Report: AE13/76

Instituut voor Actuariaat & Econometrie

Bounds for the bias of the LS estimator of σ^2 in the case of a first-order (positive) autoregressive process when the regression contains a constant term

by H. Neudecker

Classification: AMS (MOS) subject classification scheme (1970):
15A18, 26A60, 62M10, 62P20.

Keywords: Courant-Fischer min-max theorem, multiple regression.

Abstract: The LS-estimate $\frac{e'e}{n-k}$ of σ^2 and its bias are considered, in the case of regression with a constant term.
Following the procedure described in an earlier paper (AE2/76) much more satisfactory bounds are established for $\frac{1}{\sigma^2} E \frac{e'e}{n-k}$.
The computations show that the LS-estimator is biased toward zero for an impressive number of values of n, k and ρ.

Free copies are obtainable from
Interfakulteitsbureau Aktuariaat en Ekonometrie
Jodenbreestraat 23 - Amsterdam, 'phone: 020-5254212

Universiteit van Amsterdam

Bounds for the bias of the LS estimator of σ^2 in the case of a first-order (positive) autoregressive process when the regression contains a constant term.

H. Neudecker*

Introduction

In a recent paper [2] we considered the model

$$y = X\beta + \varepsilon,$$

where ε is the disturbance vector, X is of order (n,k) and rank k, the disturbance elements ε_i of ε are generated by the process

$$\varepsilon_i = \rho\varepsilon_{i-1} + \xi_i \qquad 0 < \rho < 1$$

where the ξ_i are uncorrelated random variables with zero mean and variance σ_0^2. We established bounds for $\frac{1}{\sigma^2} E \frac{e'e}{n-k}$, where e is the LS residual vector, $\sigma^2 = \frac{\sigma_0^2}{1-\rho^2}$.

This led to the interval

$$\frac{\sum_{i=1}^{n-k} \lambda_{i+k}}{n-k} \leqq \frac{1}{\sigma^2} E \frac{e'e}{n-k} \leqq \frac{\sum_{i=1}^{n-k} \lambda_i}{n-k} \qquad \text{(I)}$$

where the λ_i are the eigenvalues of the covariance matrix V of ε and further $\lambda_1 \geqq \lambda_2 \geqq \dots \geqq \lambda_n$.

Intervals were then computed for various values of n,k and ρ. They appeared to be asymmetrical around 1, and suggested a negative bias.

No restriction was put on X.

In this paper we want to incorporate a constant term in the regression, which implies that X should contain a column of ones.

In an appendix we discuss the incorporation of additional columns in X.

---

* Instituut voor Actuariaat en Econometrie, Universiteit van Amsterdam.

Bounds for tr MV

We shall apply the same procedure as before. It will, however, be a little more streamlined.

It is well-known that $\frac{1}{\sigma^2} E \frac{e'e}{n-k} = \frac{1}{n-k}$ tr MV, where $M = I-X(X'X)^{-1}X'$. X is an (n,k) matrix, hence M is idempotent of rank n-k.

Further Ms = 0, if the regression contains a constant term, where s' = (1....1). M can be expressed as BB', where $B'B = I_{n-k}$, B is of order (n,n-k). Further BB's = 0 or, equivalently, B's = 0.

We shall interpret B as a variable matrix, subject to the constraints

$$B'B = I_{n-k} \tag{1}$$

$$B's = 0 \; . \tag{2}$$

By maximizing and minimizing $\frac{1}{n-k}$ tr BB'V subject to (1) and (2) we shall then be able to establish an interval for $\frac{1}{\sigma^2} E \frac{e'e}{n-k} = \frac{1}{n-k}$ tr MV.

We form the Lagrangian function

$$\phi = \tfrac{1}{2} \operatorname{tr} BB'V - \operatorname{tr} L(B'B - I) - m'B's$$

where L and m' are Lagrange (matrix and vector) multipliers.

Necessary for an extremum is

$$0 = d\phi = \operatorname{tr} B'VdB - \operatorname{tr} \tilde{L}B'dB - \operatorname{tr} ms'dB$$

$$- \operatorname{tr}(B'B - I)\, dL - (dm)'B's,$$

where $\tilde{L} = L + L'$.

This leads to the conditions:

$$B'V = \tilde{L}B' + ms' \tag{3}$$

$$B'B = I \tag{4}$$

$$B's = 0 \; . \tag{5}$$

We postmultiply (3) by s, use (5) and get

$$m = \frac{1}{n} B'Vs \tag{6}$$

($Vs \neq 0$ as V is regular).

