Jiří Anděl

On the Bayes approach in general multiple autoregressive series

Aplikace matematiky, Vol. 18 (1973), No. 1, 18–29

Persistent URL: http://dml.cz/dmlcz/103444

SVAZEK 18 (1973) APLIKACE MATEMATIKY ČÍSLO 1

ON THE BAYES APPROACH IN GENERAL MULTIPLE AUTOREGRESSIVE SERIES

Jiří Anděl

(Received September 6, 1971)

The p-dimensional autoregressive series are treated in this paper from the Bayes viewpoint. The point estimates for the autoregressive parameters have the same form (apart from a small modification) as the maximum likelihood estimates. The posterior distributions are derived and used for testing hypotheses. The theory is applied to the model with exogenous and endogenous variables.

1. INTRODUCTION

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be given p-dimensional random vectors with zero expectations and finite covariance matrices. Let $\mathbf{Y}_{n+1}, \ldots, \mathbf{Y}_N$ be uncorrelated p-dimensional random vectors such that

$$E\mathbf{Y}_t = \mathbf{0}\,, \quad E\mathbf{Y}_t\mathbf{Y}_t' = \mathbf{I}\,, \quad E\mathbf{X}_s\mathbf{Y}_t' = \mathbf{0}$$

for $1 \leqq s \leqq n < t \leqq N$, where $\mathbf{I}$ is the unit matrix. Define random vectors $\mathbf{X}_{n+1}, \ldots$ $\ldots, \mathbf{X}_N$ by the formula

$$\sum_{j=0}^{n} \mathbf{A}_j \mathbf{X}_{t-j} = \mathbf{Y}_t\,, \quad n < t \leqq N\,, \tag{1}$$

where $\mathbf{A}_0, \ldots, \mathbf{A}_n$ are matrices of the type (p, p) such that $|\mathbf{A}_0| \neq 0$. Then $\mathbf{X}_1, \ldots, \mathbf{X}_N$ is called the p-dimensional autoregressive series. If $\mathbf{A}_n \neq 0$, we say that the series is of the order n.

The Bayes theory of one-dimensional stationary autoregressive series is given in the Champernowne's paper [3]. A conditional Bayes approach for general one-dimensional autoregressive series is suggested in [5]. The author of the present paper generalized this method to the p-dimensional model (1) under the assumption that $\mathbf{A}_0$ is diagonal [1]. This restriction is removed here. The generalization to the non-diagonal regular matrix $\mathbf{A}_0$ is not trivial and the derived statistical tests are asymptotic.

As usual, it shall be supposed that the vectors $\mathbf{X}_1, \ldots, \mathbf{X}_N$ have simultaneous normal distribution.

Denote by E_k the k-dimensional Euclidean space. If a matrix $\mathbf{A}$ is positive definite (or positive semidefinite), we write $\mathbf{A} > 0$ (or $\mathbf{A} \geqq 0$). The symbol $\operatorname{Tr} \mathbf{A}$ means the trace of $\mathbf{A}$, $|\mathbf{A}|$ denotes the determinant of $\mathbf{A}$ and $\mathbf{A}'$ is the transpose matrix to $\mathbf{A}$.

Lemma 1. *Let $\Omega = \{x_{11}, \ldots, x_{1p}, x_{22}, \ldots, x_{2p}, \ldots, x_{pp}\}$ be such a set in $E_{p(p+1)/2}$ that $\mathbf{X} = \|x_{ij}\|_{i,j=1}^{p} > 0$, where $x_{ij} = x_{ji}$ for $i > j$. Let $\mathbf{D} > 0$ be a matrix of the type (p, p). Then a constant $c_{pm} > 0$ exists for any natural $m > p$ such that*

$$\begin{aligned} f_m(x_{11}, \ldots, x_{pp}) &= c_{pm}|\mathbf{D}|^{(m-1)/2}\,|\mathbf{X}|^{(m-p-2)/2} \exp\{-\tfrac{1}{2}\operatorname{Tr}\mathbf{DX}\} \quad &&\text{on} \quad \Omega\,, \\ &= 0 &&\text{on} \quad E_{p(p+1)/2} - \Omega \end{aligned}$$

is the probability density.

Proof. See Cramér [4], § 29.5.

