UNIVERSITY OF CALIFORNIA, SANTA BARBARA
Department of Electrical and Computer Engineering

ECE 277B | Fall 2007 | Instructor: Ken Rose

Homework Assignment #5

(Due on Wednesday 11/28/2007)

Problem # 1. In class we defined linear separability for the two-category case and explored some possible ways to extend it to the multiple-category case.

- "Total linear separability" (TLS): each training subset $\mathcal{X}_i$ is linearly separable from the remainder of the training set $\mathcal{X} - \mathcal{X}_i$.
- "Pairwise linear separability" (PLS): each pair of subsets $(\mathcal{X}_i, \mathcal{X}_j)$ is linearly separable.
- "Official linear separability" (OLS): there exists a set of linear discriminant functions $g_i(\mathbf{x}) = \mathbf{a}_i^t\mathbf{x} + b_i, i = 1, \ldots, c$ such that $\mathbf{x} \in \mathcal{X}_i$ *iff* $g_i(\mathbf{x}) > g_j(\mathbf{x}), \forall j \neq i$.

Establish the relative strength of these definitions by verifying the logical relations between them, i.e., determine whether or not "TLS $\Rightarrow$ OLS", "OLS $\Rightarrow$ TLS", "PLS $\Rightarrow$ TLS", etc., are true. *Justify!*

Problem # 2. We design linear classifiers for feature vectors whose components are linearly dependent, i.e., there exist a vector α and scalar β such that

$$\alpha^T\mathbf{x} + \beta = \alpha_1 x_1 + \cdots + \alpha_l x_l + \beta = 0.$$

a) It is tempting to design a simpler classifier by reducing the dimensionality. Specifically, we wish to discard one of the components of $\mathbf{x}$. Can this be done without risk of performance degradation? (Either prove such a claim, or disprove it by a counter example.)

b) If $\mathcal{X}$ is a training set of N vectors in $\mathcal{R}^l$ satisfying the linear relation $\alpha^T\mathbf{x} + \beta = 0$, but otherwise in general position, then how many linear dichotomies of $\mathcal{X}$ are there?

Problem # 3. More on linear separability.

a) Let $\mathcal{X}$ be a set of labeled feature vectors, which is *not* linearly separable. We consider an arbitrary linear classifier a and divide $\mathcal{X}$ into two subsets: $\mathcal{X}_c$ is the subset of vectors that were correctly classified, and $\mathcal{X} - \mathcal{X}_c$ is the subset of misclassified vectors. For each of the subsets determine if it is linearly separable, not linearly separable, or could be either depending on the specific example.

b) Given two linearly separable sets (i.e., each is linearly separable), is their intersection necessarily linearly separable? How about their union?

c) Let $\mathcal{Y}$ be a linearly separable set with $\mathcal{Y}_1$ and $\mathcal{Y}_2$ its subsets corresponding to the two labels. We create new set $\mathcal{Z}_1$ that is populated by various weighted averages of sample vectors from $\mathcal{Y}_1$, i.e., vector $\mathbf{z}$ may be included in $\mathcal{Z}_1$ if

$$\mathbf{z} = \sum_{\mathbf{y}_i \in \mathcal{Y}_1} \alpha_i \mathbf{y}_i,$$

for some choice of parameters $\alpha_i \geq 0$ such that $\sum \alpha_i = 1$. All the elements of $\mathcal{Z}_1$ are of course labeled with category 1. Similarly we construct a set $\mathcal{Z}_2$ populated by arbitrary weighted averages of elements of $\mathcal{Y}_2$. The question is whether $\mathcal{Z} = \mathcal{Z}_1 \cup \mathcal{Z}_2$ is guaranteed to be linearly separable.