This leads to

$$B'VA = \tilde{L}B' \qquad (3^*)$$

where $A = I - \frac{1}{n} ss'$.

We shall rewrite (3^*) as

$$B'AVA = \tilde{L}B' , \qquad (3^{**})$$

by virtue of (5).

$\tilde{L}$ is symmetric. It can therefore be diagonalized:

$Q'\tilde{L}Q = \Lambda$, say. Q is orthogonal.

We rewrite (3^{**}), (4) and (5) as follows:

$$\tilde{B}'AVA = \Lambda\tilde{B}' \qquad (3^o)$$

$$\tilde{B}'\tilde{B} = I \qquad (4^o)$$

$$\tilde{B}'s = 0 \qquad (5^o)$$

where $\tilde{B} = BQ$.

(3^o) states that the diagonal elements of Λ are n-k eigenvalues of AVA, or equivalently VA. $\tilde{B}'$ is a matrix whose n-k rows are the corresponding eigenvectors.
As s' is an eigenvector of AVA (corresponding to zero eigenvalue), requirement (5^o) can easily be met. $\tilde{B}'$ should not contain this particular eigenvector.

Let us denote the eigenvalues of AVA by μ_i and let $\mu_1 \geqq \mu_2 \geqq \dots \geqq \mu_{n-1} > 0$.
(AVA clearly has only one zero eigenvalue).
As $\text{tr}\, BB'V = \text{tr}\, \tilde{B}\tilde{B}'V = \text{tr}\, \tilde{B}\tilde{B}'(A+ \frac{1}{n}ss')V(A+ \frac{1}{n}ss')$

$$= \text{tr}\, \tilde{B}\tilde{B}'AVA = \text{tr}\, \tilde{B}\Lambda\tilde{B}' = \text{tr}\, \Lambda,$$

tr BB'V will have as its maximum value $\sum_{i=1}^{n-k} \mu_i$, and as its minimum value $\sum_{i=1}^{n-k} \mu_{i+k-1}$.

We can thus establish the inequality

$$\frac{\sum_{i=1}^{n-k} \mu_{i+k-1}}{n-k} \leqq \frac{1}{\sigma^2} E \frac{e'e}{n-k} \leqq \frac{\sum_{i=1}^{n-k} \mu_i}{n-k} \qquad \text{(II)}$$

It is obvious that this interval is tighter than the previously established one (I), because there are more constraints.

We have computed bounds for $\frac{1}{\sigma^2} E \frac{e'e}{n-k}$ according to formula (II) for various values of n,k and ρ[1].
The results are shown in Table 1.

The results are very satisfactory. It appears that for ρ= 0.8 $\frac{e'e}{n-k}$ is biased toward zero for all values of k considered by us. For ρ = 0.5 there is a negative bias in case k=2 or 3, whereas for ρ = 0.3 a negative bias is established for k=2 only. These results hold for all values of n considered by us.
The computations show that Theil's (3, p.257) conclusion that $\frac{e'e}{n-2}$ is biased toward zero for positive ρ can to some extent be generalized for other values of k.
Increasing the number of parameters k clearly tends to undermine the conclusion about the sign of the bias for low or intermediate values of ρ. An obvious remedy for this is increasing the number of observations n.
Some of the results (for k=2 and k=4) are represented in the form of diagrams.

