Biometrical Letters
Vol. 45 (2008), No. 2, 1-8

Robustification of confidence intervals for the maximum point of a quadratic regression against autocorrelation

Dorota Kozioł[1], **Wojciech Zieliński**[2]

[1]Warsaw University of Life Sciences, Department of Agricultural Economics and IER, Nowoursynowska 166, 02-787 Warsaw, Poland, e-mail: dorota.koziol@ekonometria.info
[2]Warsaw University of Life Sciences, Department of Econometrics and Statistics, Nowoursynowska 166, 02-787 Warsaw, Poland, e-mail: wojtek.zielinski@statystyka.info

Summary

In this paper the problem of interval estimation of a maximum point of a quadratic regression function in the case of correlated random errors is considered. An approximate Student confidence interval for the maximum point is highly unrobust against autocorrelation (Kozioł and Zieliński, 2003b). In the paper a method for robustification of this interval is presented. This solution relies on modification of the data after which the minimum confidence level is close to the nominal one.

Key words: quadratic regression, autocorrelaction, maximum of quadratic regression, confidence interval, robustness

1. Introduction

Consider a quadratic regression model with repeated measurements, i.e. the model

$$Y_{ij} = \beta_0 + \beta_1 x_j + \beta_2 x_j^2 + \varepsilon_{ij}, \quad i = 1, \ldots, k, \quad j = 1, \ldots, m,$$

where i is a number of observations in the j's measurement point. Assume that ε_{ij} are normally distributed random errors with $E(\varepsilon_{ij}) = 0$, $D^2(\varepsilon_{ij}) = \sigma^2$ and

$$E(\varepsilon_{i_1 j_1}\varepsilon_{i_2 j_2}) = \begin{cases} \sigma^2 \varrho^{|j_1 - j_2|} & \text{for} \quad i_1 = i_2 \\ 0 & \text{for} \quad i_1 \neq i_2 \end{cases}.$$

The problem concerns interval estimation of $\varphi = -\beta_1/2\beta_2$, with $\beta_2 < 0$, i.e. the point at which the regression function attains its maximum.

Under the assumption of independence, ($\varrho = 0$) of random errors at least two confidence intervals for φ are known: the exact confidence interval and the approximate Student confidence interval (Kozioł and Zieliński, 2003a). The properties of both of these confidence intervals are similar in the basic model. They have been widely studied by many authors (Buonaccorsi, 1985; Buonaccorsi and Iyer, 1984; Kozioł and Zieliński, 2003a). For our investigations we choose the approximate Student confidence interval because this one always exists.

In many practical applications it appears that the ε's are not independent, for example in models of growth. In Kozioł and Zieliński (2003b) it was shown that the confidence level of a Student confidence interval is very sensitive to the autocorrelation of random errors, while its length is quite stable. The problem is what to do to make the confidence level more stable.

2. Approximate Student Confidence Interval

In matrix notation the considered model is of the form

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}, \tag{1}$$

where $\mathbf{Y} = (Y_{11}, \ldots, Y_{1m}, \ldots, Y_{k1}, \ldots, Y_{km})'$ is the vector of observations, $\mathbf{X} = \mathbf{1}_k \otimes \mathbf{U}$ with $\mathbf{U} = [1 \quad x_j \quad x_j^2]_{j=1,\ldots,m}$ is the design matrix, $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)'$ is the vector of regression coefficients and $\boldsymbol{\varepsilon} = (\varepsilon_{11}, \ldots, \varepsilon_{1m}, \ldots, \varepsilon_{k1}, \ldots, \varepsilon_{km})'$ is the vector of i.i.d random errors. Here $\mathbf{1}_k$ denotes a k-vector of ones. Assume that matrix $\mathbf{U}$ is of full rank. If so, there exists the matrix $(\mathbf{X}'\mathbf{X})^{-1}$. Let us denote the elements of $(\mathbf{X}'\mathbf{X})^{-1}$ by ν^{ij}, i.e. $(\mathbf{X}'\mathbf{X})^{-1} = [\nu^{ij}]_{i,j=0,1,2}$. Let