Lemma 2. *Denote by $\mathbf{I}$ the unit matrix of the type (p, p). Let $\mathbf{A}$, $\mathbf{B}$ be matrices of the type (p, p) such that $\mathbf{A} \geqq 0$, $\mathbf{B} \geqq 0$.*

Then

(a) $|\mathbf{I} + \mathbf{A}| \geqq 1 + \operatorname{Tr} \mathbf{A}$,

(b) $|\mathbf{I} + \mathbf{A}| \geqq 1 + |\mathbf{A}|$,

(c) $|\mathbf{I} + \mathbf{A}| \geqq 1 + \operatorname{Tr} \mathbf{A} + |\mathbf{A}|$ if $p \geqq 2$,

(d) $|\mathbf{A} + \mathbf{B}| \geqq |\mathbf{A}| + |\mathbf{B}|$.

Proof. Denote $\lambda_1, \ldots, \lambda_p$ all the roots of $\mathbf{A}$. Then $|\mathbf{A} - \lambda\mathbf{I}| = (\lambda_1 - \lambda) \ldots (\lambda_p - \lambda)$. Putting $\lambda = -1$ we get

$$|\mathbf{A} + \mathbf{I}| = (1 + \lambda_1) \ldots (1 + \lambda_p)\,. \tag{2}$$

If $\mathbf{A} \geqq 0$, then $\lambda_1 \geqq 0, \ldots, \lambda_p \geqq 0$. Formula (2) implies

$$|\mathbf{I} + \mathbf{A}| \geqq 1 + \lambda_1 + \lambda_2 + \ldots + \lambda_p = 1 + \operatorname{Tr} \mathbf{A}\,,$$

$$|\mathbf{I} + \mathbf{A}| \geqq 1 + \lambda_1\lambda_2 \ldots \lambda_p = 1 + |\mathbf{A}|$$

and for $p \geqq 2$

$$|\mathbf{I} + \mathbf{A}| \geqq 1 + \lambda_1 + \lambda_2 + \ldots + \lambda_p + \lambda_1\lambda_2 \ldots \lambda_p = 1 + \operatorname{Tr} \mathbf{A} + |\mathbf{A}|\,.$$

If $|\mathbf{A}| = |\mathbf{B}| = 0$, then (d) is obvious. If $\mathbf{A} > 0$, then

$$|\mathbf{A} + \mathbf{B}| = |\mathbf{A}^{1/2}(\mathbf{I} + \mathbf{A}^{-1/2}\mathbf{B}\mathbf{A}^{-1/2})\,\mathbf{A}^{1/2}|$$

and (d) follows from (b).

Lemma 3. *Consider random variables* V_{ij} $(1 \leqq i \leqq p, 1 \leqq j \leqq k)$ *with the simultaneous density*

$$t_{p,k,m}(\mathbf{V}) = c_{p,k,m} \left| \mathbf{I} + \frac{1}{m} \mathbf{V}\mathbf{V}' \right|^{-m/2}, \quad m \geqq 2,$$

where $c_{p,k,m}$ *is a constant and*

$$\mathbf{V} = \|v_{ij}\|_{i=1}^{p}\,{}_{j=1}^{k}.$$

Then for $m \to \infty$

$$t_{p,k,m}(\mathbf{V}) \to (2\pi)^{-pk/2} \exp\left\{-\tfrac{1}{2} \sum_{i=1}^{p} \sum_{j=1}^{k} v_{ij}^2\right\},$$

i.e., variables V_{ij} *have an asymptotically normal distribution with vanishing expectation and the unit covariance matrix.*

Proof. Obviously $|\mathbf{I} + (1/m)\,\mathbf{V}\mathbf{V}'| > 0$. Further, according to Lemma 2(a) we have for $m \geqq 2$

$$\int_{E_{pk}} \left| \mathbf{I} + \frac{1}{m} \mathbf{V}\mathbf{V}' \right|^{-m/2} \mathrm{d}v_{11} \dots \mathrm{d}v_{pk} \leqq \int_{E_{pk}} \left(1 + \frac{1}{m} \sum_{i=1}^{p} \sum_{j=1}^{k} v_{ij}^2\right)^{-m/2} \mathrm{d}v_{11} \dots \mathrm{d}v_{pk} < \infty.$$

Thus a constant $c_{p,k,m} > 0$ exists such that $t_{p,k,m}(\mathbf{V})$ is the probability density.

