Summary some fifty years before the least sum of squared residuals fitting procedure was published in 1805, Boscovich (or Bošković) proposed an alternative which minimises the (constrained) sum of the absolute residuals.

For $i = 1, 2, ..., n$, let $\{x_{i1}, x_{i2}, ..., x_{iq}, y_i\}$ represent the ith observation on a set of $q+1$ variables and suppose that we wish to fit a linear model of the form

$$y_i = x_{i1}\beta_1 + x_{i2}\beta_2 + ... + x_{iq}\beta_q + \epsilon_i$$

to these n observations. Then, for $p > 0$, the L_p-norm fitting procedure chooses values for $b_1, b_2, ..., b_q$ to minimise the L_p-norm of the residuals $[\sum_{i=1}^n |e_i|^p]^{1/p}$ where, for $i = 1, 2, ..., n$, the ith residual is defined by

$$e_i = y_i - x_{i1}b_1 - x_{i2}b_2 - ... - x_{iq}b_q.$$

The most familiar L_p-norm fitting procedure, known as the least squares procedure, sets $p = 2$ and chooses values for $b_1, b_2, ..., b_q$ to minimise the sum of the squared residuals $\sum_{i=1}^n e_i^2$.

A second choice, to be discussed in the present article, sets $p = 1$ and chooses $b_1, b_2, ..., b_q$ to minimise the sum of the absolute residuals $\sum_{i=1}^n |e_i|$

A third choice sets $p = \infty$ and chooses $b_1, b_2, ..., b_q$ to minimise the largest absolute residual $max_{i=1}^n |e_i|$.

Setting $u_i = e_i$ and $v_i = 0$ if $e_i \geq 0$ and $u_i = 0$ and $v_i = -e_i$ if $e_i < 0$, we find that $e_i = u_i - v_i$ so that the least absolute residuals (LAR) fitting problem chooses $b_1, b_2, ..., b_q$ to minimise the sum of the absolute residuals

$$\sum_{i=1}^n (u_i + v_i)$$

subject to

$$x_{i1}b_1 + x_{i2}b_2 + ... + x_{iq}b_q + U_i - v_i = y_i \quad \text{for } i = 1, 2, ..., n$$

$$\text{and} \quad U_i \geq 0, v_i \geq 0 \quad \text{for } i = 1, 2, ..., n.$$

The LAR fitting problem thus takes the form of a linear programming problem and is often solved by means of a variant of the dual simplex procedure.

Gauss has noted (when $q = 2$) that solutions of this problem are characterised by the presence of a set of q zero residuals. Such solutions are robust to the presence of outlying observations. Indeed, they remain constant under variations in the other $n - q$ observations provided that these variations do not cause any of the residuals to change their signs.

The LAR fitting procedure corresponds to the maximum likelihood estimator when the ϵ-disturbances follow a double exponential (Laplacian) distribution. This estimator is more robust to the presence of outlying observations than is the standard least squares estimator which maximises the likelihood function when the ϵ-disturbances are normal (Gaussian). Nevertheless, the LAR estimator has an asymptotic normal distribution as it is a member of Huber's class of M-estimators.

There are many variants of the basic LAR procedure but the one of greatest historical interest is that proposed in 1760 by the Croatian Jesuit scientist Rugjer (or Rudjer) Josip Bošković (1711–1787) (Latin: Rogerius Josephus Boscovich; Italian: Ruggiero Giuseppe Boscovich). In his variant of the standard LAR procedure, there are two explanatory variables of which the first is constant $x_{i1} = 1$ and the values of b_1 and b_2 are constrained to satisfy the adding-up condition $\sum_{i=1}^{n}(y_i - b_1 - x_{i2}b_2) = 0$ usually associated with the least squares procedure developed by Gauss in 1795 and published by Legendre in 1805. Computer algorithms implementing this variant of the LAR procedure with $q \geq 2$ variables are still to be found in the literature.

For an account of recent developments in this area, see the series of volumes edited by Dodge (1987, 1992, 1997, 2002). For a detailed history of the LAR procedure, analysing the contributions of Bošković, Laplace, Gauss, Edgeworth, Turner, Bowley and Rhodes, see Farebrother (1999). And, for a discussion of the geometrical and mechanical representation of the least squares and LAR fitting procedures, see Farebrother (2002).

References

Yadolah Dodge (Ed.) (1987), *Statistical Data Analysis Based on the L_1-Norm and Related Methods*, North-Holland Publishing Company, Amsterdam, The Netherlands.

Yadolah Dodge (Ed.) (1992), *L_1-Statistical Analysis and Related Methods*, North-Holland Publishing Company, Amsterdam, The Netherlands.

Yadolah Dodge (Ed.) (1997), *L_1-Statistical Procedures and Related Topics,* Institute of Mathematical Statistics, Hayward, California, USA.

Yadolah Dodge (Ed.) (2002), *Statistical Data Analysis based on the L_1-Norm and Related Methods,* Birkhäuser Publishing, Basel, Switzerland.

Richard William Farebrother (2002), *Visualizing Statistical Models and Concepts*, Marcel Dekker, New York, USA.

Reprinted with permission from Lovric, Miodrag (2011), International Encyclopedia of Statistical Science. Heidelberg: Springer Science +Business Media, LLC

Richard William Farebrother