CentER
for
Economic Research

Discussion paper

[redacted]

TILBURG UNIVERSITY

No. 8907

POSTERIOR DENSITIES FOR NONLINEAR REGRESSION WITH EQUICORRELATED ERRORS

by J. Osiewalski

February, 1989

Posterior Densities for Nonlinear Regression with Equicorrelated Errors

by

Jacek Osiewalski
Academy of Economics
ul. Rakowicka 27
31-510 Kraków
Poland

JEL classification: 211

Keywords: nonlinear models, Bayesian analysis, equicorrelation

Abstract:
For a nonlinear regression model with a constant term, it is shown that - under diffuse priors of the constant term and of the error precision - the assumption of equicorrelated errors (instead of uncorrelated ones) has no new consequences on Bayesian estimation of the (nonlinear) regression parameters (except for the constant term).

February 1989

Acknowledgement: Comments by Mark Steel are gratefully acknowledged.

1. Introduction

Main non-Bayesian results concerning linear regression with an intercept and with equicorrelated observations were obtained long time ago. Assuming joint normality of observations, Halperin (1951) showed that certain estimators and tests of significance used in regression analysis when observations are independent are equally valid in the case of equicorrelated observations. McElroy (1969) proved that, in a linear regression model with an intercept, OLS estimators are BLU if and only if the errors have the same variances and the same nonnegative coefficient of correlation between each pair; see also Balestra (1970).

Bayesian results for a linear model with equicorrelated disturbances are presented in Osiewalski (1987); it is shown that under diffuse priors of all the regression parameters and of the error variance:

1. the posterior of the correlation parameter is equal to its prior,
2. the marginal posterior of the regression parameters (except for the constant term) is the same as in the case of uncorrelated disturbances,
3. the posterior mean of the intercept is unaffected by equicorrelation, but its posterior variance is seriously affected and can be infinite for some priors.

The purpose of this paper is to extend these results to a nonlinear model with a constant term and to general prior assumptions on (nonlinear) regression parameters other than the constant term.

1.1 Notation

Throughout this note p(.) denotes a probability density function (PDF) with special notation for PDF's of gamma, normal and t distributions. For $x \epsilon R^k$, $p_N^k(x|c,W)$ denotes a k-variate normal PDF with a mean vector c and a covariance matrix W, and $p_S^k(x|\tau,c,\Upsilon)$ denotes a k-variate Student t PDF with τ degrees of freedom, a noncentrality vector (a mean vector, if $\tau > 1$) c and a precision matrix Υ. For $w \epsilon R_+^1$,

$$p_G(w|a,b) = \frac{b^a}{\Gamma(a)} w^{a-1} \exp(-bw),$$

that is a gamma PDF with parameters $a > 0$, $b > 0$.

2. Posterior densities under uncorrelated and equicorrelated errors - a comparison

We consider two cases of the following (nonlinear) regression model with normal errors:

$$y = h(Z;\vartheta) + \delta e + \mu, \quad \mu \sim N(0,\omega^{-1}V), \tag{1}$$

where $y \in R^n$ is observed; $\vartheta \in \Theta \subset R^{k-1}$, $\delta \in R^1$ and $\omega \in R^1_+$ are unknown parameters; e is a column vector of ones (thus δ is a constant term); Z is a known (nonrandom) matrix of $(n \times \tau)$ fixed values of τ explanatory variables; $h(Z;\vartheta)$ is a known vector function (without a constant term). We assume that the vector function h (as a function of ϑ only, given Z) is sufficiently well-behaved to ensure the possibility of integration with respect to ϑ. In order to save space, we will write h_ϑ instead of $h(Z;\vartheta)$.

Model (1) with a scalar covariance matrix (i.e. with $V=I_n$) is the first case considered here. Model (1) with equicorrelated errors, that is with

$$V = V_\varphi = (1-\varphi)I_n + \varphi ee' \quad , \varphi \in [0,1),$$

constitutes the second case. In both cases we assume diffuse priors on the constant term (δ) and the precision parameter (ω).

