Journal of Statistical Research *ISSN 0256 - 422 X*
2007, Vol. 41, No. 1, pp. 1–15
Bangladesh

ESTIMATION OF AUTOREGRESSIVE COEFFICIENT IN AN ARMA(1, 1) MODEL WITH VAGUE INFORMATION ON THE MA COMPONENT[1]

M. Ould Haye

School of Mathematics and Statistics
1125 Colonel By, Carleton University, Canada K1S 5B6
Email: ouldhaye@math.carleton.ca

SUMMARY

In this paper we investigate the asymptotic properties of various estimators of autocorrelation parameter of an ARMA(1,1) model when uncertain non–sample prior information on the moving average component is available. In particular we study the preliminary test and the shrinkage estimators of the autocorrelation parameter and we compare their efficiency with respect to the maximum likelihood estimator designated as the unrestricted estimator. It is shown that near the prior information on MA–parameter, both preliminary test and shrinkage estimators are superior to the MLE while they lose their superiority as the MA–parameter moves away from the prior information although preliminary test estimator gains its efficiency to some extent but the shrinkage estimator attains its lower bound of its efficiency.

Keywords and phrases: Time series, ARMA processes, sub–hypotheses, uncertain non–sample prior information, restricted, preliminary test and shrinkage estimators, asymptotic relative efficiency

1 Introduction

Classical estimators of unknown parameters are based exclusively on the sample data. The notion of non–sample information has been introduced to improve the quality of the estimators. We expect that the inclusion of additional information would lead to a better estimator. Since the seminal work of Bancroft (1944) on preliminary test estimators, many papers in the area of the so–called improved estimation have been published. Stein (1956, 1981) developed the shrinkage estimator for multivariate normal population and proved that it performs better than the usual maximum likelihood estimator in terms of the square error loss function. Saleh (2006) explored these two classes of improved estimators in a variety of contexts in his recent book. Many other researchers have been working in this area,

[1]This research was supported by NSERC, grant no 288188

notably Sclove et al. (1972), Judge and Bock (1978), and Efron and Morris (1972, 1973). This type of analysis is also useful to control overmodelling of specific models in question. More specific reasons and benefits of such analysis may be found in the recent book of Saleh (2006). To the best of our knowledge, these alternative estimators have not yet been applied to ARMA model parameter estimation.

We consider the ARMA(1,1) model given by

$$X_t - \rho X_{t-1} = \epsilon_t - \alpha \epsilon_{t-1}. \tag{1.1}$$

where ϵ_t are i.i.d. $\mathcal{N}(0, \sigma^2)$.

We are interested in the estimation of the autoregressive parameter ρ when it is suspected but one is not sure that $\alpha = \alpha_0$. Such a situation may arise when there is prior information that $\alpha = \alpha_0$ or is very close to α_0 and we want to reduce the number of parameters to be estimated in the model (1.1) and still improve on estimating ρ with better efficiency. More specifically we consider four estimators, namely 1) maximum likelihood estimator MLE, $\tilde{\rho}_n$ (the unrestricted estimator); 2) restricted estimator, $\hat{\rho}_n$ which generally performs better than the MLE when α is equal to α_0 (or very close to it), but if α is away from α_0 the restricted estimator may be considerably poorer than the MLE. (see (Saleh, 1992) in the over–modeling context); 3) the preliminary test estimator PTE, $\hat{\rho}_n^{PT}$. This estimator may be useful in case of uncertainty about the prior information $\alpha = \alpha_0$. We attempt to strike a compromise between $\tilde{\rho}_n$ and $\hat{\rho}_n$ via an indicator function depending on the size γ of the preliminary test on α and 4) the shrinkage estimator $\hat{\rho}_n^S$ (see saleh (2006)), which is basically a smoothed version of the PTE.

We obtain explicit forms of the asymptotic distributional bias (ADB) and the asymptotic distributional MSE (ADMSE) of each of these estimators. We give a detailed comparison study of the relative efficiency of these estimators.

Put

$$h(z) = 1 - \rho z \qquad \text{and} \qquad g(z) = 1 - \alpha z$$

and assume that $|\rho|$ and $|\alpha|$ are less than 1. We know (see (Dzhaparidze 1986)) that the spectral density of the process X_t is given by

$$f(\lambda) = \frac{\sigma^2}{2\pi} \left| \frac{g(z)}{h(z)} \right|^2 \qquad \text{with} \quad z = e^{i\lambda}$$

Its covariance function $r(k)$ satisfies the recurrence equation

$$r(k) - \rho r(k-1) = 0, \quad k = 2, 3, \ldots$$

Let $\theta = (\rho, \alpha, \sigma^2)'$ be the unknown parameter of the ARMA(1,1) model in (1.1). The maximum likelihood estimator MLE $(\tilde{\rho}_n, \tilde{\alpha}_n, \tilde{\sigma}_n^2)$ of θ satisfies

$$\int_{-\pi}^{\pi} I_n(\lambda)|h(z)|^2|g(z)|^{-4}\big[\cos\lambda - \tilde{\alpha}_n\big]d\lambda = 0 \tag{1.2}$$

$$\int_{-\pi}^{\pi} I_n(\lambda)|g(z)|^{-2}\big[\cos\lambda - \tilde{\rho}_n\big]d\lambda = 0 \tag{1.3}$$

$$\text{and} \quad \tilde{\sigma}_n^2 = \int_{-\pi}^{\pi} I_n(\lambda)|1-\tilde{\rho}_n|^2|1-\tilde{\alpha}_n|^2 d\lambda \tag{1.4}$$

where

$$I_n(\lambda) = \Big|\sum_{j=1}^{n} X_j e^{ij\lambda}\Big|^2$$

is the periodogram of $X_1, X_2, \ldots, X_n$.
The limit of the Fisher information matrix is given by

$$\Gamma_\theta = \begin{pmatrix} \frac{1}{1-\rho^2} & \frac{-1}{1-\rho\alpha} & 0 \\ \frac{-1}{1-\rho\alpha} & \frac{1}{1-\alpha^2} & 0 \\ 0 & 0 & \frac{1}{2\sigma^4} \end{pmatrix}$$

If we write $\theta = (\theta_1, \theta_2, \theta_3)$, each entry (j,k) of Γ_θ corresponds to

$$\frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{\partial}{\partial\theta_j}\ln f(\lambda)\frac{\partial}{\partial\theta_k}\ln f(\lambda)d\lambda$$

For example the value $-(1-\rho\alpha)^{-1}$ in Γ_θ is obtained as equal to

$$\frac{1}{4\pi}\int_{-\pi}^{\pi} \frac{\partial}{\partial\alpha}\ln f(\lambda)\frac{\partial}{\partial\rho}\ln f(\lambda)d\lambda,$$

using the variable change $z = e^{i\lambda}$, and taking $\gamma = \{z : |z| = 1\}$ and applying the residuals theorem.