If $m \to \infty$, then

$$\left(\mathbf{I} + \frac{1}{m} \mathbf{V}\mathbf{V}'\right)^{-m/2} \to \exp\left\{-\tfrac{1}{2}\mathbf{V}\mathbf{V}'\right\}$$

and

$$\left| \mathbf{I} + \frac{1}{m} \mathbf{V}\mathbf{V}' \right|^{-m/2} \to \left|\exp\left\{\tfrac{1}{2}\mathbf{V}\mathbf{V}'\right\}\right|^{-1}.$$

Denote by $\lambda_1, \dots, \lambda_p$ all the roots of $\frac{1}{2}\mathbf{V}\mathbf{V}'$. The matrix $\exp\{\frac{1}{2}\mathbf{V}\mathbf{V}'\}$ has the roots $\exp\{\lambda_1\}, \dots, \exp\{\lambda_p\}$. (From $\frac{1}{2}\mathbf{V}\mathbf{V}' \geqq 0$ it follows that $\lambda_1 \geqq 0, \dots, \lambda_p \geqq 0$.) The product of the roots gives the value of the determinant.

Therefore, we obtain

$$\left|\exp\left\{-\tfrac{1}{2}\mathbf{V}\mathbf{V}'\right\}\right|^{-1} = \exp\{-\lambda_1 - \dots - \lambda_p\} = \exp\left\{-\operatorname{Tr} \tfrac{1}{2}\mathbf{V}\mathbf{V}'\right\} =$$

$$= \exp\left\{-\tfrac{1}{2} \sum_{i=1}^{p} \sum_{j=1}^{k} v_{ij}^2\right\}.$$

The rest of the proof is clear.

3. THE POINT ESTIMATES

Rewrite the model (1) in the form

(3) $$\mathbf{X}_t = \sum_{j=1}^{n} \mathbf{U}_j \mathbf{X}_{t-j} + \mathbf{A}_0^{-1} \mathbf{Y}_t, \quad n < t \leqq N,$$

where

$$U_j = -A_0^{-1} A_j, \quad 1 \leqq j \leqq n. \tag{4}$$

Lemma 4. *Let the random vectors* $X_1, \ldots, X_n, Y_{n+1}, \ldots, Y_N$ *have a simultaneous normal distribution. Given* $X_1 = x_1, \ldots, X_n = x_n$, *then the conditional density of the random vectors* $X_{n+1}, \ldots, X_N$ *is*

$$f(x_{n+1}, \ldots, x_N \mid x_1, \ldots, x_n) = (2\pi)^{-(N-n)p/2} |G|^{(N-n)/2} \,. \tag{5}$$
$$. \exp\{-\tfrac{1}{2} \sum_{t=n+1}^{N} (x_t - \sum_{j=1}^{n} U_j x_{t-j})' G(x_t - \sum_{j=1}^{n} U_j x_{t-j})\},$$

where $G = A_0' A_0$.

Proof. One starts with the simultaneous density of $X_1, \ldots, X_n, Y_{n+1}, \ldots, Y_N$. After the substitution (3) the density of $X_1, \ldots, X_N$ is obtained and then the evaluation of the conditional density gives the formula (5).

As it is well known, the density f depends on A_0 only through $G = A_0' A_0$. The estimate of A_0 cannot be determined uniquely from an estimate of G without an additional knowledge of the structure of A_0. But, fortunately, it is known in some applications that A_0 is an upper (or lower) triangular matrix and this enables us to construct estimates for A_0.

Introduce the matrix

$$U = \|U_1, \ldots, U_n\| \,.$$

We shall suppose that the elements of the matrices G and U are random variables which were realized before constructing the series $X_1, \ldots, X_N$. Let the prior density of the elements of G, U be proportional to $|G|^{-1/2}$ if $G > 0$ and zero otherwise, independently of $X_1, \ldots, X_n$.

The chosen prior density sometimes is called "vague". The following theorems can be considered as mathematical consequences of our assumption, but there are some reasons leading to our prior density. This density is a straightforward generalization of that used in the one-dimensional case (see [5]) and in the special p-dimensional case (see [1]). Furthermore, the point estimates of the autoregressive parameters have the usual form as those derived by the maximum likelihood method (see [6], p. 40 and p. 70).

Given $X_1 = x_1, \ldots, X_N = x_N$, then in view of the Bayes theorem the posterior density of G and U is

$$g(G, U \mid x) = c|G|^{(N-n-1)/2} \,.$$
$$. \exp\{-\tfrac{1}{2} \sum_{t=n+1}^{N} (x_t - \sum_{j=1}^{n} U_j x_{t-j})' G(x_t - \sum_{j=1}^{n} U_j x_{t-j})\}$$

for $G > 0$, where x denotes the condition $X_1 = x_1, \ldots, X_N = x_N$ and c is a constant.

