3

Adaptive LAD + LS Regression

3.1 Introduction

Arthanari and Dodge (1981) introduced an estimation method in the linear model based on a convex combination of a least squares and of a least absolute deviations estimators with a fixed weight $0 \leq \delta \leq 1$. They also provided two algorithms using a mathematical programming approach for finding estimates in linear regression. However, in optimizing such a convex combination, the experimenter is required to fix the value of $\delta$ or vary it at different values up to complete satisfaction in an ad hoc fashion.

Adaptive combination of least squares and the least absolute deviations estimators was first introduced by Dodge and Jurečková (1987). They showed that the choice of $\delta$ could be adapted in order to achieve the minimum asymptotic variance over a scale family of densities. In the present chapter, we focus on this adaptive linear regression for estimation of parameters.

Two historical real data sets will be examined in this chapter to show the effectiveness of the proposed methods.
3.2 Convex Combination of LAD and LS Regressions

Suppose the errors have the distribution with density \( f(z) \) of the following form:

\[
f(z) = (1 - \delta) \frac{1}{\sqrt{2\pi}} e^{-z^2/2} + \delta \frac{1}{2\sigma} e^{-|z|/\sigma},
\]

where \( 0 \leq \delta \leq 1 \) and \( \sigma > 0 \). We can regard \( f(z) \) as a contaminated normal distribution. When \( \delta = 0 \), \( f(z) \) is the usual normal distribution with mean zero and variance one. When \( \delta = 1 \), \( f(z) \) is Laplace (double exponential) with standard deviation \( \sigma \).

As an alternative to least squares, we consider the minimization of the convex combination of criteria,

\[
G(\beta) = (1 - \delta) \|Y - X\beta\|_2^2 + \delta \|Y - X\beta\|_1,
\]

where \( 0 \leq \delta \leq 1 \). When \( \delta = 0 \), we have the usual least squares method. When \( \delta = 1 \), the problem becomes the LAD minimization problem. This direct combination of least squares and the least absolute deviations leads to the following \( \rho \) and \( \psi \):

\[
\rho(x) = (1 - \delta)x^2 + \delta |x|
\]

and

\[
\psi(x) = \begin{cases} 
2(1 - \delta)x - \delta & x < 0 \\
2(1 - \delta)x + \delta & x > 0,
\end{cases}
\]

where \( 0 \leq \delta \leq 1 \).

Minimizing the convex combination of LS and LAD by means of mathematical programming can be written as follows:

\[
\text{minimize } (1 - \delta)z'z + \delta \sum_{i=1}^{n} z_i \\
\text{subject to } -z \leq Y - X\beta \leq z, \\
\beta \text{ unrestricted in sign, } z \geq 0,
\]

where \( Y, X, \beta, \) and \( z \) are defined as in (1.1). The optimal solution to the problem and methods for finding the regression coefficients based on the modification of the simplex method are available. This procedure is robust against small deviations from the normal distribution. However, since the values of \( \delta \) are fixed in advance and need to be chosen by the user, it is difficult to achieve good properties for the estimator obtained in this way.