The covariance matrix of $(\tilde{\rho}_n, \tilde{\alpha}_n, \tilde{\sigma}_n^2)$ is given by

$$\Sigma_n = \frac{1}{n(\rho-\alpha)^2}$$
$$\begin{pmatrix} (1-\rho^2)(1-\rho\alpha)^2 & (1-\rho^2)(1-\alpha^2)(1-\rho\alpha) & 0 \\ (1-\rho^2)(1-\alpha^2)(1-\rho\alpha) & (1-\alpha^2)(1-\rho\alpha)^2 & 0 \\ 0 & 0 & 2\sigma^4(\rho-\alpha)^2 \end{pmatrix}$$
$$+o(1/n).$$

Also we know (see (Dzhaparidze 1986)) that as $n \to \infty$,

$$\begin{pmatrix} \sqrt{n}(\tilde{\rho}_n - \rho) \\ \sqrt{n}(\tilde{\alpha}_n - \alpha) \end{pmatrix} \Longrightarrow \mathcal{N}(0, \Sigma) \tag{1.5}$$

where

$$\Sigma = \frac{1}{(\rho-\alpha)^2}\begin{pmatrix} (1-\rho^2)(1-\rho\alpha)^2 & (1-\rho^2)(1-\alpha^2)(1-\rho\alpha) \\ (1-\rho^2)(1-\alpha^2)(1-\rho\alpha) & (1-\alpha^2)(1-\rho\alpha)^2 \end{pmatrix}. \tag{1.6}$$

If we approximate the vector $\sqrt{n}(\tilde{\rho}_n-\rho, \tilde{\alpha}_n-\alpha)'$ by its asymptotic normal distribution with the previous matrix we get after some calculations

$$\mathbb{E}(\tilde{\rho}_n-\rho|\tilde{\alpha}_n-\alpha) = \frac{1-\rho^2}{1-\rho\alpha}(\tilde{\alpha}_n-\alpha) \tag{1.7}$$

and

$$\mathbb{E}\big[\mathrm{Var}(\tilde{\rho}_n|\tilde{\alpha}_n-\alpha)\big] = (1-\rho^2). \tag{1.8}$$

The equality (1.7) suggests one to consider the restricted estimator of ρ when $\alpha=\alpha_0$ as

$$\hat{\rho}_n = \tilde{\rho}_n - \frac{1-\tilde{\rho}_n^2}{1-\tilde{\rho}_n\alpha_0}(\tilde{\alpha}_n-\alpha_0). \tag{1.9}$$

Now due to uncertainty that $\alpha=\alpha_0$, we test $\mathcal{H}_0$: $\alpha=\alpha_0$ versus $\mathcal{H}_a$: $\alpha\neq\alpha_0$ based on the test–statistic

$$\mathcal{L}_n = \frac{n(\tilde{\alpha}_n-\alpha_0)^2(\tilde{\rho}_n-\alpha_0)^2}{(1-\alpha_0^2)(1-\tilde{\rho}_n\alpha_0)^2} = \frac{n(\tilde{\alpha}_n-\alpha_0)^2(\rho-\alpha_0)^2}{(1-\alpha_0^2)(1-\rho\alpha_0)^2} + o_P(1) \tag{1.10}$$

which under $\mathcal{H}_0$ and given (1.5) and (1.6), has asymptotically a χ_1^2 distribution.

Next we define the preliminary test estimator (PTE) as

$$\begin{aligned} \hat{\rho}_n^{PT} &= \hat{\rho}_n 1\mathrm{I}_{\{\mathcal{L}_n<\chi_1^2(\gamma)\}} + \tilde{\rho}_n 1\mathrm{I}_{\{\mathcal{L}_n\geq\chi_1^2(\gamma)\}} \\ &= \tilde{\rho}_n - (\tilde{\rho}_n-\hat{\rho}_n)1\mathrm{I}_{\{\mathcal{L}_n<\chi_1^2(\gamma)\}} \\ &= \tilde{\rho}_n - \frac{1-\tilde{\rho}_n^2}{1-\tilde{\rho}_n\alpha_0}(\tilde{\alpha}_n-\alpha_0)1\mathrm{I}_{\{\mathcal{L}_n<\chi_1^2(\gamma)\}} \\ &= \tilde{\rho}_n - \frac{1-\rho^2}{1-\rho\alpha_0}(\tilde{\alpha}_n-\alpha_0)1\mathrm{I}_{\{\mathcal{L}_n<\chi_1^2(\gamma)\}} + o_P(1) \end{aligned} \tag{1.11}$$

Instead of making extreme choices between $\hat{\rho}_n$ and $\tilde{\rho}_n$ using the PTE, a nicer compromise would be a smooth choice that depends on the value of $\mathcal{L}_n$. This can be reached by considering

$$\begin{aligned} \hat{\rho}_n^S &= \tilde{\rho}_n - (\tilde{\rho}_n-\hat{\rho}_n)\frac{c}{\sqrt{\mathcal{L}_n}} \quad \text{where } c \text{ is some constant} \\ &= \tilde{\rho}_n - \frac{c(1-\alpha_0^2)^{1/2}(1-\tilde{\rho}_n^2)}{\sqrt{n}|\tilde{\rho}_n-\alpha_0||\tilde{\alpha}_n-\alpha_o|}(\tilde{\alpha}_n-\alpha_0) \\ &= \tilde{\rho}_n - \frac{c(1-\alpha_0)^{1/2}(1-\rho^2)}{\sqrt{n}|\rho-\alpha_0||\alpha-\alpha_0|}(\tilde{\alpha}_n-\alpha_0) + o_P(1/\sqrt{n}) \end{aligned} \tag{1.12}$$

2 Bias and MSE under Contiguous Alternatives

We consider the alternative $\mathcal{H}_n : \alpha_n = \alpha_0 + \frac{\delta}{\sqrt{n}}$ where $\delta \neq 0$. We give two theorems: one for the bias and the other one is for the MSE.