Introduce vectors

$$^\circ\mathbf{x}_t = \begin{Vmatrix} \mathbf{x}_{t-1} \\ \vdots \\ \mathbf{x}_{t-n} \end{Vmatrix}, \quad n < t \leqq N$$

and matrices

$$\mathbf{C}_0 = \sum_{t=n+1}^{N} \mathbf{x}_t \mathbf{x}'_t, \quad \mathbf{C} = \sum_{t=n+1}^{N} \mathbf{x}_t {}^\circ\mathbf{x}'_t,$$

$$\mathbf{S} = \sum_{t=n+1}^{N} {}^\circ\mathbf{x}_t {}^\circ\mathbf{x}'_t, \quad \mathbf{S}_0 = \begin{Vmatrix} \mathbf{C}_0 & \mathbf{C} \\ \mathbf{C}' & \mathbf{S} \end{Vmatrix}.$$

Theorem 5. *Suppose* $\mathbf{S}_0 > 0$. *Then*

$$g(\mathbf{G}, \mathbf{U} \mid \mathbf{x}) = c|\mathbf{G}|^{(N-n-1)/2} \exp\{-\tfrac{1}{2} \operatorname{Tr} \mathbf{DG}\} \tag{6}$$

for $\mathbf{G} > 0$, *where*

$$\mathbf{D} = (\mathbf{U} - \mathbf{U}^*)\, \mathbf{S}(\mathbf{U} - \mathbf{U}^*)' + \mathbf{C}_0 - \mathbf{CS}^{-1}\mathbf{C}',$$

$$\mathbf{U}^* = \mathbf{CS}^{-1}. \tag{7}$$

The modus of the posterior distribution is $\mathbf{G}^*$, $\mathbf{U}^*$, *where*

$$\mathbf{G}^* = (N - n - 1)(\mathbf{C}_0 - \mathbf{CS}^{-1}\mathbf{C}')^{-1}. \tag{8}$$

Proof. Obviously

$$g(\mathbf{G}, \mathbf{U} \mid \mathbf{x}) = c|\mathbf{G}|^{(N-n-1)/2} \exp\{-\tfrac{1}{2} \sum_{t=n+1}^{N} (\mathbf{x}_t - \mathbf{U}{}^\circ\mathbf{x}_t)'\, \mathbf{G}(\mathbf{x}_t - \mathbf{U}{}^\circ\mathbf{x}_t)\} \tag{9}$$

for $\mathbf{G} > 0$. We have

$$\sum_{t=n+1}^{N} \operatorname{Tr} \mathbf{U}'\mathbf{GU}^* {}^\circ\mathbf{x}_t {}^\circ\mathbf{x}'_t - \sum_{t=n+1}^{N} \operatorname{Tr} \mathbf{U}'\mathbf{Gx}_t {}^\circ\mathbf{x}'_t =$$

$$= \operatorname{Tr} \mathbf{U}'\mathbf{GU}^*\mathbf{S} - \operatorname{Tr} \mathbf{U}'\mathbf{GC} = \operatorname{Tr} \mathbf{U}'\mathbf{G}(\mathbf{U}^*\mathbf{S} - \mathbf{C}) = 0,$$

and thus

$$\sum_{t=n+1}^{N} {}^\circ\mathbf{x}'_t \mathbf{U}'\mathbf{GU}^* {}^\circ\mathbf{x}_t - \sum_{t=n+1}^{N} {}^\circ\mathbf{x}'_t \mathbf{U}'\mathbf{Gx}_t = 0.$$

Using this fact we obtain after some computation

$$\sum_{t=n+1}^{N} (\mathbf{x}_t - \mathbf{Ux}_t)'\, \mathbf{G}(\mathbf{x}_t - \mathbf{Ux}_t) = \operatorname{Tr} \mathbf{GD} = \operatorname{Tr} \mathbf{DG}.$$

If $\mathbf{S}_0 > 0$ then $\mathbf{S} > 0$ and it follows immediately that $\mathbf{U} = \mathbf{U}^*$ maximizes g for any $\mathbf{G} > 0$.

