



Bias-corrected support vector machine with Gaussian kernel in high-dimension, low-sample-size settings

Yugo Nakayama¹ · Kazuyoshi Yata² · Makoto Aoshima²

Received: 28 September 2018 / Revised: 15 April 2019 / Published online: 15 July 2019
© The Institute of Statistical Mathematics, Tokyo 2019

Abstract

In this paper, we study asymptotic properties of nonlinear support vector machines (SVM) in high-dimension, low-sample-size settings. We propose a bias-corrected SVM (BC-SVM) which is robust against imbalanced data in a general framework. In particular, we investigate asymptotic properties of the BC-SVM having the Gaussian kernel and compare them with the ones having the linear kernel. We show that the performance of the BC-SVM is influenced by the scale parameter involved in the Gaussian kernel. We discuss a choice of the scale parameter yielding a high performance and examine the validity of the choice by numerical simulations and actual data analyses.

Keywords Geometric representation · HDLSS · Imbalanced data · Radial basis function kernel

We are very grateful to the associate editor and the reviewer for their constructive comments. The research of the second author was partially supported by Grant-in-Aid for Scientific Research (C), Japan Society for the Promotion of Science (JSPS), under Contract Number 18K03409. The research of the third author was partially supported by Grants-in-Aid for Scientific Research (A) and Challenging Research (Exploratory), JSPS, under Contract Numbers 15H01678 and 17K19956.

✉ Makoto Aoshima
aoshima@math.tsukuba.ac.jp

Yugo Nakayama
n-yougo@math.tsukuba.ac.jp

Kazuyoshi Yata
yata@math.tsukuba.ac.jp

¹ Graduate School of Pure and Applied Sciences, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8571, Japan

² Institute of Mathematics, University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Ibaraki 305-8571, Japan