Theorem 1. *We have*

$$\sqrt{n}(\tilde{\alpha}_n - \alpha_0) \Longrightarrow \mathcal{N}\big(\delta, \frac{1}{(\rho-\alpha_0)^2}(1-\alpha_0^2)(1-\rho\alpha_0)^2\big) \tag{2.1}$$

i)

$$\lim_{n\to\infty} \mathbb{E}[\sqrt{n}(\tilde{\rho}_n - \rho)] = 0$$

ii)

$$\begin{aligned}\lim_{n\to\infty} \mathbb{E}\big[\sqrt{n}(\hat{\rho}_n - \rho)\big] &= \lim_{n\to\infty} \mathbb{E}\big[\sqrt{n}(\tilde{\rho}_n - \rho\big] - \lim_{n\to\infty} \mathbb{E}\Big[\frac{1-\tilde{\rho}_n^2}{1-\tilde{\rho}_n\alpha_0}\sqrt{n}(\tilde{\alpha}_n - \alpha_o)\Big] \\ &= 0 - \frac{1-\rho^2}{1-\rho\alpha_o}\delta = -\frac{1-\rho^2}{1-\rho\alpha_o}\delta = -C(\rho,\alpha_0)\Delta\end{aligned}$$

where we put

$$C(\rho,\alpha_0) = \frac{(1-\rho^2)(1-\alpha_0)^{1/2}}{\rho-\alpha_0} \qquad \text{and} \qquad \Delta = \frac{\delta(\rho-\alpha_0)}{(1-\alpha_0^2)^{1/2}(1-\rho\alpha_0)}$$

iii)

$$\lim_{n\to\infty} \mathbb{E}\big[\sqrt{n}(\hat{\rho}_n^{PT} - \rho)\big] = -C(\rho,\alpha_0)\Delta H_3(\chi_1^2(\gamma), \Delta^2)$$

where $H_m(x,\Delta^2)$ denotes the cdf at x of a noncentral χ^2 distribution with m degrees of freedom and noncentrality parameter $\Delta^2/2$ and where $\chi_m^2(\gamma)$ is the γ–level critical value under $H_m(x,0)$.

iv)

$$\lim_{n\to\infty} \mathbb{E}\big[\sqrt{n}(\hat{\rho}_n^S - \rho)\big] = -\frac{c(1-\alpha_0^2)^{1/2}(1-\rho^2)}{|\rho-\alpha_0|}[2\Phi(\Delta)-1] = -c|C(\rho,\alpha_0)|[2\Phi(\Delta)-1]$$

where Φ is the cdf of the standard normal distribution.

Theorem 2.

i)
$$\lim_{n\to\infty} \mathbb{E}\big[n(\tilde{\rho}_n - \rho)^2\big] = \frac{1}{(\rho-\alpha_0)^2}(1-\rho^2)(1-\rho\alpha_0)^2$$

and hence

$$\begin{aligned}A\mathrm{Var}(\sqrt{n}(\tilde{\alpha}_n) &= \frac{1}{(\rho-\alpha_0)^2}(1-\alpha_0^2)(1-\rho\alpha_0)^2 \\ &= \frac{1-\alpha_0^2}{1-\rho^2}A\mathrm{Var}(\sqrt{n}\tilde{\rho}_n),\end{aligned} \tag{2.2}$$

ii)

$$\lim_{n\to\infty}\mathbb{E}\big[n(\hat\rho_n-\rho)^2\big]=(1-\rho^2)+C^2(\rho,\alpha_0)\Delta^2$$

iii)

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}\big[n(\hat\rho_n^{PT}-\rho)^2\big] &= \mathrm{Var}(\sqrt{n}\tilde\rho_n)\Big[1-K(\rho,\alpha_0)\Big\{H_3(\chi_1^2(\gamma),\Delta^2)\\ &\quad -\Delta^2\Big(H_3(\chi_1^2(\gamma),\Delta^2)-H_5(\chi_1^2(\gamma),\Delta^2)\Big)\Big\}\Big]\end{aligned}$$

where for a random variable Y_n, $A\mathrm{Var}(Y_n)$ represents the asymptotic variance and

$$K(\rho,\alpha_0)=\frac{(1-\rho^2)(1-\alpha_0^2)}{(1-\rho\alpha_0)^2}$$

vi)

$$\lim_{n\to\infty}\mathbb{E}\big[n(\hat\rho_n^{S}-\rho)^2\big]=A\mathrm{Var}(\sqrt{n}\tilde\rho_n)\left[1+K(\rho,\alpha_0)\frac{(1-\rho\alpha_0)^2}{(\rho-\alpha_0)^2}\frac{2}{\pi}\Big(1-2e^{-\Delta^2/2}\Big)\right]$$

Proof. (Theorem 1)

Of course (2.1) is straightforward from (1.5).

i) follows from (1.7).
ii) Obvious.
iii) We use the following result: (see Judge and Bock (1978) and Saleh (2006)). If $Z\sim\mathcal{N}(\Delta,1)$ then

$$\mathbb{E}(Zf(Z^2))=\Delta\mathbb{E}\big[f(\chi_3^2(\Delta^2))\big] \tag{2.3}$$

and

$$\mathbb{E}(Z^2f(Z^2))=\mathbb{E}\big[f(\chi_3^2(\Delta^2))\big]+\Delta^2\mathbb{E}\big[f(\chi_5^2(\Delta^2))\big] \tag{2.4}$$

We will use these equalities with the function

$$f(Z)=1\mathrm{I}_{\{Z<\chi_1^2(\gamma)\}}$$

and referring to (1.5),

$$Z=\frac{(\rho-\alpha_0)\sqrt{n}(\tilde\alpha_n-\alpha_0)}{(1-\alpha_0^2)^{1/2}(1-\rho\alpha_0)}=\big[A\mathrm{Var}(\sqrt{n}\tilde\alpha_n)\big]^{-1/2}\sqrt{n}(\tilde\alpha_n-\alpha_0) \tag{2.5}$$

