ELSEVIER

Statistics & Probability Letters 21 (1994) 27–28

STATISTICS & PROBABILITY LETTERS

A note on min–maxbias estimators in approximately linear models

Gilbert W. Bassett Jr.[1]

University of Illinois at Chicago, Department of Economics (m\c 144), 601 S. Morgan (Rm 2103), Chicago, IL 60607-7121, USA

Received July 1993; revised September 1993

Abstract

The approximately linear model represents deviations from the ideal linear model by a vector contained in a prescribed bias-ball. In a recent paper Mathew and Nordstrom (1993) proposed min–maxbias estimators in which a criterion function is defined by maximizing errors over the bias-ball. When the Chebyshev norm defines the bias-ball they found the least absolute deviation or L_1 estimator to be identical to its maxbias version. This was thought to be a robustness property since it contrasts with least squares where the maxbias criterion is a combination of L_1 and the sum of squares. In this paper it is shown, however, that equivalence between the L_1 estimator and its min–maxbias version is not special to L_1 and that the equivalence is valid for estimates that are not robust. Hence, while the L_1 estimate does have desirable robustness properties the equivalence to its min–maxbias version cannot be counted as one of them.

Keywords: Approximately linear model; Least squares; Least absolute deviation

In the approximately linear model, $y = b + X\beta + e$, deviations from the linear model are represented by an unknown bias vector $b \in C$, which is contained in a prescribed bias-ball. In a recent paper Mathew and Nordstrom (1993) considered the case in which the bias-ball is given by $C = \{b \mid L_\infty(Db) \leqslant 1\}$, where D is a known diagonal matrix, and $L_\infty(z_1, \dots, z_n) = \max\{|z_i|: i = 1, \dots, n\}$ is the Chebyshev norm. They proposed a new class of estimators defined in terms of a criterion function that was specified by maximizing the deviations b over the bias-ball. These "min–maxbias" estimators are of the form $\hat{\beta} = \operatorname{argmin}_\theta m(y - X\theta)$, where $m(\cdot)$ is the maxbias function, $m(z) = \max\{h(W(z - b)) \mid b \in C\}$, defined for a given diagonal matrix W, and a given norm h. They initially took h to be least squares, $L_2^2(z_1, \dots, z_n) = \sum z_i^2$, and they showed that $m(z)$ is then a weighted linear combination of least squares and least absolute values,

$$m(z) = L_2^2(Wz) + 2L_1(WDz) + \text{constant},$$

where $L_1(z_1, \dots, z_n) = \sum |z_i|$. This estimator, which combines the L_1 and L_2 criterion functions, had been previously considered by Arthanari and Dodge (1981), Dodge (1984) and Dodge and Jureckova (1987).

The case in which h is the L_1 instead of the L_2 function was then considered. It was found that $m(z) = L_1(Wz) + \text{constant}$, and hence

$$\operatorname{argmin} m(z) = \operatorname{argmin} L_1(Wz).$$

[1] Email: U09006@uicvm.uic.edu.

SSDI 0167-7152(93)E0193-W

This means that the least absolute deviation estimator is identical to its min–maxbias cousin. This was favorably interpreted as another one of L_1's desirable robustness properties; for other such properties see Bassett (1991).

It is shown below that the equivalence between the L_1 estimator and its min–maxbias version however is not special to L_1 and that the equivalence is valid for estimates that are not robust. Hence, while the L_1 estimate does have desirable robustness properties the equivalence to its min–maxbias version cannot be counted as one of them.

For one example, let h be the Chebyshev norm so that

$$m(z) = \max\{\max_{i \leqslant 1 \leqslant n} \{w_i|z_i - b_i|\}: \max|b_i d_i| \leqslant 1\}.$$

The maximum is achieved at $b_i = -d^{-1}\operatorname{sgn}(z_i)$ so

$$m(z) = \max_i \{w_i|z_i| + w_i d_i^{-1}\}$$

from which it follows that

$$\operatorname{argmin} m(z) = \operatorname{argmin} L_\infty(Wz).$$

Hence, the Chebyshev estimate is also the same as its min–maxbias cousin.

For another example take $D = W = I$, let h be an arbitrary norm, and suppose that C is defined in terms of h, $C = \{b \mid h(Db) \leqslant 1\}$. Since h is a norm, $h(z - b) \leqslant h(z) + h(b)$, so

$$\begin{aligned} m(z) &= \max\{h(z-b) \mid h(b) \leqslant 1\} \\ &\leqslant \max\{h(z) + h(b) \mid h(b) \leqslant 1\} \\ &= h(z) + 1. \end{aligned}$$

On the other hand, $m(z) \geqslant h(z - b)$, for any b, $h(b) \leqslant 1$, so that at

$$b = -\frac{1}{h(z)} z,$$

we have $h(z - b) = h(z) + 1$, and as a result, $m(z) = h(z) + 1$. This says that when the bias-ball and the M-estimate are defined using the same norm, the M-estimate the min–maxbias estimate are identical. Hence, the result for L_1 is not one of its special features since it holds for other estimators.

References

Arthanari, T.S. and Y. Dodge (1981), *Mathematical Programming in Statistics* (Wiley, New York).

Bassett, G.W. (1991), Equivariant, monotone, 50% breakdown estimators, *The Amer. Statist.* 135–137.

Dodge, Y. (1984), Robust estimation of regression coefficients by minimizing a convex combination of least squares and least absolute deviations, *Comput. Statist. Quart.* **1**, 139–153.

Dodge, Y. and J. Jureckova (1987), Adaptive combination of least squares and least absolute deviations estimators, in: Y. Dodge, ed. *Statistical Data Analysis Based on L_1 Norm and Related Methods* (North-Holland, Amsterdam) 275–284.

Mathew, T. and K. Nordstrom (1993), Least squares and least absolute deviation procedures in approximately linear models, *Statist. Probab. Lett.* **16**, 153–158.