A. If $V = I_n$, then the conditional data density $p(y| Z, \vartheta, \delta, \omega)$ is the following n-variate normal density:

$$p(y|Z, \vartheta, \delta, \omega) = p_N^n (y|h_\vartheta + \delta e, \omega^{-1}I_n).$$

Under the prior

$$p(\vartheta,\ \delta,\ \omega) \propto \omega^{-1} p(\vartheta), \quad \vartheta\varepsilon\Theta,\ \delta\varepsilon R^1,\ \omega\varepsilon R^1_+,$$

which uses Jeffreys' rule to express vague prior knowledge about ω, we obtain the following posterior:

$$p(\vartheta,\ \delta,\ \omega|y,\ Z) \propto p(\vartheta)\ \omega^{\frac{n}{2}-1} \exp[-\tfrac{\omega}{2}\ (y-h_\vartheta-\delta e)'(y-h_\vartheta-\delta e)].$$

Denoting

$$d_\vartheta = \frac{1}{n}\ e'\ (y-h_\vartheta),$$

$$s_\vartheta = (y-h_\vartheta)'(I_n - \frac{1}{n}\ ee')(y-h_\vartheta),$$

we can write the identity

$$(y-h_\vartheta-\delta e)'\ (y-h_\vartheta-\delta e) = s_\vartheta + n(\delta-d_\vartheta)^2.$$

Now it is possible to present the joint posterior density in the form:

$$p(\vartheta,\ \delta,\ \omega|y,\ Z) \propto p(\vartheta)\ s_\vartheta^{-\frac{n-1}{2}}\ p_G(\omega|\tfrac{n-1}{2},\ \tfrac{s_\vartheta}{2})\ p^1_N\ (\delta|d_\vartheta,\ \tfrac{1}{n\omega}). \tag{2}$$

From (2) we easily obtain the following marginal and conditional posterior densities:

$$p(\vartheta|y,Z) \propto p(\vartheta)\ s_\vartheta^{-\frac{n-1}{2}},$$

$$p(\omega|y,Z,\vartheta) = p_G(\omega|\ \tfrac{n-1}{2},\ \tfrac{s_\vartheta}{2}),$$

$$p(\delta|y,Z,\vartheta,\omega) = p^1_N(\delta|d_\vartheta,\ \tfrac{1}{n\omega}),$$

$$p(\delta|y,Z,\vartheta) = {}_0\!\int^\infty p(\delta|y,Z,\vartheta,\omega)\ p(\omega|y,Z,\vartheta)\ d\omega = p^1_S(\delta|n-1,\ d_\vartheta,\ n\ \tfrac{n-1}{s_\vartheta}).$$

In the case of a linear model with diffuse prior, that is when

$$\tau = k-1, \qquad \Theta = R^{k-1}, \qquad h_\vartheta = Z\vartheta, \qquad p(\vartheta) = \text{const},$$

one obtains the well-known posterior results for $\beta = [\delta\ \vartheta']'$, as in Zellner (1971) pp. 66-69.

B. If we assume equicorrelated disturbances, then we have one more unknown parameter, $\varphi\epsilon[0,1)$, and the conditional data density $p(y|Z, \vartheta, \delta, \omega, \varphi)$ is the following normal density:

$$p(y|Z, \vartheta, \delta, \omega, \varphi) = p_N^n\ (y|h_\vartheta + \delta e, \omega^{-1}\ V\varphi),$$

$$V\varphi = (1-\varphi)\ I_n + \varphi ee'.$$

Under the prior

$$p(\vartheta, \delta, \omega, \varphi)\ \alpha\ \omega^{-1}\ p(\vartheta,\varphi) = \omega^{-1}p(\vartheta)p(\varphi|\vartheta),\ \vartheta\epsilon\Theta,\ \delta\epsilon R^1,\ \omega\epsilon R^1_+,\ \varphi\epsilon[0,1),$$

