On R-Methods in Errors-in-Variables Models

Silvelyn Zwanzig, Uppsala University, Sweden, zwanzig@math.uu.se

Abstract

Rank estimators are defined as minimizer of a dispersion measure, which includes the residuals and their rank. The naive application of rank estimators to errors-in-variables models delivers biased estimators. In the paper it is shown that rank estimates based on orthogonal residuals are consistent.

Keywords: consistency, measurement errors, nuisance parameters, rank estimators

1. Introduction

Rank test statistics have the advantage that the null distribution is independent of the underlying distribution. This distribution property does not hold for rank estimators, but there is the hope that they are more robust than least squares estimators. Recently the application of rank estimation methods to errors-in-variables models got more and more attention. Sen and Saleh (2010) showed that the naive use of the Theil–Sen estimator causes the same bias as the naive least squares estimator. In Zwanzig (2012) the multivariate linear structural model is considered and a consistent rank estimator based on Kendall's tau is derived. Here we are now interested in the functional model. A correction of the Jaeckel's dispersion is proposed for defining a consistent rank estimator.

The paper is organized as follows. First we introduce rank estimates based on orthogonal residuals. In Section 3 the limit of a general rank statistic is derived. In Section 4 this result is applied to the simple linear functional model with normal errors and the consistency of a estimator minimizing this rank statistic is shown.

2. Model

Consider a simple linear functional errors-in-variables model

$$y_i = \beta_0 \xi_i + \epsilon_i,\ x_i = \xi_i + \delta_i,\ i = 1, ..., n, \tag{1}$$

where the errors are $\epsilon_i \sim N(0, \sigma^2)$ and $\delta_i \sim N(0, \sigma^2)$ mutually independent. The unknown design points $\xi_i, i = 1, ..., n$ are centered, fixed and for sufficiently large n exist positive constants δ and M, such that

$$\sum_{i=1}^{n} \xi_i = 0, \quad 0 < \delta \le \frac{1}{n}\sum_{i=1}^{n} \xi_i^2 < M, \ \ \frac{1}{n}\sum_{i=1}^{n} |\xi_i|^3 < M. \tag{2}$$

A naive R-estimator is defined by

$$\widehat{\beta}_{naive,R} \in \arg\min_{\beta} D_{naive}(\beta),\ D_{naive}(\beta) = \frac{1}{n}\sum_{i=1}^{n} (y_i - \beta x_i) a_n(R_i),$$

where $D_{naive}(\beta)$ coincides with Jaeckel's dispersion in the regression model $y_i = x_i\beta_0 + \epsilon_i$. R_i denotes the rank of the vertical residuals $y_i - \beta x_i$, the scores $a_n(i) = \varphi(\frac{i}{n+1})$ are generated by a score function $\varphi : [0, 1] \to R$, which is even, $\varphi(t) = -\varphi(1 - t)$, bounded and has bounded derivatives of first and second order φ', φ''.

A rank estimator which is more adapted to the errors-in-variables model can defined by using the orthogonal distance

$$d^2(y,x) = \min_{\xi} \left(|y-\beta\xi|^2 + |x-\xi|^2 \right) = \frac{1}{1+\beta^2} |y-\beta x|^2 .$$

Define the orthogonal residual $d(y_i, x_i)$ such that it has the same sign as the vertical residual $y_i - \beta x_i$, then both residuals have the same rank.

An orthogonal R-estimator is defined by

$$\widehat{\beta}_{orth,R} \in \arg\min_{\beta} D_{orth}(\beta), \quad D_{orth}(\beta) = \frac{1}{n} \sum_{i=1}^{n} d(y_i, x_i)\, a_n(R_i),$$

where R_i is the rank of $y_i - \beta x_i$. It holds

$$D_{orth}(\beta) = \frac{1}{\sqrt{1+\beta^2}} D_{naive}(\beta). \tag{3}$$

3. Limit of a General Rank Statistic

Consider the rank statistic

$$D = \frac{1}{n} \sum_{j=1}^{n} Z_j a_n(R_j),$$

where R_j is the rank of Z_j. The random variables Z_j, $j = 1, ..., n$, are independently distributed with continuous distribution functions F_j. Furthermore we assume $\sum_{j=1}^{n} EZ_j = 0$ and $\delta \le \frac{1}{n} \sum_{j=1}^{n} EZ_j^2 < M$ for sufficiently large n.