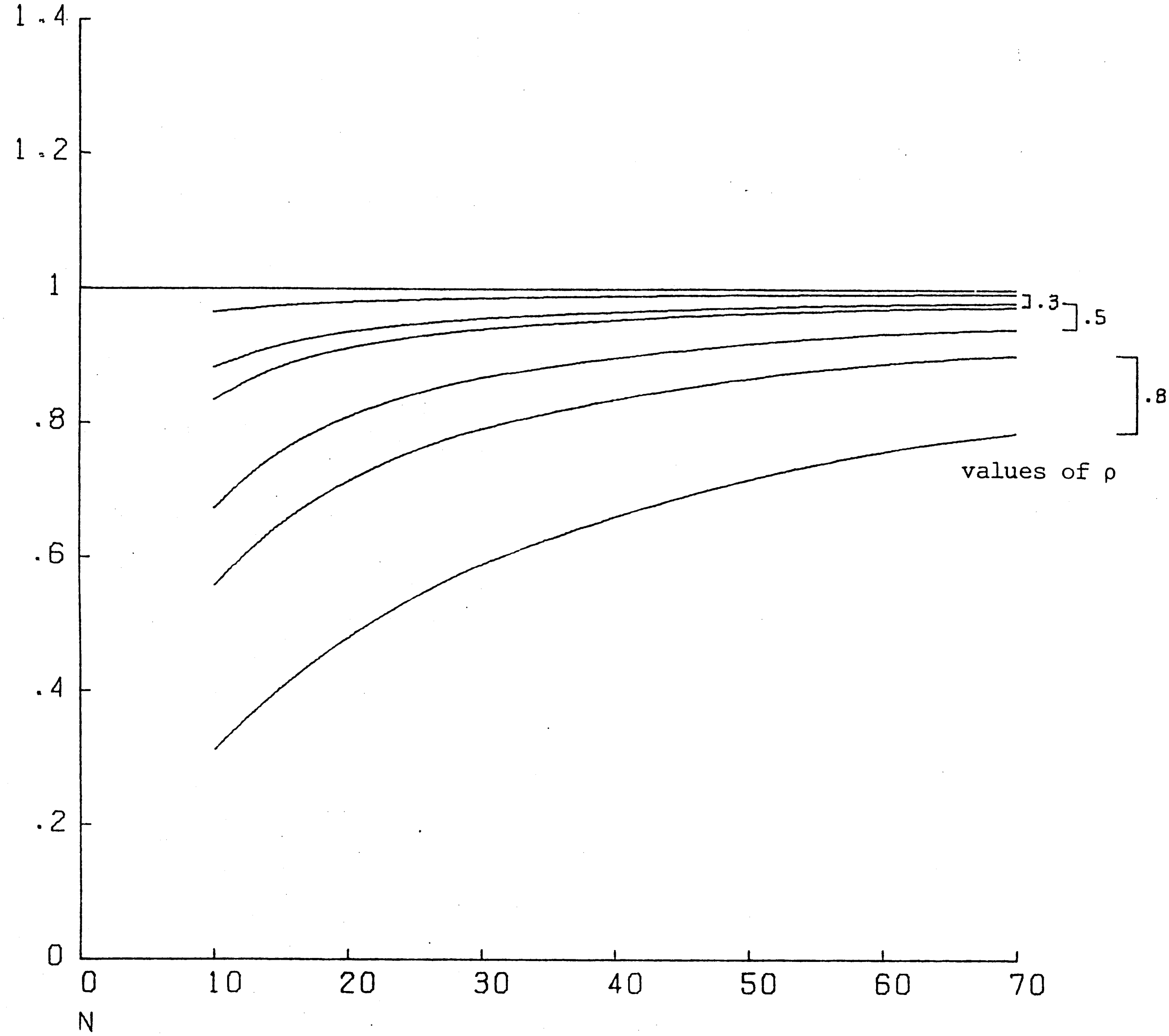
---

1) Thanks are due to Mrs Els de Bakker for performing the computations.