where $p(\varphi|\vartheta)$ is proper (for every $\vartheta\epsilon\Theta$), we obtain the following joint posterior density

$$p(\vartheta, \delta, \omega, \varphi|y, Z)\ \alpha\ p(\vartheta)\ p(\varphi|\vartheta)|V\varphi|^{-\frac{1}{2}}\ \omega^{\frac{n}{2}-1}\ \exp[-\tfrac{\omega}{2}\ (y-h_\vartheta-\delta e)'V\varphi^{-1}$$

$$(y-h_\vartheta-\delta e)].$$

Now let us take into account that

$$V\varphi^{-1} = \frac{1}{1-\varphi}\ [\ I_n - \frac{\varphi}{1+(n-1)\varphi}\ ee'],$$

see e.g. Graybill (1969) p. 172; we have

$$(y-h_\vartheta-\delta e)'\ V\varphi^{-1}\ (y-h_\vartheta-\delta e) =$$

$$= \frac{1}{1-\varphi}\ \{\ (y-h_\vartheta-\delta e)'\ (y-h_\vartheta-\delta e) - \frac{\varphi}{1+(n-1)\varphi}\ [(y-h_\vartheta-\delta e)'e]^2\} =$$

$$= \frac{1}{1-\varphi} s_\vartheta + \frac{n}{1+(n-1)\varphi} (\delta - d_\vartheta)^2.$$

The joint posterior density can be presented as

$$p(\vartheta, \delta, \omega, \varphi | y, Z) \propto p(\varphi|\vartheta) p(\vartheta) |V\varphi|^{-\frac{1}{2}} \omega^{\frac{n}{2}-1} \exp[- \frac{\omega s_\vartheta}{2(1-\varphi)}]$$

$$\exp[- \frac{1}{2} \frac{n\omega}{1+(n-1)\varphi} (\delta - d_\vartheta)^2].$$

Since the determinant of $V\varphi$ takes the form

$$|V\varphi| = [1 + (n-1)\varphi] (1-\varphi)^{n-1},$$

see e.g. Graybill (1969) p. 172, we can write

$$p(\vartheta, \delta, \omega, \varphi | y, Z) \propto p(\varphi|\vartheta) p(\vartheta) s_\vartheta^{-\frac{n-1}{2}} p_G(\omega | \frac{n-1}{2}, \frac{s_\vartheta}{2 (1-\varphi)})$$

$$p_N^1 (\delta | d_\vartheta, \frac{1+(n-1)\varphi}{n\omega}),$$

or equivalently

$$p(\vartheta,\delta,\omega,\varphi|y,Z) = p(\vartheta|y,Z) p(\varphi|y,Z,\vartheta) p(\omega|y,Z,\vartheta,\varphi) p(\delta|y,Z,\vartheta,\varphi,\omega)$$

where

$$p(\vartheta|y,Z) \propto p(\vartheta) s_\vartheta^{-\frac{n-1}{2}},$$

$$p(\varphi|y,Z,\vartheta) = p(\varphi|\vartheta),$$

$$p(\omega|y,Z,\vartheta,\varphi) = p_G(\omega | \frac{n-1}{2}, \frac{s_\vartheta}{2 (1-\varphi)}),$$

$$p(\delta|y,Z,\vartheta,\varphi,\omega) = p_N^1 (\delta | d_\vartheta, \frac{1 + (n-1)\varphi}{n\omega}).$$

First, let us notice that the conditional posterior of φ given ϑ is identical to the conditional prior of φ (given ϑ); we can learn from the data about φ only through prior dependence between φ and ϑ:

$$p(\varphi|y,\ Z) = \int_{\Theta} p(\varphi|\vartheta)\ p(\vartheta|y,\ Z)\ d\vartheta.$$

In the case of prior independence, i.e. if $p(\varphi|\vartheta) = p(\varphi)$, we also have posterior independence between φ and ϑ, and we cannot learn about φ from the data at all. Let us also notice that the marginal posterior of ϑ is the same as in the case of uncorrelated errors, so the form of correlation assumed here has no influence on the estimation of ϑ, provided, of course, that we use the same (marginal) prior $p(\vartheta)$ in both cases.