That is, asymptotically $Z\sim\mathcal{N}(\Delta,1)$. We obtain

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}\big[\sqrt{n}(\hat\rho_n^{PT}-\rho)\big] &= -\frac{1-\rho^2}{1-\rho\alpha_0}\lim_{n\to\infty}\mathbb{E}\Big[(\sqrt{n}(\tilde\alpha_n-\alpha_0)1\mathrm{I}_{\{\mathcal{L}_n<\chi_1^2(\gamma)\}}\Big]\\ &= -\frac{1-\rho^2}{1-\rho\alpha_0}\frac{(1-\alpha_0^2)^{1/2}(1-\rho\alpha_0)}{(\rho-\alpha_0)}\Delta H_3(\chi_1^2(\gamma),\Delta^2)\\ &= -C(\rho,\alpha_0)\Delta H_3(\chi_1^2(\gamma),\delta^2).\end{aligned}$$

vi) Let U be a standard normal distribution. iv) follows from

$$\begin{aligned}\mathbb{E}\Big(\frac{\tilde{\alpha}_n-\alpha_0}{|\tilde{\alpha}_n-\alpha_0|}\Big) &= P\big[\sqrt{n}(\tilde{\alpha}_n-\alpha_0)>0\big]-P\big[\sqrt{n}(\tilde{\alpha}_n-\alpha_0)<0\big]\\ &= P(U>-\Delta)-P(U<-\Delta)=\Phi(\Delta)-1\end{aligned}$$

$\square$

Proof. (Theorem 2) i) can be shown from (1.6) since $\alpha=\alpha_n\to\alpha_0$.
ii) We have $\lim_{n\to\infty}\mathbb{E}\big[n(\hat{\rho}_n-\rho)^2\big]=\lim_{n\to\infty}\mathrm{Var}\big[\sqrt{n}(\hat{\rho}_n-\rho)\big)\big]+\lim_{n\to\infty}\mathbb{E}\big[\sqrt{n}(\hat{\rho}_n-\rho)\big]^2$. We only need to show that $\lim_{n\to\infty}\mathrm{Var}\big[\sqrt{n}(\hat{\rho}_n-\rho)\big]=(1-\rho^2)$ since $\lim_{n\to\infty}\mathbb{E}\big[n(\hat{\rho}_n-\rho)\big]$ is given by ii) in theorem 1.

$$\begin{aligned}\lim_{n\to\infty}\mathrm{Var}\big[\sqrt{n}(\hat{\rho}_n-\rho)\big] &= \lim_{n\to\infty}\mathrm{Var}\big[\sqrt{n}(\tilde{\rho}_n-\rho)-\frac{1-\tilde{\rho}_n^2}{1-\tilde{\rho}_n\alpha_0}\sqrt{n}(\tilde{\alpha}_n-\alpha_0)\big]\\ &= \frac{1}{(\rho-\alpha_0)^2}\Bigg[(1-\rho^2)(1-\rho\alpha_0)^2\\ &\quad+\left(\frac{1-\rho^2}{1-\rho\alpha_0}\right)^2(1-\alpha_0^2)(1-\rho\alpha_0)^2\\ &\quad-2\frac{1-\rho^2}{1-\rho\alpha_0}(1-\rho^2)(1-\alpha_0^2)(1-\rho\alpha_0)\Bigg]\\ &= (1-\rho^2)\end{aligned}$$

iii) Using (2.4), (2.3), and (1.7), and some conditional expectation arguments, we get

$$\begin{aligned}\lim_{n\to\infty}\mathbb{E}\big[n(\hat{\rho}_n^{PT}-\rho)^2\big] &= \lim_{n\to\infty}\mathbb{E}\left(\big[\sqrt{n}(\tilde{\rho}_n-\rho)-\frac{1-\rho^2}{1-\rho\alpha_0}\sqrt{n}(\tilde{\alpha}_n-\alpha_0)1\mathrm{I}_{\{\mathcal{L}_n<\chi_1^2(\gamma)\}}\big]^2\right)\\ &= A\mathrm{Var}(\sqrt{n}\tilde{\rho}_n)-\left(\frac{1-\rho^2}{1-\rho\alpha_0}\right)^2A\mathrm{Var}(\sqrt{n}\tilde{\alpha}_n)H_3(\chi_1^2(\gamma),\Delta^2)\\ &\quad+\Delta^2\left(\frac{1-\rho^2}{1-\rho\alpha_0}\right)^2\\ &\quad A\mathrm{Var}(\sqrt{n}\tilde{\alpha}_n)\big[2H_3(\chi_1^2(\gamma),\Delta^2)-H_5(\chi_1^2(\gamma),\Delta^2)\big].\end{aligned}$$

Replacing $A\mathrm{Var}(\sqrt{n}\tilde{\alpha}_n)$ by its value in (2.2) we get the right hand side of iii).
vi) We will use the fact that if $Z\sim\mathcal{N}(\Delta,1)$, then

$$\mathbb{E}(|Z|)=\sqrt{\frac{2}{\pi}}e^{-\Delta^2/2}+\Delta\big[2\Phi(\Delta)-1\big] \tag{2.6}$$