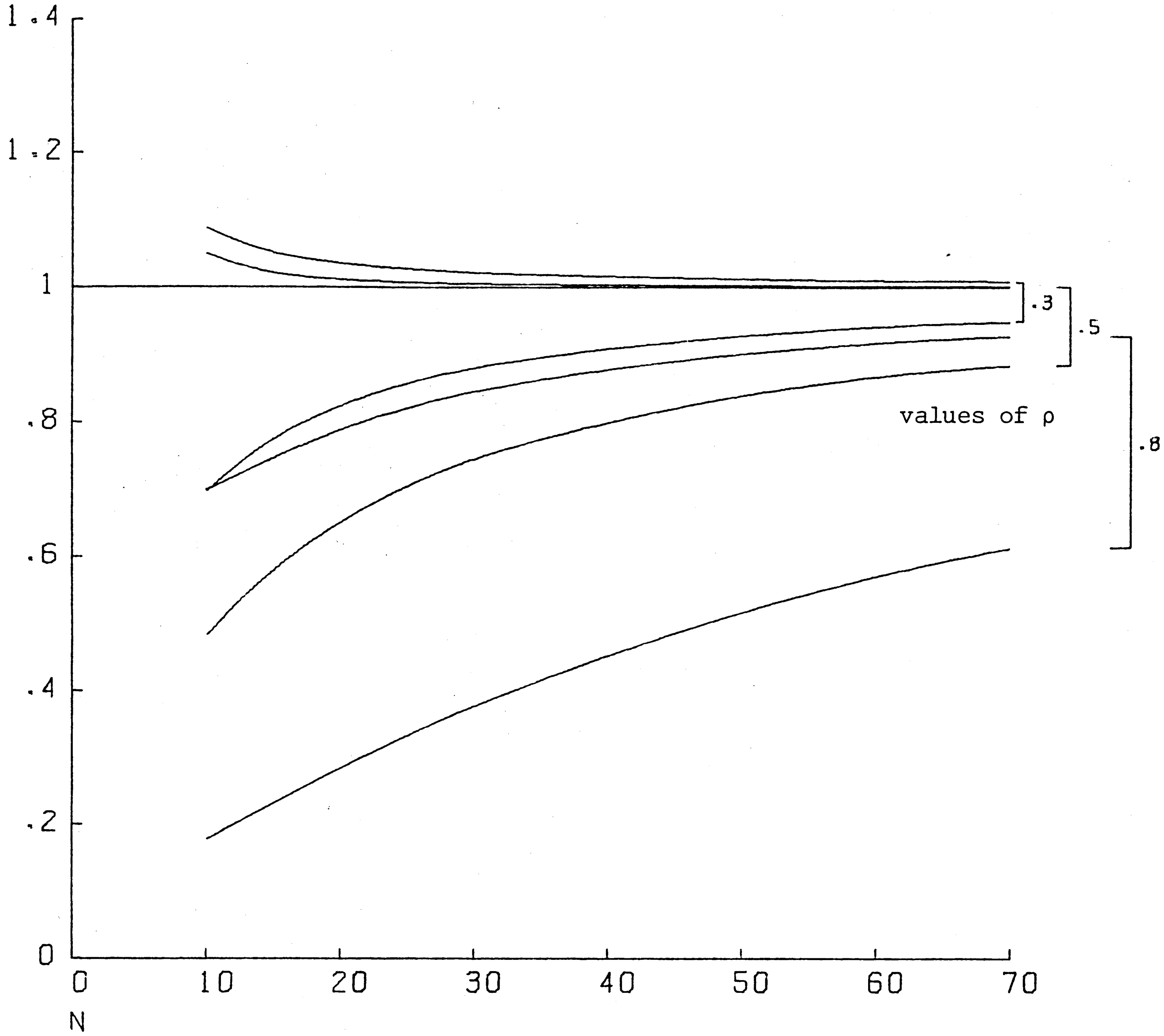
Table 1

Bounds of $\frac{1}{\sigma^2} E \frac{e'e}{n-k}$ for various n, k and ρ

		ρ = 0.3		ρ = 0.5		ρ = 0.8	
	k	lower bound	upper bound	lower bound	upper bound	lower bound	upper bound
	2	0.83486	0.96478	0.67256	0.88242	0.31244	0.55709
n=10	3	0.75721	1.02059	0.55693	0.95668	0.22049	0.61917
	4	0.69650	1.08728	0.48309	1.04905	0.17759	0.69924
	5	0.64638	1.16567	0.42971	1.16465	0.15154	0.80555
	2	0.88555	0.97554	0.75854	0.91770	0.40571	0.65081
n=15	3	0.82734	1.01065	0.65645	0.96534	0.29148	0.69538
	4	0.77573	1.05024	0.58008	1.02010	0.23307	0.74745
	5	0.72898	1.09479	0.51916	1.08331	0.19567	0.80895
	2	0.91311	0.98130	0.81064	0.93694	0.48068	0.71473
n=20	3	0.86776	1.00681	0.72353	0.97202	0.35496	0.75008
	4	0.82518	1.03480	0.65196	1.01091	0.28521	0.78962
	5	0.78481	1.06541	0.59112	1.05489	0.23869	0.83409
	2	0.93020	0.98487	0.84500	0.94893	0.54122	0.76010
n=25	3	0.89346	1.00489	0.77046	0.97671	0.41107	0.78952
	4	0.85791	1.02647	0.70551	1.00685	0.33342	0.82164
	5	0.82335	1.04969	0.64785	1.03958	0.27970	0.85681
	2	0.94177	0.98729	0.86915	0.95711	0.59061	0.79358
n=30	3	0.91105	1.00376	0.80463	0.99010	0.46040	0.81881
	4	0.88082	1.02131	0.74619	1.00469	0.37762	0.84593
	5	0.85101	1.03996	0.69271	1.03103	0.31839	0.87512
	2	0.96514	0.99226	0.92004	0.97388	0.71856	0.86852
n=50	3	0.94687	1.00188	0.87926	0.98749	0.60468	0.88462
	4	0.92848	1.01189	0.83974	1.00165	0.51822	0.90142
	5	0.90996	1.02227	0.80141	1.01640	0.44935	0.91894
	2	0.97517	0.99444	0.94266	0.98123	0.78808	0.90383
n=70	3	0.96227	1.00123	0.91327	0.99089	0.69352	0.91565
	4	0.94923	1.00822	0.88419	1.00083	0.61436	0.92784
	5	0.93605	1.01540	0.85543	1.01107	0.54671	0.94039



Bounds for $E \frac{e'e}{n-k} / \sigma^2$ for k=2 and ρ=.3,.5,.8.