The presence of unknown $\varphi\epsilon$ [0,1) affects only the posterior densities of the precision parameter ω and the constant term δ.
In the case of uncorrelated disturbances we have $p(\omega|y,\ Z,\ \vartheta) =$ $= p_G(\omega|\frac{n-1}{2},\ \frac{s_\vartheta}{2\ (1-\varphi)})$ and thus, for example, $E(\omega|y,\ Z,\ \vartheta) = \frac{n-1}{s_\vartheta}$.

In the case of equicorrelation disturbances we have

$$p(\omega|y,\ Z,\ \vartheta) = {}_0\!\int^1 p_G(\omega|\frac{n-1}{2},\ \frac{s_\vartheta}{2\ (1-\varphi)})\ p(\varphi|\vartheta)\ d\varphi,$$

$$E(\omega|y,\ Z,\ \vartheta) = \frac{n-1}{s_\vartheta}\ [1 - E(\varphi|\vartheta)].$$

The assumption of equicorrelation influences the posterior of δ, but not its mean; we have

$$p(\delta|y,\ Z,\ \vartheta,\ \varphi) = {}_0\!\int^\infty p(\delta|y,\ Z,\ \vartheta,\ \varphi,\ \omega)\ p(\omega|y,\ Z,\ \vartheta,\ \varphi)\ d\omega =$$

$$= p_S^1\ (\delta|n-1,\ d_\vartheta,\ n\ \frac{n-1}{s_\vartheta}\cdot\ \frac{1-\varphi}{1+(n-1)\varphi}),$$

$$p(\delta|y,\ Z,\ \vartheta) = {}_0\!\int^1 p_S^1\ (\delta|n-1,\ d_\vartheta,\ n\ \frac{n-1}{s_\vartheta}\cdot\ \frac{1-\varphi}{1+(n-1)\varphi})\ p(\varphi|\vartheta)\ d\varphi,$$

$$E(\delta|y,\ Z,\ \vartheta) = d_\vartheta = E(\delta|y,\ Z,\ \vartheta,\ \varphi),$$

where the conditional posterior mean d_ϑ is the same as in the case of uncorrelated errors. The higher moments of δ, however, are different and can be infinite. For example:

$$E(\delta^2|y,\ Z,\ \vartheta) = {}_0\int^1 \left[\frac{s_\vartheta}{n(n-3)} \cdot \frac{1+(n-1)\varphi}{1-\varphi} + d_\vartheta^2\right] p(\varphi|\vartheta)\ d\varphi =$$

$$d_\vartheta^2 + \frac{s_\vartheta}{n(n-3)} \left[{}_0\int^\infty \frac{p(\varphi|\vartheta)}{1-\varphi}\ d\varphi + (n-1)\ {}_0\int^1 \frac{\varphi}{1-\varphi}\ p(\varphi|\vartheta)\ d\varphi\right],$$

and for $p(\varphi|\vartheta) = 1$ we have

$${}_0\int^1 \frac{d\varphi}{1-\varphi} = +\infty, \qquad {}_0\int^1 \frac{\varphi}{1-\varphi}\ d\varphi = +\infty.$$

Usually, however, the precision ω and the constant term δ are treated as nuisance parameters. Whenever the model with equicorrelated observations seems appropriate and the elements of ϑ are the only parameters of interest, we can rely on the posterior results under assumption of uncorrelated observations. The same conclusions were reached in Osiewalski (1987), but only for the case of a linear model with diffuse prior, that is for

$$\tau = k-1, \qquad \Theta = R^{k-1}, \qquad h_\vartheta = Z\vartheta, \qquad p(\vartheta,\varphi)\ \alpha\ p(\varphi).$$

3. Concluding remarks

For a (nonlinear) regression model with a constant term δ, with equicorrelated errors and with diffuse priors of δ and of the error precision ω, it is shown that the marginal posterior of ϑ, the vector of the (nonlinear) regression parameters other than δ, is exactly the same as in the case of uncorrelated errors.