$$\hat{\boldsymbol{\beta}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{Y} = \frac{1}{k}(\mathbf{U}'\mathbf{U})^{-1}(\mathbf{1}_k' \otimes \mathbf{U}')\mathbf{Y},$$

and

$$\begin{aligned} S^2 &= \mathbf{Y}'(\mathbf{I} - \mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}')\mathbf{Y}/(n-3) = \\ &= \mathbf{Y}'(\mathbf{I} - \tfrac{1}{k}\mathbf{1}_k\mathbf{1}_k' \otimes \mathbf{U}(\mathbf{U}'\mathbf{U})^{-1}\mathbf{U}')\mathbf{Y}/(n-3) \end{aligned}$$

be LSE estimators of $\boldsymbol{\beta}$ and σ^2, respectively ($n = km$). Assuming that $\boldsymbol{\varepsilon} \sim N_n(\mathbf{0}, \sigma^2\mathbf{I}_n)$ we have

$$\hat{\boldsymbol{\beta}} \sim N_3\left(\boldsymbol{\beta}, \sigma^2(\mathbf{X}'\mathbf{X})^{-1}\right), \qquad (n-3)S^2 \sim \sigma^2\chi^2(n-3),$$

and $\hat{\boldsymbol{\beta}}$ and S^2 are stochastically independent. Let $\hat{\varphi} = -\hat{\beta}_1/2\hat{\beta}_2$ be the point estimator of the point φ of the regression function.

The approximate Student confidence interval is based on the fact (Serfling, 1980) that $\hat{\varphi}$ is asymptotically normal:

$$\hat{\varphi} = -\hat{\beta}_1/2\hat{\beta}_2 \sim AN(\varphi; \sigma^2\omega^2),$$

where

$$\omega^2 = \frac{4\varphi^2\nu^{22} + 4\varphi\nu^{12} + \nu^{11}}{(2\beta_2)^2}.$$

Application of the classical Student technique gives the following approximate confidence interval for φ:

$$(\hat{\varphi} \pm t(\alpha, n-3)S\hat{\omega}), \tag{2}$$

where $\hat{\omega}^2 = (4\hat{\varphi}^2\nu^{22} + 4\hat{\varphi}\nu^{12} + \nu^{11})/(2\hat{\beta}_2)^2$ and $t(\alpha, n-3)$ is the critical value of t distribution.

The above confidence interval is constructed under the assumption $\varrho = 0$. Now assume that the correlation matrix of random errors $\boldsymbol{\varepsilon}$ is of the form $\sigma^2(\mathbf{I}_k \otimes \boldsymbol{\Sigma})$ with $\boldsymbol{\Sigma} = [\varrho^{|i-j|}]_{i,j=1,\ldots,m}$. Such a correlation structure is typical for $AR(1)$ processes. In Kozioł and Zieliński (2003b) simulation studies showed that if $\varrho > 0$, then the confidence level decreases by up to 70% of the nominal level, i.e. the assumed level in the case $\varrho = 0$. Hence the confidence interval (2) may be considered as unrobust against correlation.

In what follows a method of robustification of this confidence interval is presented. Because $\boldsymbol{\Sigma}$ is p.d. there exists an upper triangle matrix $\mathbf{W}$ such that $\boldsymbol{\Sigma} = \mathbf{W}'\mathbf{W}$. Let $\mathbf{V} = \mathbf{W}^{-1}$. The matrix $\mathbf{V}$ is of the form

$$\mathbf{V} = \frac{1}{\sqrt{1-\varrho^2}} \begin{bmatrix} \sqrt{1-\varrho^2} & -\varrho & 0 & 0 & \cdots & 0 \\ 0 & 1 & -\varrho & 0 & \cdots & 0 \\ 0 & 0 & 1 & -\varrho & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \end{bmatrix}.$$

Let $\mathbf{L} = \mathbf{I}_k \otimes \mathbf{V}'$. Our proposition is to use $\mathbf{Z} = \mathbf{LY}$ instead of $\mathbf{Y}$. Then the considered model $\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$ becomes

$$\mathbf{Z} = \mathbf{LX}\boldsymbol{\beta} + \mathbf{L}\boldsymbol{\varepsilon}, \tag{3}$$

where $\mathbf{Z} = (Z_{11}, \ldots, Z_{1m}, \ldots, Z_{k1}, \ldots, Z_{km})'$ is the vector of modified observations. Matrix $\mathbf{U}$ is of full rank, so there exists the matrix $((\mathbf{LX})'(\mathbf{LX}))^{-1} =$

