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

# Properties of $L^1$ residuals

**Stephan Morgenthaler**

*Swiss Federal Institute of Technology, Lausanne, Switzerland*

*Abstract*: The paper discusses the behavior of residuals from least-absolute-deviations (or $L^1$) fits of linear models. Particular emphasis is given to data arising by way of designed experiments. The paper argues that the $L^1$ method of fitting such models should be discouraged. The method is inefficient when compared to other robust methods while not being any simpler to compute. The residuals obtained by $L^1$ fitting exhibit several weaknesses. First of all they are ambiguous in the sense that there are a multitude of $L^1$ fits, sometimes quite far apart. Second, typical algorithms produce as many exact zero residuals as there are contrasts fitted. As a result, the non zero residuals do not give an accurate reflection of the errors that occurred during the experimental runs.

*Key words*: Least absolute deviations, factorial designs, outliers detection, uniqueness of fit.

AMS subject classification: 62K15, 62J05.

## 1 Introduction

Let $y = X\theta + \varepsilon$ be a linear model with uncorrelated, centered, and homoskedastic errors $\varepsilon_1, \ldots, \varepsilon_n$. As indicated, we take $n$ to be the number of observations, whereas $p$ denotes the dimension of the regression parameter $\theta$. The least-squares residuals are $r = (I - H)y$, where $\hat{y} = Hy = X(X^TX)^{-1}X^Ty$ is the least-squares fit. If the error distribution has two moments, it follows that $E(r) = X\theta - HX\theta = 0$, and $Var(r) = \sigma^2(I - H)$. The use of such residuals for outlier detection and other diagnostic purposes has been explored in great detail in the statistical literature (see for example Belsley, Kuh and Welsch, 1980; Cook and Weisberg, 1982).

Residuals from an $L^1$ fit are not so easily described. Throughout this article, we denote by $\tilde{\theta}$ a parameter fit obtained by minimizing the least-