From (1.12), we can write (in a similar way as in the calculations in iii) and applying (2.6) with Z as in (2.5),

$$\begin{aligned}
\lim_{n\to\infty}\mathbb{E}\big[n(\hat\rho_n^S-\rho)^2\big] &= \lim_{n\to\infty}\mathbb{E}\Bigg(\Big[\sqrt{n}(\tilde\rho_n-\rho)-\frac{c(1-\alpha_0^2)^{1/2}(1-\tilde\rho_n^2)}{|\tilde\rho_n-\alpha_0|}\frac{\tilde\alpha_n-\alpha_0}{|\tilde\alpha_n-\alpha_0|}\Big]^2\Bigg)\\
&= A\mathrm{Var}(\sqrt{n}\tilde\rho_n)+\frac{c^2(1-\alpha_0^2)(1-\rho^2)^2}{(\rho-\alpha_0)^2}\\
&\quad -2\frac{c(1-\alpha_0^2)^{1/2}(1-\rho^2)}{|\rho-\alpha_0|}\frac{1-\rho^2}{1-\rho\alpha_0}\Bigg[\Big(A\mathrm{Var}(\sqrt{n}\tilde\alpha_n)\Big)^{1/2}\\
&\quad \Bigg(\sqrt{\frac{2}{\pi}}e^{-\Delta^2/2}+\Delta\big[2\Phi(\Delta)-1\big]\Bigg)-\delta\big[2\Phi(\Delta)-1\big]\Bigg]\\
&= A\mathrm{Var}(\sqrt{n}\tilde\rho_n)+\frac{(1-\alpha_0^2)(1-\rho^2)^2}{(\rho-\alpha_0)^2}\Big[c^2-2c\sqrt{\frac{2}{\pi}}e^{-\Delta^2/2}\Big]
\end{aligned}$$

The value of c that minimizes this quantity is $c=\sqrt{\frac{2}{\pi}}e^{-\Delta^2/2}$. Making c independent of Δ^2, we chose $c=\sqrt{2/\pi}$. Plugging this optimal value in the previous equality we get the minimal value for MSE

$$\lim_{n\to\infty}\mathbb{E}\big[n(\hat\rho_n^S-\rho)^2\big]=A\mathrm{Var}(\sqrt{n}\tilde\rho_n)+\frac{(1-\alpha_0^2)(1-\rho^2)^2}{(\rho-\alpha_0)^2}\frac{2}{\pi}\Big[1-2e^{-\Delta^2/2}\Big]$$

which can be written in the form of the right hand side of vi). □

3 MSE under Fixed Alternative Hypothesis

In this section we show that under fixed alternative

$$\alpha=\alpha_0+\delta \tag{3.1}$$

there is no gain in terms of MSE when we consider the PTE estimator.

Theorem 3. *Under (3.1) we have*

$$\lim_{n\to\infty}\mathbb{E}\Big[n(\hat\rho_n^{PT}-\tilde\rho_n)^2\Big]=0$$

Proof. We have

$$\begin{aligned}
\mathbb{E}\Big[n(\hat\rho^{PT}-\tilde\rho)^2\Big] &= n\Big(\frac{1-\tilde\rho_n^2}{1-\tilde\rho_n\alpha_0}\Big)^2(\tilde\alpha_n-\alpha_0)^2 1\mathrm{I}_{\{\mathcal{L}_n<\chi_1^2(\gamma)\}}\\
&= n\frac{(\tilde\alpha_n-\alpha_0)^2}{1-\alpha_0^2}\Big(\frac{\rho-\alpha_0}{1-\rho\alpha_0}\Big)^2\Big(\frac{1-\rho^2}{\rho-\alpha_0}\Big)^2(1-\alpha_0^2)1\mathrm{I}_{\{\mathcal{L}_n<\chi_1^2(\gamma)\}}+o_P(1)\\
&= \Big(\frac{1-\rho^2}{\rho-\alpha_0}\Big)^2(1-\alpha_0^2)\mathcal{L}_n 1\mathrm{I}_{\{\mathcal{L}_n<\chi_1^2(\gamma)\}}+o_P(1).
\end{aligned}$$

Thus we need to show that

$$\lim_{n\to\infty} \mathbb{E}\Big[\mathcal{L}_n 1\mathrm{I}_{\{\mathcal{L}_n<\chi_1^2(\gamma)\}}\Big] = 0. \tag{3.2}$$

Under the fixed alternative hypothesis (3.1), $\mathcal{L}_n = Z^2$ where asymptotically

$$Z \sim \mathcal{N}(\Delta, 1) \qquad \text{with} \qquad \Delta^2 = \frac{n\delta^2(\rho-\alpha_0)^2}{(1-\alpha_0^2)(1-\rho\alpha_0)^2}.$$

Now using (2.4) and writing $\Delta^2 = na$ and $\chi_1^2(\gamma) = b$ the expectation in (3.2) can be written as

$$P(\chi_3^2(na) \le b) + naP(\chi_5^2(na) \le b) \tag{3.3}$$

With $U \sim \mathcal{N}(0,1)$, clearly each probability in (3.3) is bounded by

$$\begin{aligned} P[(U-na)^2 \le b] &= P[-\sqrt{b}+na \le U \le \sqrt{b}+na] \le P[U \ge na-\sqrt{b}] \\ &\sim P[U \ge na] \sim \frac{1}{\sqrt{2\pi}}\frac{1}{na}e^{-(na)^2/2} \quad \text{as } n\to\infty. \end{aligned}$$

The last equivalence is well known, which may be shown via integration by parts. Therefore (3.2) is established.

We mention that this is not the case for the smoothed version ρ_n^S or for the restricted estimator $\hat{\rho}_n$. Actually we can see that

$$\lim_{n\to\infty} \mathbb{E}\Big[n(\hat{\rho}_n^S - \tilde{\rho}_n)^2\Big] = \frac{c^2(1-\alpha_0^2)(1-\rho^2)^2}{(\rho-\alpha_0)^2}.$$

□

4 Comparative Study

4.1 Comparing Quadratic Bias Functions

We note that from theorem 1, the usual PMLE estimate $\tilde{\rho}_n$ is asymptotically unbiased and the alternative estimates $\hat{\rho}_n$, $\hat{\rho}_n^{PT}$ and $\hat{\rho}_n^S$ have the following quadratic bias functions:

$$i) \qquad B(\hat{\rho}_n)^2 = C^2(\rho,\alpha_0)\Delta^2,$$

$$ii) \qquad B(\hat{\rho}_n^{PT})^2 = C^2(\rho,\alpha_0)\Delta^2 H_3^2,$$

and

$$iii) \quad B(\hat{\rho}_n^S)^2 = c^2C^2(\rho,\alpha_0)[2\Phi(\Delta)-1]^2.$$

We note that for $\Delta = 0$,

$$B^2(\tilde{\rho}_n) = B^2(\hat{\rho}_n) = B^2(\rho_n^{PT}) = B^2(\rho_n^S) = 0$$

