L_1-*Statistical Procedures and Related Topics*
IMS Lecture Notes – Monograph Series (1997) Volume 31

Exact algorithms for computing the least median of squares estimate in multiple linear regression

José Agulló

Universidad de Alicante, Spain

Abstract: We propose two finite algorithms to compute the exact least median of squares (LMS) estimates of parameters of a linear regression model with p coefficients. The first algorithm is similar to Stromberg's (1993) exact algorithm. It is based on the exact fit to subsets of p cases and uses impossibility conditions to avoid unnecessary calculations. The second one is based on a branch and bound (BAB) technique. Empirical results suggest that the proposed algorithms are faster than the finite exact algorithms described earlier in the literature.

Key words: Branch and bound, exact algorithms, high breakdown regression, least median of squares, robust regression

AMS subject classification: 62F35, 62J05, 65U05.

1 Introduction

In this paper we consider the multiple linear regression model

$$Y = Z\theta + \epsilon, \tag{1}$$

where Y is an $n \times 1$ vector of dependent variables, θ is a $p \times 1$ vector of unknown parameters, Z is an $n \times p$ design matrix of predictors, and ϵ is an $n \times 1$ vector of true residuals. We denote the ith component of Y and the ith row of Z by y_i and z_i^t, respectively. We suppose that the design matrix Z is fixed and has rank p. Sometimes we also assume that any $p \times p$ submatrix of Z is nonsingular; in this case we say that Z verifies the *Haar condition*, or that the observations are in *general position*. An estimate of θ, say $\hat{\theta}$, gives n residuals $e_i(\hat{\theta}) = y_i - z_i^t\hat{\theta}$. The most well-known