Algorithms for Weak and Wide Separation of Sets

Michael E. Houle

School of Computer Science, McGill University, McConnell Engineering Building, 3480 University Street, Montréal, Québec, Canada, H3A 2A7

Abstract

Much attention has been given to the problem of finding specific types of separating surfaces for sets of points. In the pattern recognition setting of two-category point classification, linear separators are often sought for use as discriminant functions. Many strategies for obtaining these separators exist. However, it often happens that the two point sets cannot be separated by a simple hyperplane. In these instances, higher-order surfaces such as hyperspheres or hypercones are sometimes considered as candidate separators, or linear discriminant functions based on statistical considerations are employed. However, these approaches do not concern themselves with geometric alternatives to exact linear separators in the event that none can be found. One such alternative, explored in this paper, is to use discriminant functions that correctly classify the greatest number of points of the two sets.

Let R and G be disjoint finite sets of points in E^d. Let P be the union of R and G, and let n be the number of points of P. A hyperplane h can be said to ***partition*** P into two ***semi-spaces***, $P_1 = P \cap h_1$ and $P_2 = P \cap h_2$, if h_1 and h_2 are the two open half-spaces defined by h, and h does not contain any point of P. Similarly, the same hyperplane h also partitions R into point sets R_1 and R_2, and G into G_1 and G_2, such that $P_1 = R_1 \cup G_1$ and $P_2 = R_2 \cup G_2$. We shall call the sets $R_1 \cup G_2$ and $R_2 \cup G_1$ the (linearly separable) ***components*** of R and G with respect to h, as each of these point sets is linearly separable with the hyperplane h as separator. A component of maximum size we call a ***maximal*** component; one of minimum size we call a ***minimal*** component.

With these definitions, if hyperplane h is used as a linear discriminant function for R and G, then h correctly classifies exactly one of the two components of R and G with respect to it. If R and G are not strictly linearly separable, it might be useful to consider as a discriminant function a hyperplane such that the total number of misclassified points is minimized; that is, a hyperplane such that the components of R and G are of maximum and minimum size. Such a hyperplane we call a ***weak*** strict linear separator for R and G. In contrast, a hyperplane that misclassifies no points of R and G we call a ***strong*** strict linear separator.

Some of the well-known methods for obtaining linear separators, such as straightforward linear programming, too often yield extreme separators whose effectiveness as a discriminant function is diminished. One might prefer instead a separator that does not approach the points it separates. Accordingly, if a separator h is such that the minimum orthogonal distance to the points of its maximal component is maximized, then we shall say that h is a ***widest*** separator.

In this paper, algorithms shall be presented for finding various types of strict linear separators for point sets and sets of hyperspheres. The results are summarized in the following table:

Problem	*Object Class*	*Time*	*Space*
Strong Linear Separation	Hyperspheres	$O(n)$	$O(n)$
Weak Linear Separation	Hyperspheres	$O(n^{d+1})$	$O(n^{d+1})$
	Points	$O(n^d)$	$O(n)$
Widest Strong Linear Separation	Points	$O(n)$	$O(n)$
Widest Weak Linear Separation	Points	$O(n^{d+1})$	$O(n^{d+1})$