Homework #2, EE556, Fall 2018

Due: 9/13/18; Please hand in Problems 1,3, and 6.

Problem #1

Prove that the decision regions of a linear machine are convex by showing that if $\underline{x_1} \in R_i$ and $\underline{x_2} \in R_i$, then $\lambda\underline{x_1} + (1-\lambda)\underline{x_2} \in R_i$.

Problem #2

i) Using the method of Lagrange multipliers (for constrained optimization), show that the distance from a point $\underline{x_1}$ to the closest point on the hyperplane $\underline{w}^T\underline{x} + w_0 = 0$ is $|g(\underline{x_1})|/||\underline{w}||$.
ii) Determine an equation specifying the location of this nearest point on the hyperplane.

Problem #3

Consider the multiclass pairwise linear discriminant approach where discriminants separate one class from all others. Give an example in two dimensions where this approach fails to separate classes, even though all classes are pairwise linearly separable.

Problem #4

Consider a d-dimensional binary vector $\underline{x}$ with entries that are either 0 or 1. Suppose we assign $\underline{x}$ to class 1 if the number of ones in the vector is odd, and to class 2 otherwise (this is the d-bit parity problem).

a) Show that these classes are not linearly separable if $d > 1$.

b) Come up with a method that uses *multiple* linear discriminants and solves this problem. For each class, the largest discriminant function output, amongst the set of discriminants belonging to that class, is taken as the class's discriminant function value. The values from each of the two classes are then compared to select the winning class. This is called a "piecewise linear" discriminant function. Show that this approach can be used to solve the parity problem and specify all the discriminant functions and how they are used in the decisionmaking.

Problem #5

Give a 1-D example (depicted in 2-D weight space) involving 2 classes and 3 data points where the patterns are not linearly separable.

Problem #6

Consider the problem of maximum likelihood estimation for the mean vector, μ, and covariance matrix, Σ, of a multivariate Gaussian density. Using the gradient results given in class, derive the ML estimate for μ and Σ.

Problem #7

In class we derived the ML estimates for the mean and variance of a univariate Gaussian density. Verify that this solution is, indeed, a maximum of the likelihood function. To do this: i) compute the Hessian matrix of second order partial derivatives; 2) evaluate the Hessian at the ML solution; 3) show that the resulting matrix is negative definite.