$(\mathbf{X}'(\mathbf{I}_k \otimes \mathbf{\Sigma}^{-1})\mathbf{X})^{-1} = \frac{1}{k}(\mathbf{U}'\mathbf{V}\mathbf{V}'\mathbf{U})^{-1}$. Denote an element of this matrix by ν^{st}, i.e. $(\mathbf{X}'(\mathbf{I}_k \otimes \mathbf{\Sigma}^{-1})\mathbf{X})^{-1} = [\nu^{st}]_{s,t=0,1,2}$. Let

$$\begin{aligned}\tilde{\boldsymbol{\beta}} &= [(\mathbf{1}_k \otimes \mathbf{V}'\mathbf{U})'(\mathbf{1}_k \otimes \mathbf{V}'\mathbf{U})]^{-1}(\mathbf{1}_k \otimes \mathbf{V}'\mathbf{U})'\mathbf{Z} \\ &= \tfrac{1}{k}[\mathbf{U}'\mathbf{V}\mathbf{V}'\mathbf{U}]^{-1}(\mathbf{1}_k' \otimes \mathbf{U}'\mathbf{V}\mathbf{V}')\mathbf{Y},\end{aligned}$$

and

$$S^2 = \mathbf{Z}'(\mathbf{I}_{km} - \frac{1}{k}\mathbf{1}_k\mathbf{1}_k' \otimes \mathbf{V}'\mathbf{U}(\mathbf{U}'\mathbf{V}\mathbf{V}'\mathbf{U})^{-1}\mathbf{U}'\mathbf{V})\mathbf{Z}/(n-3)$$

be LSE estimators of $\boldsymbol{\beta}$ and σ^2 respectively ($n = km$) in model (3). Note that $\mathbf{L}\boldsymbol{\varepsilon} \sim N_{km}(0, \sigma^2\mathbf{I}_{km})$. Hence

$$\tilde{\boldsymbol{\beta}} \sim N_3\left(\boldsymbol{\beta}, \frac{\sigma^2}{k}[\mathbf{U}'\mathbf{V}\mathbf{V}'\mathbf{U}]^{-1}\right), \qquad (n-3)S^2 \sim \sigma^2\chi^2(n-3),$$

and $\tilde{\boldsymbol{\beta}}$ and S^2 are stochastically independent. It is easy to check that, although $\tilde{\boldsymbol{\beta}} \neq \hat{\boldsymbol{\beta}}$, $\tilde{\varphi} = -\tilde{\beta}_1/2\tilde{\beta}_2 = \hat{\varphi}$.

We construct the Student confidence interval (2) in the modified model. Because the value of autocorrelation ϱ is unknown, it is estimated by the popular estimator

$$r = \frac{\frac{1}{m-1}\sum_{j=1}^{m-1}\sum_{i=1}^{k}(Y_{ij} - \bar{Y}_{\bullet j})(Y_{i,j+1} - \bar{Y}_{\bullet,j+1})}{\frac{1}{m-1}\sum_{j=1}^{m-1}\sqrt{\sum_{i=1}^{k}(Y_{ij} - \bar{Y}_{i\bullet})^2\sum_{i=1}^{k-1}(Y_{i+1,j} - \bar{Y}_{i+1,\bullet})^2}},$$

where $\bar{Y}_{\bullet j} = \frac{1}{k}\sum_{i=1}^{k} Y_{ij}$, $(j = 1, \ldots, m)$ and $\bar{Y}_{i\bullet} = \frac{1}{m}\sum_{j=1}^{m} Y_{ij}$, $(i = 1, \ldots, k)$. The above estimator is a modification of the estimators of autocorrelation coefficient proposed by von Neumann and Durbin-Watson (see: Chow 1983; Greene 2000).

We investigated the properties of the confidence interval (2). To estimate the confidence level as well as the length the Monte Carlo method was applied.

