Penalized Bregman Divergence for Large Dimensional Regression and Classification

by

Chunming Zhang University of Wisconsin Department of Statistics Madison, WI 53706 cmzhang@stat.wisc.edu

Abstract

Regularization methods are characterized by loss functions measuring data fits and penalty terms constraining model parameters. The commonly used quadratic loss for linear models is not suitable for classification with binary responses, whereas the log-likelihood function for likelihood models is not readily applicable to other models, such as the quasi-likelihood model, where the exact distribution of observations is unknown or not fully specified. To broaden the scope of regularization, we introduce the penalized-BD by replacing the negative log-likelihood in the conventional penalized-likelihood with a broad and important class of Bregman divergence (BD), which encompasses nearly all of the commonly used loss functions in the regression analysis, classification procedures and machine learning literature. We investigate new statistical properties of the class of penalized-BD estimators with the number p_n of parameters diverging or even nearly comparable with the sample size n, and develop statistical inference tools. It is shown that the penalized-BD estimator, combined with appropriate penalties, achieves the same oracle property as the penalized-likelihood estimator, but asymptotically does not rely on the complete specification of the underlying distribution. Furthermore, the choice of loss functions in the penalized-BD classifiers has an asymptotically relatively negligible impact on classification performance. We illustrate the proposed penalized quasi-likelihood regression and penalized-BD for classification with simulation evaluation and real data application.