Define the matrix

$$\boldsymbol{L} = \begin{Vmatrix} \boldsymbol{I} & -\boldsymbol{C}\boldsymbol{S}^{-1} \\ \boldsymbol{0} & \boldsymbol{I} \end{Vmatrix}.$$

The matrix $\boldsymbol{L}$ is regular, $\boldsymbol{S}_0 > 0$, and therefore

$$\boldsymbol{L}\boldsymbol{S}_0\boldsymbol{L}' = \begin{Vmatrix} \boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}' & * \\ * & \boldsymbol{0} \end{Vmatrix} > 0,$$

where symbol $*$ denotes the blocks which are not interesting for our purpose. Thus we proved $\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}' > 0$.

It remains to prove that $\boldsymbol{G}^*$ maximizes $g(\boldsymbol{G}, \boldsymbol{U}^* \mid \boldsymbol{x})$. We use Watson's method (see Rao [7], Chap. 8 a 5). Denote by $\lambda_1, \ldots, \lambda_p$ all the roots of the equation

$$|(N - n - 1)^{-1}\boldsymbol{G}(\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}') - \lambda\boldsymbol{I}| = 0.$$

Then

$$|(N - n - 1)^{-1}\boldsymbol{G}(\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')| = \lambda_1\lambda_2\ldots\lambda_p,$$

$$\operatorname{Tr}\boldsymbol{G}(\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}') = (N - n - 1)(\lambda_1 + \ldots + \lambda_p).$$

From $\boldsymbol{G} > 0$, $\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}' > 0$ it follows that $\lambda_1 > 0, \ldots, \lambda_p > 0$. This can be proved directly or as a corollary of Exercise 1.9 in Rao's book [7], Chap. 1. Consequently

$$2[\log g(\boldsymbol{G}^*, \boldsymbol{U}^* \mid \boldsymbol{x}) - \log g(\boldsymbol{G}, \boldsymbol{U}^* \mid \boldsymbol{x})] =$$

$$= -(N - n - 1)\log|(N - n - 1)^{-1}\boldsymbol{G}(\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')| - (N - n - 1)p +$$

$$+ \operatorname{Tr}\boldsymbol{G}(\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}') = (N - n - 1)[-\log(\lambda_1\lambda_2\ldots\lambda_p) - p + \lambda_1 \ldots + \lambda_p] \geqq 0.$$

The last inequality follows from

$$-\log x - 1 + x \geqq 0 \quad \text{for} \quad x > 0.$$

We remark that the assumption $\boldsymbol{S}_0 > 0$ is not very restrictive and is often fulfilled in applications.

The values of $\boldsymbol{G}^*$ and $\boldsymbol{U}^*$ can be used as the point estimates of the matrices $\boldsymbol{G}$ and $\boldsymbol{U}$.

4. MARGINAL POSTERIOR DISTRIBUTIONS AND TESTING HYPOTHESES

Theorem 6. *Suppose* $\boldsymbol{S}_0 > 0$ *and let* $g(\boldsymbol{G}, \boldsymbol{U} \mid \boldsymbol{x})$ *be defined by formula* (6). *If* $N \geqq (n + 1)p$, *then the posterior distribution of* $\boldsymbol{G}$ *is Wishart distribution*

$$W_p(N - n - np + p, \boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}').$$

Define the matrix

$$\boldsymbol{V} = (N - n + p)^{1/2}(\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')^{-1/2}(\boldsymbol{U} - \boldsymbol{U}^*)\boldsymbol{S}^{1/2}. \tag{10}$$

If $N - n + p \geqq 2$, then $\boldsymbol{V}$ has the posterior distribution given by the density $t_{p,np,N-n+p}(\boldsymbol{V})$ mentioned in Lemma 3.

Proof. The marginal posterior density $h(\boldsymbol{G} \mid \boldsymbol{x})$ of $\boldsymbol{G}$ is evaluated by the formula

$$h(\boldsymbol{G} \mid \boldsymbol{x}) = \int_{E_{np^2}} g(\boldsymbol{G}, \boldsymbol{U} \mid \boldsymbol{x}) \, \mathrm{d}u_{11} \ldots \mathrm{d}u_{p,np} .$$

In order to evaluate the integral we use the substitution

$$\boldsymbol{Q} = \boldsymbol{G}^{1/2}(\boldsymbol{U} - \boldsymbol{U}^*) ,$$

the Jacobian of which is (in the absolute value) $|\boldsymbol{G}|^{-np/2}$. Thus we obtain

$$h(\boldsymbol{G} \mid \boldsymbol{x}) = c_1 |\boldsymbol{G}|^{(N-n-np-1)/2} \exp \{-\tfrac{1}{2} \operatorname{Tr} \boldsymbol{G}(\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')\}$$

for $\boldsymbol{G} > 0$. If $N \geqq (n + 1) p$, $h(\boldsymbol{G} \mid \boldsymbol{x})$ is the density of the Wishart distribution $W_p(N - n - np + p, \boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')$ (see Rao [7], Exercise 11.6, Chap. 8, or Anderson [2], Chap. 7.2).