3. Simulation studies

In simulation studies we confine ourselves to the interval $x \in [-1;1]$. Note that every finite interval for x may be reduced to $[-1;1]$. We chose

$$x_i = -1 + \frac{2}{9}i, \qquad i = 0, 1, \dots, 9,$$

i.e. ten equally distributed points over the considered interval ($m = 10$). Such a choice should model time points which are equidistant (for example ten weeks). We observe $k = 5$ courses of the regression function. Hence we have 5×10 observations. Also we took $\sigma^2 = 0.1$. On such observations we build a confidence interval for the maximum and note its length and whether it hits a true maximum point. This procedure was repeated 1000 times and as a result we obtained the mean length as well as the empirical confidence level.

In our simulations we consider the 100 quadratic regression functions given below:

β_2	β_1									
−0.1	0.00	0.02	0.04	0.06	0.08	0.10	0.12	0.14	0.16	0.18
−0.5	0.00	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90
−1.0	0.00	0.20	0.40	0.60	0.80	1.00	1.20	1.40	1.60	1.80
−1.5	0.00	0.30	0.60	0.90	1.20	1.50	1.80	2.10	2.40	2.70
−2.0	0.00	0.40	0.80	1.20	1.60	2.00	2.40	2.80	3.20	3.60
−2.5	0.00	0.50	1.00	1.50	2.00	2.50	3.00	3.50	4.00	4.50
−3.0	0.00	0.60	1.20	1.80	2.40	3.00	3.60	4.20	4.80	5.40
−3.5	0.00	0.70	1.40	2.10	2.80	3.50	4.20	4.90	5.60	6.30
−4.0	0.00	0.80	1.60	2.40	3.20	4.00	4.80	5.60	6.40	7.20
−4.5	0.00	0.90	1.80	2.70	3.60	4.50	5.40	6.30	7.20	8.10
−5.0	0.00	1.00	2.00	3.00	4.00	5.00	6.00	7.00	8.00	9.00
$x_{\max}$	0	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9

The columns contain functions with the same maximum point and of different flatness. The rows contain functions of the same flatness and with different maximum points. By symmetry we consider only positive maximum points. The constant β_0 is not important and we put it equal to 10.

Also we have to adopt some values for autocorrelation. To enhance our simulations we consider autocorrelations ranging from −0.9 to 0.9 in steps of 0.1, i.e. we consider 19 autocorrelation values.

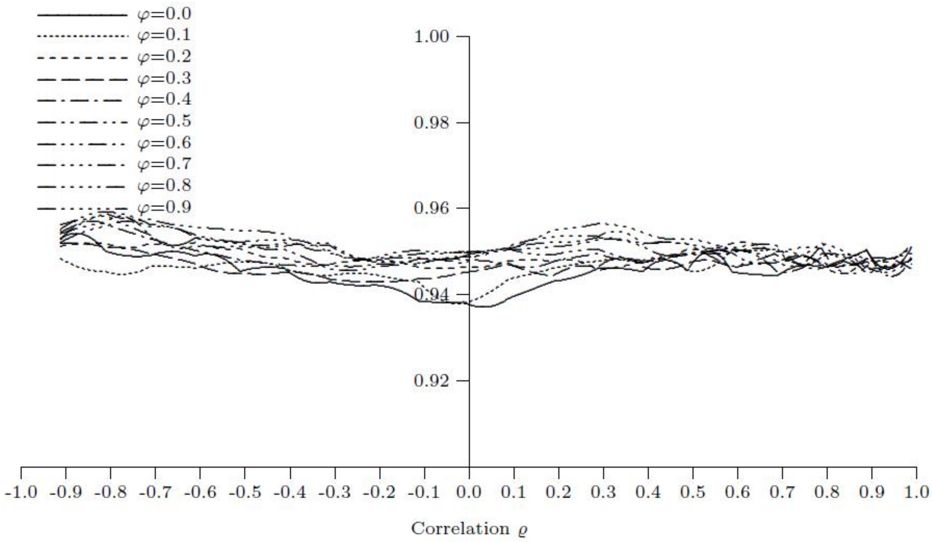

Figure 1. Empirical confidence level of Student confidence interval.