Since normality of errors and nonrandomness of explanatory variables (Z matrix) were assumed, let us note that both these assumptions can be relaxed. We could consider random Z and under the assumption of a Baysesian cut our results would remain wholly valid; for the definition of a Bayesian cut see Florens and Mouchart (1985). If we assumed, more generally, elliptical errors instead of normal ones, then only the posterior of the error precision ω would be affected by this change, but the (marginal) posterior of the rest of the parameters would remain unchanged; see Osiewalski (1988). On the other hand, the presence of the constant term (δ) in the model and the form of the priors of δ and ω (diffuse!) seem to be the crucial assumptions for obtaining the results presented in the paper.

References

Balestra, P., 1970, On the efficiency of ordinary least-squares in regression models, Journal of the American Statistical Association 65, 1330-1337.

Florens, J.-P. and M. Mouchart, 1985, Conditioning in dynamic models, Journal of Time Series Analysis 6, 15-34.

Graybill, F.A., 1969, Introduction to Matrices with Applications in Statistics (Wadsworth, Belmont).

Halperin, M., 1951, Normal regression theory in the presence of intra-class correlation, Annals of Mathematical Statistics 22, 573-579.

McElroy, F.W., 1967, A necessary and sufficient condition that ordinary least-squares estimators be best linear unbiased, Journal of the American Statistical Association 62, 1302-1304.

Osiewalski, J., 1987, Regresja liniowa z jednakowo skorelowanymi skladnikami losowymi - ujecie bayesowskie (Linear regression with equicorrelated errors - Bayesian approach), Przeglad Statystyczny 34, 233-242.

Osiewalski, J., 1988, A note on Bayesian estimation and prediction in nonlinear regressions with elliptical errors, mimeo.

Zellner, A., 1971, An Introduction to Bayesian Inference in Econometrics (Wiley, New York).

Discussion Paper Series, CentER, Tilburg University, The Netherlands:

No.	Author(s)	Title
8801	Th. van de Klundert and F. van der Ploeg	Fiscal Policy and Finite Lives in Interdependent Economies with Real and Nominal Wage Rigidity
8802	J.R. Magnus and B. Pesaran	The Bias of Forecasts from a First-order Autoregression
8803	A.A. Weber	The Credibility of Monetary Policies, Policymakers' Reputation and the EMS-Hypothesis: Empirical Evidence from 13 Countries
8804	F. van der Ploeg and A.J. de Zeeuw	Perfect Equilibrium in a Model of Competitive Arms Accumulation
8805	M.F.J. Steel	Seemingly Unrelated Regression Equation Systems under Diffuse Stochastic Prior Information: A Recursive Analytical Approach
8806	Th. Ten Raa and E.N. Wolff	Secondary Products and the Measurement of Productivity Growth
8807	F. van der Ploeg	Monetary and Fiscal Policy in Interdependent Economies with Capital Accumulation, Death and Population Growth
8901	Th. Ten Raa and P. Kop Jansen	The Choice of Model in the Construction of Input-Output Coefficient Matrices
8902	Th. Nijman and F. Palm	Generalized Least Squares Estimation of Linear Models Containing Rational Future Expectations
8903	A. van Soest, I. Woittiez, A. Kapteyn	Labour Supply, Income taxes and Hours Restrictions in The Netherlands
8904	F. van der Ploeg	Capital Accumulation, Inflation and Long-Run Conflict in International Objectives
8905	Th. van de Klundert and A. van Schaik	Unemployment Persistence and Loss of Productive Capacity: a Keynesian approach
8906	A.J. Markink and F. van der Ploeg	Dynamic Policy of Linear Models with Rational Expectations of Future Events: A Computer Package
8907	J. Osiewalski	Posterior Densities for Nonlinear Regression with Equicorrelated Errors

P.O. BOX 90153, 5000 LE TILBURG, THE NETHERLANDS