Theorem 1. *Under the assumptions above on the random variables and on the score function*

1. $$ED = \frac{1}{n} \sum_{j=1}^{n} E\left(Z_j \varphi(\frac{\mu_j}{n+1}) \right) + O(n^{-1}),\ \mu_j = 1 + \sum_{i=1, i\neq j}^{n} F_i(Z_j).$$

2. *For linear score functions $\varphi(t) = at + b$ it holds*

$$ED = \frac{a}{n(n+1)} \sum_{j=1}^{n} \sum_{i=1, i\neq j}^{n} E(Z_j(F_i(Z_j)).$$

Proof: We have

$$ED = \frac{1}{n} \sum_{j=1}^{n} E\left(Z_j E\left(a_n(R_j) \mid Z_j \right) \right). \tag{4}$$

The rank R_j of Z_j can be presented as sum, counting how many members of the sample $Z_1, ..., Z_n$ are less or equal to Z_j

$$R_j = \sum_{i=1}^{n} u(Z_j - Z_i),\ u(x) = \begin{cases} 1 & for \quad x \ge 0 \\ 0 & else \end{cases}. \tag{5}$$

The conditional distribution of $R_j \mid_{Z_j} -1$ is a Poisson binomial distribution with $p_i = F_i(Z_j)$, $i = 1, ..., n,\ i \neq j$ and with mean $\mu_j - 1$ and variance σ_j^2

$$\mu_j = E(R_j \mid Z_j) = 1 + \sum_{i=1, i\neq j}^{n} F_i(Z_j), \quad \sigma_j^2 = \sum_{i=1, i\neq j}^{n} F_i(Z_j)(1 - F_i(Z_j)).$$

Applying the following Taylor expansion to the scores $a_n(k)$

$$\varphi(\frac{\mu_j}{n+1}) + \frac{1}{n+1}(k-\mu_j)\varphi'(\frac{\mu_j}{n+1}) + \frac{1}{(n+1)^2}(k-\mu_j)^2 \varphi''(\frac{\mu_j}{n+1} + \theta_j)$$

we obtain

$$E\left(a_n\left(R_j\right) \mid Z_j\right) = \varphi(\frac{\mu_j}{n+1}) + rest(1),$$

with

$$rest(1) = \frac{M}{\left(n+1\right)^2}\sigma_j^2 \leq O(n^{-1}).$$

For linear score functions no Taylor expansion is needed and because of $\sum EZ_j = 0$ we get the second statement. □

The variance of D is more complicated to study. An adaption of the proof of Theorem 3.1 in Hajek (1968) does not work. The monotony properties of the conditional covariances between rank scores $a_n(R_j)$ cannot be transformed to terms $Z_j a_n(R_j)$. We give a direct proof for linear score functions.

Theorem 2. *It holds*

$$Var(\frac{1}{n(n+1)}\sum_{j=1}^{n} Z_j R_j) = O(n^{-1}).$$

Proof: We have

$$Var(\sum_{j=1}^{n} Z_j R_j) = \sum_{j_1=1}^{n}\sum_{j_2=1}^{n} Cov(Z_{j_1}R_{j_1}, Z_{j_2}R_{j_2}).$$

From (5) it follows

$$Var(\sum Z_j R_j) = \sum_{j_1=1}^{n}\sum_{j_2=1}^{n}\sum_{i_1=1}^{n}\sum_{i_2=1}^{n} Cov(Z_{j_1}u(Z_{j_1}-Z_{i_1}), Z_{j_2}u(Z_{j_2}-Z_{i_2}).$$

The sum has n^4 terms. But terms, where all indices are different have covariance zero. Thus $Var(\sum Z_j R_j)$ is of order n^3. □

Summarizing we obtain

$$\frac{1}{n(n+1)}\sum_{j=1}^{n} Z_j R_j = \frac{1}{n(n+1)}\sum_{j=1}^{n}\sum_{i=1, i\neq j}^{n} E(Z_j(F_i(Z_j)) + o_p(1). \tag{6}$$

Note, the convergence is uniformly in $\beta \in B$, B compact, for continuously parameterized distributions $F_j = F_{j,\beta}$.

4. Consistency of the Orthogonal Rank Estimate

Let us apply (6) to $D_{naive}(\beta)$ with $Z_j = y_j - \beta x_j \sim N((\beta_0 - \beta)\xi_j, (1+\beta^2)\sigma^2)$ and use (3).

