Qualitative Robustness of Bootstrap Approximations for Kernel Based Methods

Andreas Christmann
University of Bayreuth, Bayreuth, Germany
andreas.christmann@uni-bayreuth.de

Matias Salibian-Barrera*
The University of British Columbia, Vancouver, Canada matias@stat.ubc.ca

Stefan Van Aelst
Ghent University, Gent, Belgium Stefan.VanAelst@UGent.be

Support vector machines (SVMs) are a very popular method in modern statistical learning theory and practice, where they are typically used for classification and regression purposes. SVMs can be thought as penalized M-estimators, and in the last decade their robustness properties have been studied. For example, it is know that if the loss function is Lipschitz continuous and the kernel is bounded, the resulting SVMs have a bounded influence function, bounded maxbias, and are qualitatively robust. Although there are many theoretical results available dealing with the consistency of SVMs and their rate of convergence, less is known about their asymptotic distribution and how to estimate it.

The bootstrap (Efron, 1979) provides a consistent estimator for the distribution of a wide range of statistics. Recently it has been shown that this is also the case for SVMs. Here we study the robustness properties of these bootstrap distribution estimators for support vector machines. More specifically, we show that if T is an estimator based on a continuous operator from the space of probability measures over a compact metric space into a complete separable metric space, then bootstrap approximations for the distribution of T are stable, in the sense of being qualitatively robust. Intuitively, this means that the bootstrap distribution estimates are not severely affected by the presence of outliers in the data.

Key Words: Kernel-based methods, Support Vector Machines, Bootstrap, Robustness