Bounds for $E\,\frac{e'e}{n-k}/\sigma^2$ for k=4 and ρ=.3,.5,.8.

Appendix The incorporation of additional columns in X

It is easy to see that the incorporation of columns orthogonal to s in X will further tighten the interval.

Let us incorporate a column t orthogonal to s.[2)]

We then have to add an expression $\ell'B't$ say to the Lagrangian function ϕ.

This leads to $\psi = \phi - \ell'B't$, where ℓ' is a Lagrange vector multiplier.

We now get the following necessary conditions:

$$B'V = \tilde{L}B' + ms' + \ell t' \qquad (7)$$

$$B'B = I \qquad (8)$$

$$B's = 0 \qquad (9)$$

$$B't = 0 \;. \qquad (10)$$

Further $s't = 0$. (11)

We postmultiply (7) by s and t, using (9), (10) and (11), and find:

$$m = \frac{1}{n} B'Vs \qquad (12)$$

$$\ell = \frac{1}{t't} B'Vt \;. \qquad (13)$$

(7) will now become:

$$B'VC = \tilde{L}B' \qquad (7^*)$$

where $C = I - \frac{1}{n} ss' - \frac{1}{t't} tt'$.

We shall rewrite (7^*) as

$$B'CVC = \tilde{L}B' \;, \qquad (7^{**})$$

by virtue of (9) and (10).

We finally arrive at

$$\tilde{B}'CVC = \Lambda\tilde{B}' \qquad (7^o)$$

$$\tilde{B}'\tilde{B} = I \qquad (8^o)$$

$$\tilde{B}'s = 0 \qquad (9^o)$$

$$\tilde{B}'t = 0 \qquad (10^o)$$

where $\tilde{B} = BQ$.

Further $s't = 0$ (11)

---

2) An obvious choice for t is a vector representing a linear trend e.g. $t' = \left[-\tfrac{1}{2}(n-1) \;\ldots\; 0 \;\ldots\; \tfrac{1}{2}(n-1)\right]$ for odd n.

Both s' and t' are eigenvectors of CVC (corresponding to the two zero eigenvalues). $\tilde{B}'$ should therefore not contain these two particular eigenvectors.

Let the n-2 non-zero eigenvalues of CVC be denoted by ν_i where $\nu_1 \geqq \nu_2 \geqq \cdots \geqq \nu_{n-2}$, then tr BB'V will have as its maximum value $\sum_{i=1}^{n-k} \nu_i$ and as its minimum value $\sum_{i=1}^{n-k} \nu_{i+k-2}$.

We have now found the interval

$$\frac{\sum_{i=1}^{n-k} \nu_{i+k-1}}{n-k} \leqq \frac{1}{\sigma^2} E \frac{e'e}{n-k} \leqq \frac{\sum_{i=1}^{n-k} \nu_i}{n-k} \qquad \text{(III)}$$

This interval is again tighter than (II).

We can apply this theorem because $C = I - \frac{1}{n} ss' - \frac{1}{t't} tt'$

$= I - Z(Z'Z)^{-1}Z$, where $Z = (s \vdots t)$.

The theorem implies:

$$\lambda_{i+2} \leqq \nu_i \leqq \lambda_i \quad , \qquad i=1 \ldots n-2$$

hence

$$\sum_{i=1}^{n-k} \nu_i \leqq \sum_{i=1}^{n-k} \lambda_i$$

$$\sum_{i=1}^{n-k} \nu_{i+k-2} \geqq \sum_{i=1}^{n-k} \lambda_{i+k}$$

References

[1] Anderson, T.W.: The Statistical Analysis of time series.
New York; Wiley, 1971

[2] Neudecker, H.: Bounds for the bias of the LS estimator of σ^2 in the case of a first-order autoregressive process (positive autocorrelation). Report AE2/76, Instituut voor Aktuariaat en Ekonometrie, 1976

[3] Theil, H.: Principles of Econometrics.
New York; Wiley, 1971