i.e. all the estimators are asymptotically unbiased. Also as $\Delta^2 \to \infty$, we get $B^2(\rho_n^{PT}) \to 0$ while $B^2(\rho_n^S) \to \frac{2}{\pi}K^2(\rho, \alpha_0)$, if we keep the value $\frac{2}{\pi}$ for c. Hence, as $\Delta^2 \to \infty$,

$$0 = B^2(\rho_n^{PT}) < B^2(\rho_n^S) < B^2(\hat{\rho}_n),$$

and for a given Δ, we always have $B^2(\rho_n^{PT}) < B^2(\hat{\rho}_n)$ but the other comparisons will depend on both Δ and the level γ. Clearly we observe that in terms of smallest bias, the PTE estimator is better than the unrestricted estimator but the shrinkage estimator can give good results if we decide to choose the constant c very small. However, the usual PMLE remains the best one since it is asymptotically unbiased. The following graphical presentation in Figure (1) shows these properties.

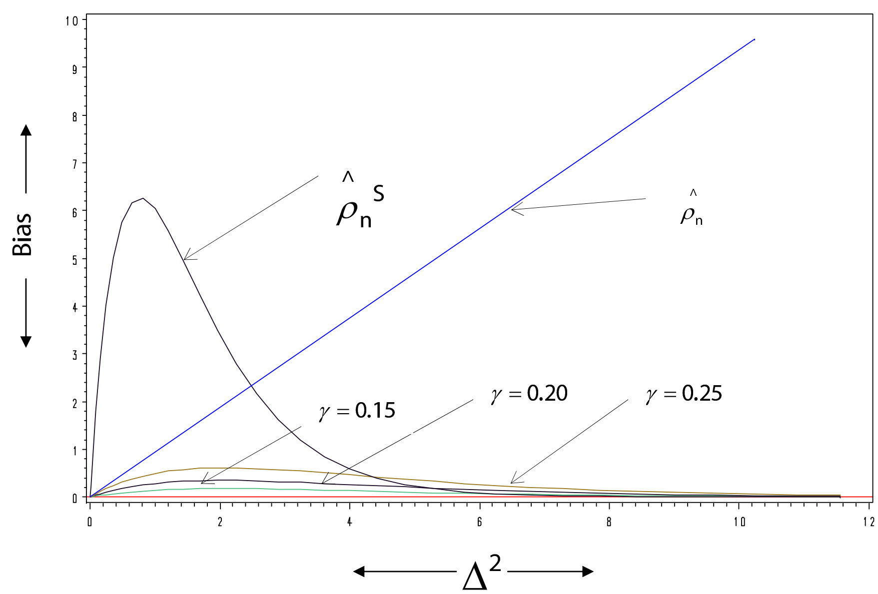

Figure 1: Graph of the asymptotic quadratic bias of the three estimators.

4.2 Asymptotic Relative Efficiency (ARE)

In this section we compare the proposed estimators to the PMLE which we call here unrestricted estimator (UE). This comparison will be based on the ADMSE and the so-called asymptotic relative efficiency (ARE).

Comparing the PTE Against the UE:

Denote by $\mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n)$ the ARE of $\hat{\rho}_n^{PT}$ with respect to $\tilde{\rho}_n$. That is the quotient of their reciprocal MSE's. So we have

$$\mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n) = \frac{\mathrm{MSE}(\sqrt{n}\tilde{\rho}_n)}{\mathrm{MSE}(\sqrt{n}\hat{\rho}_n^{PT})} = \Big[1 - K(\rho, \alpha_0)H_3 + \Delta^2(2H_3 - H_5)\Big]^{-1}$$

In order to make notations easier, we omitted the arguments $(\chi_1^2(\gamma), \Delta^2)$ in the functions H_3 and H_5.

From the previous equation we conclude the following: The graph of $\mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n)$ as a function of Δ^2 for fixed γ, is decreasing crossing the 1–line to a minimum at $\Delta^2 = \Delta^2_{\min}(\gamma)$, then it increases towards the 1–line as $\Delta^2 \to \infty$. Thus, if $\Delta^2 \leq (H_3/(2H_2 - H_3)$ then $\hat{\rho}_n^{PT}$ is better than $\tilde{\rho}$ while if $\Delta^2 \geq (H_3/(2H_2 - H_3)$, $\tilde{\rho}$ is better than $\hat{\rho}_n^{PT}$.

The maximum of $\mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n)$ at $\Delta^2 = 0$ is given by

$$\max_{\Delta^2} \mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n) = \Big[1 - KH_3\Big]^{-1}, \quad K = \frac{(1-\rho^2)(1-\alpha_0^2)}{(1-\rho\alpha_0)^2}$$

where $H_3 = H_3(\chi_1^2(\gamma), 0)$ for all $\gamma \in A$, the set of all possible values of γ and $K = K(\rho, \alpha_0)$.

The value of $\max_{\Delta^2} \mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n)$ decreases as γ increases, while if

$\gamma = 0$ and Δ^2 varies, the graph of $\mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n) = 1$ and $\mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n)$ intersects at $\Delta^2 = 1$. In general $\mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n)$ at $\gamma = \gamma_1$ and γ_2 intersect within the interval $0 \leq \Delta^2 \leq 1$, the value of Δ^2 at the intersection increases as γ increases, therefore, for two values of γ, the $\mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n)$'s will always be below the 1–line.

In order to obtain optimum level of significance γ^* for the application of PTE, we prefix a minimum guaranteed efficiency, say E_0 and follow the following procedure

i) If $0 \leq \Delta^2 \leq 1$, use $\tilde{\rho}_n$ because $\mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n)$ is always ≥ 1 in this region of Δ^2. However, Δ^2 is generally unknown and there is no way to choose uniformly best estimator and look for the value of γ in the set

$$A_\gamma = \{\gamma \ \mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n) \geq E_0\}.$$

ii) The PT estimator chosen maximizes $\mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n)$ over all $\gamma \in A_\gamma$ and Δ^2. Thus we solve the equation

$$\min_{\Delta^2} \mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n) = \mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n)(\gamma, \Delta_0) = E_0.$$

The solution γ^* obtained this way gives a PTE with minimum guaranteed efficiency of E_0 which may increase to $\mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n)$ at $\Delta^2 = 0$.