Theorem 3. *Under (1),(2) and for a linear score function $\varphi(t) = t$ and for arbitrary fixed K and for some constant C*

$$\begin{aligned} D_{orth}\left(\beta\right) &= D_{lead}\left(\beta\right) + O(n^{-1}) + rest_1(\beta), \\ D_{lead}\left(\beta\right) &= \frac{1}{2\sqrt{\pi}}(1 + \frac{\left(\beta_0-\beta\right)^2}{\sqrt{1+\beta^2}\sigma}\frac{1}{n}\sum_{i=1}^{n}\xi_i^2) + rest_2(\beta), \end{aligned}$$

where $\sup_{\beta\in B}|rest_1(\beta)| = o_P(1)$, $B = \{\beta : |\beta| \leq K\}$ *and* $|rest_2(\beta)| < C\left|\beta_0 - \beta\right|^3$.

Proof: Denote by Φ the distribution function of $N(0,1)$ and by φ the density of $N(0,1)$. Note, because we suppose now a linear score function, a denotation confusion cannot

arise. It holds

$$EZ_jF_i(Z_j) = \int z\ \Phi\left(\frac{z-(\beta_0-\beta)\xi_i}{\sqrt{1+\beta^2\sigma}}\right)\frac{1}{\sqrt{1+\beta^2\sigma}}\varphi\left(\frac{z-(\beta_0-\beta)\xi_j}{\sqrt{1+\beta^2\sigma}}\right)dz.$$

Transform the variable $z=\sqrt{1+\beta^2\sigma}\ u+(\beta_0-\beta)\xi_j$ and set

$$\Delta_{(j-i)} = \frac{(\beta_0-\beta)\,(\xi_j-\xi_i)}{\sqrt{1+\beta^2\sigma}}.$$

Then $EZ_jF_i(Z_j) = \sqrt{1+\beta^2\sigma}\ A+(\beta_0-\beta)\xi_j\ B$ with

$$A=\int u\ \Phi(u+\Delta_{(j-i)})\ \varphi(u)du,\ B=\int \Phi(u+\Delta_{(j-i)})\ \varphi(u)du.$$

Applying a Taylor expansion to Φ with $\varphi'(u)=-u\varphi(u)$

$$\Phi(u+\Delta)=\Phi(u)+\Delta\varphi(u)-\frac{1}{2}\Delta^2u\varphi(u)+\frac{1}{3}\Delta^3\varphi''(u+\theta)$$

we obtain

$$A=\int u\ \Phi(u)\ \varphi(u)du+\Delta_{(j-i)}\int u\ \varphi(u)^2du-\frac{1}{2}\Delta^2_{(j-i)}\int u^2\ \varphi(u)^2du+r_A$$

and

$$B=\int \Phi(u)\ \varphi(u)du+\Delta_{(j-i)}\int \varphi(u)^2du-\frac{1}{2}\Delta^2_{(j-i)}\int\ u\ \varphi(u)^2du+r_B.$$

By partial integration we have $\int u\ \Phi(u)\ \varphi(u)du=\int\varphi(u)^2du=\frac{1}{2\sqrt{\pi}}$. Furthermore $\int u\varphi(u)^2du=0$ and $\int u^2\ \varphi(u)^2du=\frac{1}{4\sqrt{\pi}}$ and $\int\Phi(u)\ \varphi(u)du=\frac{1}{2}$. Hence $EZ_jF_i(Z_j)$ equals

$$\sqrt{1+\beta^2\sigma}\frac{1}{2\sqrt{\pi}}\ \left(1-\frac{1}{4}\Delta^2_{(j-i)}\right)+(\beta_0-\beta)\xi_j\frac{1}{2}\left(1+\frac{1}{\sqrt{\pi}}\Delta_{(j-i)}\right)+rest_2(\beta).$$

Now calculate $ED=D1+D2$ with

$$D1=-\frac{1}{n(n+1)}\sum_{i=1}^{n}E(Z_i(F_i(Z_i))$$

and

$$D2=\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}E(Z_j(F_i(Z_j))+O(n^{-1}).$$

