



KATHOLIEKE UNIVERSITEIT LEUVEN

Statistics Seminar

Joint organization statistics research groups Faculty of Science and Faculty of Economics and Applied Economics
Leuven Statistics Research Centre (LSTAT)

Professor Dr Ricardo Maronna

University of La Plata and CICPBA, Argentina

“Robust ridge regression”

Thursday October 22, 2009
12:00—13:00

Location: Room B00.05, Department of Mathematics, Celestijnenlaan 200B, Heverlee.
Supporting research project: GOA-project 2007/04

Abstract: see next page.

Robust ridge regression
Ricardo Maronna
University of La Plata and CICPBA

Ridge regression, being based on the minimization of a quadratic loss function, is sensitive to outliers. Current proposals for robust ridge regression estimates are sensitive to “bad leverage observations”, cannot be employed when the number of predictors p is larger than the number of observations n , and have a low robustness when the ratio p/n is large. In this work a penalized regression MM estimate for the model

$$y_i = \beta_0 + \mathbf{x}'_i \beta_1 + e_i$$

is proposed, of the form

$$\hat{\beta} = (\hat{\beta}_0, \hat{\beta}_1) = \arg \min_{\beta} \left\{ \hat{\sigma}_{\text{ini}}^2 \sum_{i=1}^n \rho \left(\frac{r_i(\beta)}{\hat{\sigma}_{\text{ini}}} \right) + \lambda \|\beta_1\|^2 \right\},$$

where $r_i(\beta) = y_i - \beta_0 - \mathbf{x}'_i \beta_1$ are the residuals and $\hat{\sigma}_{\text{ini}}$ is the residual scale from an initial estimate, and ρ is a bounded loss function.

The initial estimate is a penalized S estimate which obtained by a modification of the Peña-Yohai deterministic procedure. The penalty λ is chosen by cross-validation in order to minimize a robust estimate of prediction error. A fast iterative algorithm is proposed. The MM estimate can be computed for $p > n$ and is robust for large p/n .

The advantages of the proposed approach over its competitors are demonstrated through both simulated and real data.