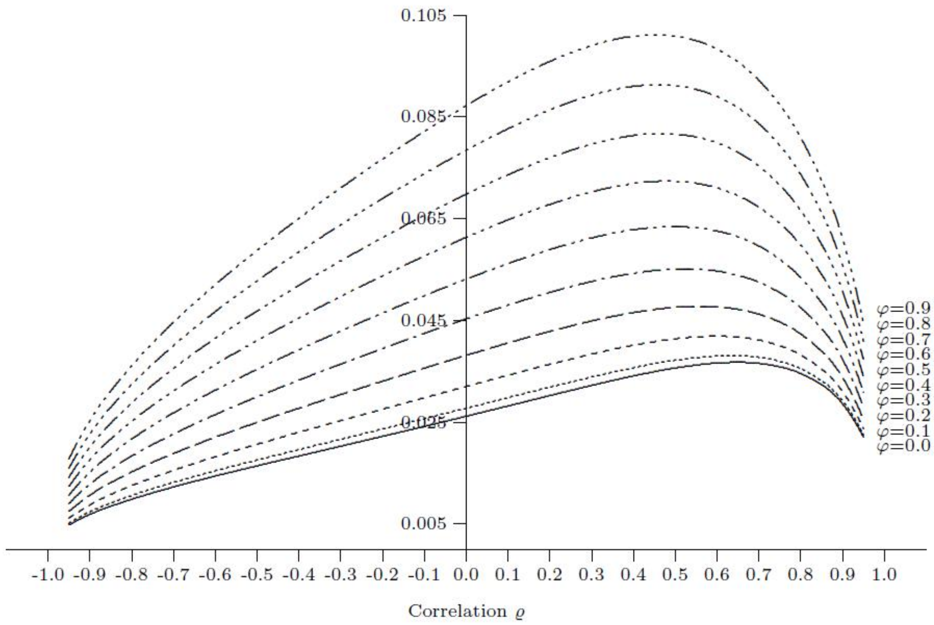

Figure 2. Empirical length of Student confidence interval.

4. Results

The results of the simulations are shown in the figures. For presentation we chose only functions with different maximum points and the same flatness ($\beta_2 = -5$). For other functions the results are similar.

The confidence level of the confidence interval was highly unrobust against autocorrelation (Kozioł and Zieliński, 2003b). If $\varrho > 0$, then the confidence level decreases by up to 70% of the nominal level, i.e. the assumed level in the case $\varrho = 0$. The comparison of confidence level is shown in Figure 1. Values of autocorrelation are on the X axis, the confidence level is on the Y axis. The comparison of confidence intervals for different autocorrelations shows that data modification has affected the confidence level significantly: the confidence level is close to the nominal one ($1 - \alpha = 0.95$) for all values of autocorrelation.

The lengths of confidence interval were robust against autocorrelation (Kozioł and Zieliński, 2003b). The comparison of length of confidence interval is shown in Figure 2. Values of autocorrelation are on the X axis, and the length of confidence interval is on the Y axis. The comparison of length of confidence interval for different autocorrelations shows that the length of the interval is affected by the data modification. The interval lengths increase. This is a cost of robustification of the confidence level.

The data modification using Cholesky decomposition is a good method of robustification of approximate Student confidence intervals against autocorrelation.

The results presented were obtained for $AR(1)$ processes. It may be expected that results for any matrix $\mathbf{\Sigma}$ with known structure should be similar.

References

Buonaccorsi J.P., (1985): Ratios of linear combinations in the general linear model. Comm. in Statist.14: 635–650.

Buonaccorsi J.P., Iyer H.K. (1984): A comparison of confidence regions and designs in estimation of a ratio. Comm. in Statist.13: 723–741.

Chow G.C. (1983): Econometrics. Mc-Graw-Hill Book Company.

Greene W.H. (2000): Econometric Analysis. Prentice Hall Inc..

Kozioł D., Zieliński W. (2003a): Comparison of confidence intervals for maximum of a quadratic regression function. Biometrical Letters 40: 57–64.

Kozioł D., Zieliński W. (2003b): Robustness of confidence intervals for the maximum point of a quadratic regression against autocorrelation. Biometrical Letters 40: 65–72.

Serfling R.J. (1980): Approximation Theorems of Mathematical Statistics. John Wiley & Sons.