The sample table in Table (1) gives the minimum and maximum $\mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n)$ for chosen values of K over $\gamma = 0.05(0.05)0.5$.

To illustrate the use of the table, let us set $E_0 = 0.8$ as the minimum guaranteed $\mathrm{ARE}(\hat{\rho}_n^{PT} : \tilde{\rho}_n)$ for $K = 0.5$. Then we look for 0.8 or near it for $K = 0.5$ and find $\gamma^* = 0.2$. Hence, using a PTE with $\gamma^* = 0.2$ we obtain a minimum guaranteed ARE of 0.8 with a possible maximum ARE of 1.21 if Δ^2 is close to 0.

Table 1: Maximum and minimum ARE of the PTE for selected K, γ values

γ	K	0.10	0.20	0.30	0.40	0.50	0.60	0.70	0.80	0.90	1.00
0.05	ARE^{max}	1.08	1.17	1.28	1.41	1.56	1.76	2.02	2.36	2.85	3.58
	ARE^{min}	0.87	0.77	0.69	0.63	0.58	0.53	0.49	0.46	0.43	0.41
	Δ^2	4.84	4.84	4.84	4.84	4.84	4.84	4.84	4.84	4.84	4.84
0.10	ARE^{max}	1.06	1.13	1.20	1.29	1.39	1.51	1.65	1.81	2.02	2.28
	ARE^{min}	0.91	0.84	0.78	0.72	0.68	0.64	0.60	0.57	0.54	0.51
	Δ^2	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00	4.00
0.15	ARE^{max}	1.05	1.10	1.15	1.22	1.28	1.36	1.45	1.55	1.66	1.79
	ARE^{min}	0.94	0.88	0.83	0.79	0.75	0.71	0.68	0.65	0.62	0.59
	Δ^2	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61	3.61
0.20	ARE^{max}	1.04	1.08	1.12	1.16	1.21	1.27	1.32	1.39	1.46	1.54
	ARE^{min}	0.95	0.91	0.87	0.83	0.80	0.77	0.74	0.71	0.69	0.66
	Δ^2	3.24	3.24	3.24	3.24	3.24	3.24	3.24	3.24	3.24	3.24
0.25	ARE^{max}	1.03	1.06	1.09	1.12	1.16	1.20	1.24	1.28	1.33	1.38
	ARE^{min}	0.96	0.93	0.90	0.87	0.84	0.81	0.79	0.77	0.75	0.72
	Δ^2	3.24	3.24	3.24	3.24	3.24	3.24	3.24	3.24	3.24	3.24
0.30	ARE^{max}	1.02	1.05	1.07	1.09	1.12	1.15	1.18	1.21	1.24	1.28
	ARE^{min}	0.97	0.95	0.92	0.90	0.87	0.85	0.83	0.81	0.79	0.78
	Δ^2	2.89	2.89	2.89	2.89	2.89	2.89	2.89	2.89	2.89	2.89
0.35	ARE^{max}	1.02	1.03	1.05	1.07	1.09	1.11	1.13	1.16	1.18	1.20
	ARE^{min}	0.98	0.96	0.94	0.92	0.90	0.89	0.87	0.85	0.84	0.82
	Δ^2	2.89	2.89	2.89	2.89	2.89	2.89	2.89	2.89	2.89	2.89
0.40	ARE^{max}	1.01	1.03	1.04	1.05	1.07	1.08	1.10	1.11	1.13	1.15
	ARE^{min}	0.98	0.97	0.95	0.94	0.93	0.91	0.90	0.89	0.87	0.86
	Δ^2	2.89	2.89	2.89	2.89	2.89	2.89	2.89	2.89	2.89	2.89
0.45	ARE^{max}	1.01	1.02	1.03	1.04	1.05	1.06	1.07	1.08	1.10	1.11
	ARE^{min}	0.99	0.98	0.97	0.95	0.94	0.93	0.92	0.91	0.90	0.89
	Δ^2	2.89	2.89	2.89	2.89	2.89	2.89	2.89	2.89	2.89	2.89
0.50	ARE^{max}	1.01	1.01	1.02	1.03	1.04	1.04	1.05	1.06	1.07	1.08
	ARE^{min}	0.99	0.98	0.97	0.97	0.96	0.95	0.94	0.94	0.93	0.92
	Δ^2	2.56	2.56	2.56	2.56	2.56	2.56	2.56	2.56	2.56	2.56

Comparing the Shrinkage Estimator Against the UE

We have

$$\mathrm{ARE}(\hat{\rho}_n^S : \tilde{\rho}_n) = \left[1 + \frac{(1-\rho^2)(1-\alpha_0^2)}{(\rho-\alpha_0)^2}\frac{2}{\pi}\Big[1 - 2e^{-\Delta^2/2}\Big]\right]^{-1}$$

Following a similar discussion as we did for the relative efficiency for PTE, we can see from this equation that for

$$\Delta^2 \le \ln 4 \qquad \hat{\rho}_n^S \text{ is better than } \tilde{\rho}_n$$

and for

$$\Delta^2 \ge \ln 4 \qquad \tilde{\rho}_n \text{ is better than } \hat{\rho}_n^S$$

Also, as $\Delta^2 \to \infty$,

$$\mathrm{ARE}(\hat{\rho}_n^S : \tilde{\rho}_n) \to \left[1 + \frac{(1-\rho^2)(1-\alpha_0^2)}{(\rho-\alpha_0)^2}\frac{2}{\pi}\right]^{-1} \le 1$$

which is the lower bound of $\mathrm{ARE}(\hat{\rho}_n^S : \tilde{\rho}_n)$.