We have

$$D1=-\frac{1}{(n+1)2\sqrt{\pi}}\sqrt{1+\beta^2\sigma}$$

and $D2$ equals

$$\sqrt{1+\beta^2\sigma}\frac{1}{2\sqrt{\pi}}\ \left(1-\frac{1}{4}\Delta^2\left(\beta\right)\right)+\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}\Delta_{(j-i)}(\beta_0-\beta)\xi_j+rest_2(\beta)+O(n^{-1})$$

with $\Delta^2\left(\beta\right)=\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}\Delta^2_{(j-i)}$. Under (2) $|rest_2(\beta)|<C\left|\beta-\beta_0\right|^3$. Using $\sum\xi_i=0$ we get

$$\Delta^2\left(\beta\right)=\left(\beta_0-\beta\right)^2\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}\left(\xi_j-\xi_i\right)^2=0.$$

and

$$\frac{1}{n^2}\sum_{j=1}^{n}\sum_{i=1}^{n}\Delta_{(j-i)}(\beta_0-\beta)\xi_j=(\beta_0-\beta)^2\,\frac{1}{n}\sum_{j=1}^{n}\xi_j^2.$$

□

Note, under symmetry assumptions the proof works also for location families, but then we will have another constant as $\frac{1}{2\sqrt{\pi}}$.

We obtain the following result.

Theorem 4. *For a rank estimator*

$$\widehat{\beta}_{orth,R}\in\arg\min_{\beta}\frac{1}{\sqrt{1+\beta^2}}\frac{1}{n}\sum_{i=1}^{n}(y_i-x_i\beta)R_i$$

where R_i denotes the rank of $y_i-x_i\beta$, it holds

$$\widehat{\beta}_{orth,R}\to\beta_0\ in\ \ probability.$$

Proof: First we show that there exist a K and $\widehat{\beta}_{orth,R}\in B=\{\beta:|\beta|\le K\}$. In Jaeckel (1972) it is shown, that $D_{naive}(\beta)$ is positive, convex and piecewise linear. Thus for sufficiently large K it exists a positive constant a and

$$D_{orth}(\beta)\ge a\frac{1}{\sqrt{1+\beta^2}}|\beta|,\quad for\ \ |\beta|>K.$$

The function $\frac{1}{\sqrt{1+\beta^2}}\beta$ is monoton increasing for $\beta>K$, hence the minimizers of $D_{orth}(\beta)$ are in B.

Second under (2) for $\beta\in B$ and $|\beta-\beta_0|>\varepsilon$, there exists a constant c_0 such for all $\varepsilon,\ 0<\varepsilon<\frac{\delta}{\sqrt{1+K^2}}\frac{1}{C}-c_0$

$$D_{lead}(\beta)-D_{lead}(\beta_0)\ge c_0\varepsilon^2.$$

Because $D_{orth}\left(\widehat{\beta}_{orth,R}\right)\le D_{orth}(\beta_0)$, we have

$$\begin{aligned}
& D_{lead}\left(\widehat{\beta}_{orth,R}\right)-D_{lead}(\beta_0)\\
\le\ & D_{lead}\left(\widehat{\beta}_{orth,R}\right)-D_{lead}(\beta_0)+D_{orth}(\beta_0)-D_{orth}\left(\widehat{\beta}_{orth,R}\right)\\
\le\ & 2\sup_{\beta\in B,|\beta-\beta_0|>\varepsilon}|D_{orth}(\beta)-D_{lead}(\beta)|.
\end{aligned}$$

Hence

$$P\left(\left|\widehat{\beta}_{orth,R}-\beta_0\right|>\varepsilon\right)\le P\left(\sup_{\beta\in B,|\beta-\beta_0|>\varepsilon}|D_{orth}(\beta)-D_{lead}(\beta)|\ge\frac{c_0}{2}\varepsilon^2\right)$$

and Theorem 3 delivers the result.

□

Note, the leading term of the naive rank estimator does not attain the minimum at β_0, what implies the inconsistency of the naive Theil–Sen estimator. This result was already shown by Sen and Saleh (2010) by another method.

References

Hajek J.(1968) "Asymptotic normality of simple linear rank statistics under alternatives," *Ann. Math.Stat.*, 39, 325-346.

Jaeckel, L.A. (1972) "Estimating regression coefficients by minimizing the dispersion of the residuals," *Ann.Math.Stat.*,43, 1449-1458.

Sen and Saleh (2010) "The Theil–Sen estimator in a measurement error perspective", in Festschrift in honor of Professor Jana Jureckova, *Inst. Math. Stat., Collect.* Vol. 7, Math. Statist. Beachwood, Ohio, 224-243.

Zwanzig, S. (2012) "On a consistent rank estimate in a linear structural model," *Tatra Mt. Math. Publ.* 51(2012),191-202.