
Escaping the Convex Hull with Extrapolated Vector Machines.

Patrick Haffner

AT&T Labs-Research, 200 Laurel Ave, Middletown, NJ 07748
haffner@research.att.com

Abstract

Maximum margin classifiers such as Support Vector Machines (SVMs) critically depends upon the convex hulls of the training samples of each class, as they implicitly search for the minimum distance between the convex hulls. We propose Extrapolated Vector Machines (XVMs) which rely on extrapolations outside these convex hulls. XVMs improve SVM generalization very significantly on the MNIST[7] OCR data. They share similarities with the Fisher discriminant: maximize the inter-class margin while minimizing the intra-class disparity.

1 Introduction

Both intuition and theory[10] seem to support that the best linear separation between two classes is the one that maximizes the margin. But is this always true? In the example shown in Fig.(1), the maximum margin hyperplane is W_0 ; however, most observers would say that the separating hyperplane W_1 has better chances to generalize, as it takes into account the expected location of additional training samples. Traditionally, to take this into account, one would estimate the distribution of