The maximum $\mathrm{ARE}(\hat{\rho}_n^S : \tilde{\rho}_n)$ is $\Big[1 - \frac{2}{\pi}K\frac{(1-\rho\alpha_0)^2}{(\rho-\alpha_0)^2}\Big]^{-1} \ge \big[1 - KH_3(\chi_1^2(\gamma), 0)\big]^{-1}$

if

$$H_3(\chi_1^2(\gamma), 0) \le \frac{2}{\pi}\frac{(1-\rho\alpha_0)^2}{(\rho-\alpha_0)^2}.$$

Comparing the Restricted Estimator RE Against UE

From vi) in theorem 2, we can rewrite the asymptotic MSE of $\sqrt{n}\hat{\rho}_n$ as

$$(1-\rho^2)\Big[1 + (1-\rho^2)\frac{1-\alpha_0^2}{(\rho-\alpha_0)^2}\Delta^2\Big]$$

and we can write the MSE of $\sqrt{n}\tilde{\rho}_n$ as

$$(1-\rho^2)\frac{(1-\rho\alpha_0)^2}{(\rho-\alpha_0)^2} = (1-\rho^2)\Big[1 + \frac{(1-\rho^2)(1-\alpha_0^2)}{(\rho-\alpha_0)^2}\Big]$$

and therefore we can conclude

$$\text{if } \Delta^2 \le 1 \qquad \text{then RE is better than UE}$$

and

$$\text{if } \Delta^2 \ge 1 \qquad \text{then UE is better than RE}$$

The graphs in the Figure (2) of the ARE's depict the ARE properties of the estimators.

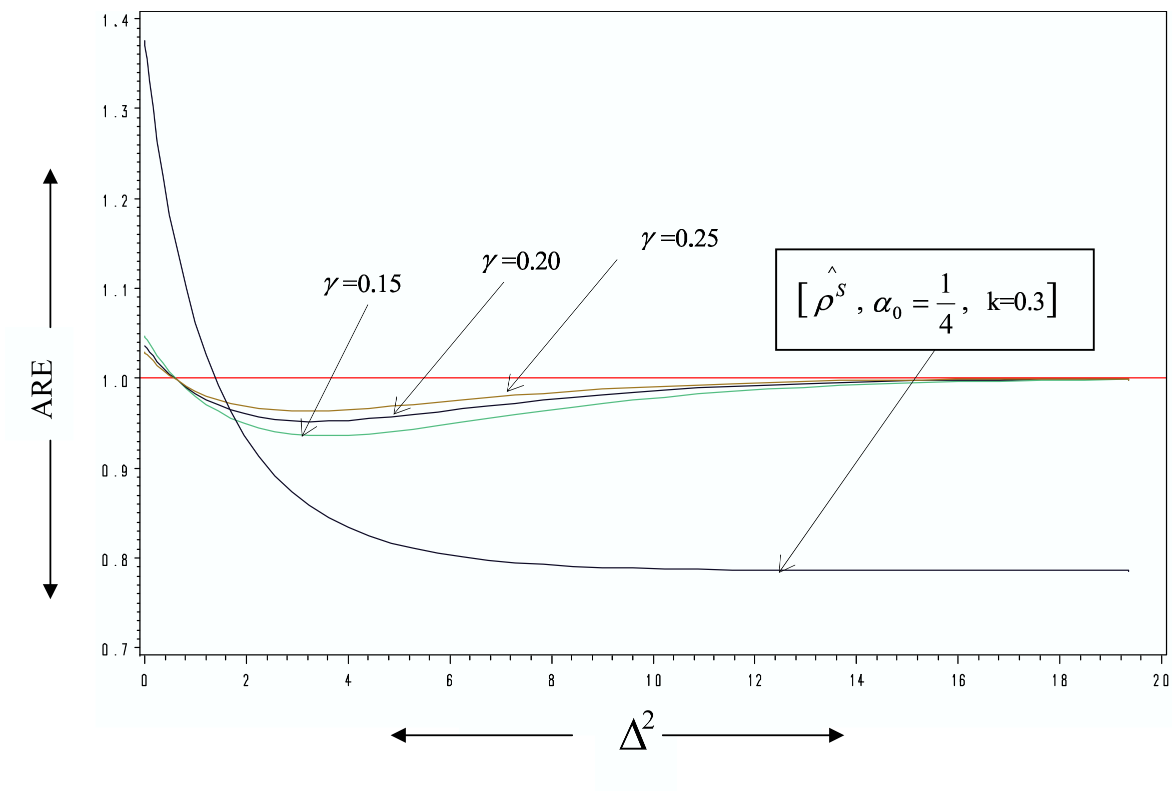

Figure 2: Graph of ARE for shrinkage estimator for selected values of γ, α, K.

Acknowledgments

The author is grateful Professor A. K. Md. E. Saleh for his invaluable help and his insightful suggestions about this paper.

References

[1] Bancroft, T.A. (1944). On biases in estimation due to the use of preliminary tests of significance. *Ann. Math. Statist.* **15**, 190–204.

[2] Dzhaparidze, K. (1986). *Parameter Estimation and Hypothesis Testing in Spectral Analysis of Stationary Time Series.* Springer–Verlag.

[3] Efron, B. and Morris, C. (1972). Limiting the risk of Bayes and Empirical Bayes estimators. Part II: the empirical Bayes case. *J. Amer. Statist. Assoc.*. **67**, 130–139.

[4] Efron, B. Morris, C. (1973). Stein's estimation rule and competitors. An empirical Bayes approach. *J. Amer. Statist. Assoc.* **68**, 117–130.

[5] Judge, G.G. and Bock, M.E. (1978). The Statistical Implications of Pre–test and Stein–rule Estimation in Econometrics. North–Holland, New–York.

[6] Saleh A.K.Md.E. (1992.) On Shrinkage Estimation of The Parameters of An Autoregressive Gaussian Process. *Theory of Probab. and its Appl.* 250–260.

[7] Saleh, A.K.Md.E. (2006). Theory of Preliminary Test and Stein–Type Estimation With Applications. Wiley–InterSience. Wiley & Sons Inc. New York.

[8] Sclove, S.L., Morris, C. and Rao, C.R. (1972). Non–optimality of preliminary test estimators for the mean of a multivariate normal distribution. *Ann. Math. Statis.* **43**, 1481–1490.

[9] Stein, C. (1956). Inadmissibility of the usual estimator for the mean of multivariate normal distribution. Proceedings of the 3rd Berkeley Symposium on Math. Statist. and Probab. Univ. of California Press, Berkeley. **1**, 197–206.

[10] STEIN, C. (1981). Estimation of the mean of a multivariate normal distribution. *Ann. of Statist.* **9**, 1135–1151.