the reference signal vector X_k, and the residual noise N_k. Fix k and assume that N_l is independent Gaussian with zero mean and variance $\sigma_l^2 = \lambda^{l-k}$, $l = 1, 2, \dots, k$, $\lambda \in [0, 1]$.

   (a) Derive the maximum likelihood estimator $\hat{W}_k$ of W based on past and present observations $\{(X_i, Y_i)\}_{i=1}^k$ and show that the RLS algorithm gives an exact update $\hat{W}_k \to \hat{W}_{k+1}$.

   (b) Now assume that W has an *a priori* Gaussian distribution with mean vector μ_W and covariance matrix R_W. Find the MAP estimate of W, give an exact update algorithm $\hat{W}_k \to \hat{W}_{k+1}$, and compare to standard RLS. Comment on the behavior of your algorithm for $R_W \to 0$ and $R_W \to \infty$.

   (c) Under what model for N_k does the LMS algorithm result from the approach of (a)? Using this model, derive an analog to LMS for random W via the approach of (b).

2. [33] **Robust RLS/LMS via MLE**: Under the same assumptions as in Problem 1, we now assume that the residual N_k is i.i.d. with a heavy tailed Student-t distribution, i.e the density of N_k is $p_{N_l}(z) = \alpha \left(1 + (z/\lambda^{l-k})^2\right)^{-(\lambda^{l-k}+1)/2}$, where α is a normalizing constant depending on λ. This formulation is motivated by the desire to reduce the influence of large residual errors on the weight update algorithm.

   (a) Write down the likelihood function for W based on $\{(X_i, Y_i)\}_{i=1}^k$. Compare this to the Gaussian likelihood function of Problem (1a) when the residuals $Y_i - W^T X_i$ are small, $i = 1, \dots, k$. Compare the influence of an outlier (large residual) on this likelihood vs the Gaussian likelihood of (1a) ? Do your results make sense?

   (b) Derive an iterative maximum likelihood algorithm for W using steepest descent (gradient search), Fisher scoring, or other favorite algorithm, and compare to the RLS and LMS recursions.

   (c) Use Matlab or other routine to simulate the algorithm performance and compare to RLS/LMS.

3. [33] **Generalized Linear Models**: As in class you are given that the primary i.i.d. signal Y_k has a specified distribution with parameter θ. From this you are to find the linkage function $g(\theta)$ and the corresponding non-linear representation: $Y_k = g^{-1}(W^T X_k) + N_k$ for the primary signal in terms of a linear combination $W^T X_k$ of the reference signal components. The optimal weight vector W is then obtained by maximizing the log-likelihood expressed in terms of $W^T X_k$. For each of the following find the linkage function and give an iterative procedure for finding W:

   (a) Y_k follows a Gaussian distribution with density $p_Y(y) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{1}{2}(y-\theta)^2\right), \quad -\infty < y < \infty$.

   (b) Y_k follows a geometric distribution with mass function $p_Y(y) = P(Y = y) = \frac{1}{1-\theta}\theta^y, \quad y = 0, 1, 2 \ldots$. Compare to the Poisson example derived in class.

   (c) Y_k follows an exponential distribution with density $p_Y(y) = \theta e^{-\theta y}$, $y \geq 0$. Compare to the result of (a) and (b).