We obtain the marginal posterior density $q(\boldsymbol{U} \mid \boldsymbol{x})$ of the matrix $\boldsymbol{U}$ using Lemma 1. We get

$$(11) \quad q(\boldsymbol{U} \mid \boldsymbol{x}) = \int_{\Omega} g(\boldsymbol{G}, \boldsymbol{U} \mid \boldsymbol{x}) \, \mathrm{d}g_{11} \ldots \mathrm{d}g_{pp} = c_2 |\boldsymbol{D}|^{-(N-n+p)/2} = \\ = c_2 |\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}' + (\boldsymbol{U} - \boldsymbol{U}^*) \, \boldsymbol{S}(\boldsymbol{U} - \boldsymbol{U}^*)'|^{-(N-n+p)/2} .$$

Let us use the linear substitution (10), the Jacobian of which is constant. Then we get the density

$$q_1(\boldsymbol{V} \mid \boldsymbol{x}) = c_3 |\boldsymbol{I} + (N - n + p)^{-1} \boldsymbol{V}\boldsymbol{V}'|^{-(N-n+p)/2} ,$$

which is the same as $t_{p,np,N-n+p}(\boldsymbol{V})$. The constant c_3 depends, of course, on $\boldsymbol{x}$, but it is fixed in the Bayes approach.

Theorem 6 enables us to test some statistical hypotheses. The simplest case is the testing hypothesis $\boldsymbol{U} = \boldsymbol{U}^0$, where $\boldsymbol{U}^0$ is a given matrix. The test statistic will be the sum of the squares of the elements of the matrix

$$\boldsymbol{V}_0 = (N - n + p)^{1/2} \, (\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')^{-1/2} \, (\boldsymbol{U}^0 - \boldsymbol{U}^*) \, \boldsymbol{S}^{1/2} .$$

This sum equals to

$$\operatorname{Tr} \boldsymbol{V}_0\boldsymbol{V}_0' = (N - n + p) \operatorname{Tr} [(\boldsymbol{U}^0 - \boldsymbol{U}^*) \, \boldsymbol{S}(\boldsymbol{U}^0 - \boldsymbol{U}^*) \, (\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')^{-1}] .$$

The last formula is more appropriate for the evaluation.

If the hypothesis $\boldsymbol{U} = \boldsymbol{U}^0$ is true, then $\boldsymbol{V}_0$ has the distribution with the density $t_{p,np,N-n+p}(\boldsymbol{V})$. Unfortunately, the exact distribution of the statistic $\operatorname{Tr} \boldsymbol{V}_0\boldsymbol{V}_0'$ is not known. Some work on this problem would be useful. We give only the asymptotic solution of the problem.

In view of Lemma 3 the elements of $\boldsymbol{V}_0$ have for $N \to \infty$ an asymptotically normal distribution with vanishing expectation and with the unit covariance matrix. Thus the sum of the squares of the elements of $\boldsymbol{V}_0$ has an asymptotically χ^2-distribution with np^2 degrees of freedom. If the value $\operatorname{Tr} \boldsymbol{V}_0 \boldsymbol{V}_0'$ exceeds the α per cent critical value of $\chi^2(np^2)$, we reject the hypothesis $\boldsymbol{U} = \boldsymbol{U}^0$ on the level α per cent.

If we put particularly $\boldsymbol{U}^0 = \boldsymbol{0}$ then it is tested that random vectors $\boldsymbol{X}_t$ are mutually independent. (The covariance matrix $E\boldsymbol{X}_t\boldsymbol{X}_t'$ is not specified.)

We defined $\boldsymbol{U} = \|\boldsymbol{U}_1, \ldots, \boldsymbol{U}_n\|$. We often want to test the hypothesis $\boldsymbol{U}_n = \boldsymbol{U}_n^0$, where $\boldsymbol{U}_n^0$ is a given matrix. This hypothesis does not specify the values of the matrices $\boldsymbol{U}_1, \ldots, \boldsymbol{U}_{n-1}$. Particularly, the hypothesis $\boldsymbol{U}_n = \boldsymbol{0}$ means that the order of the autoregressive model is smaller than n. In order to construct such a test we derive the marginal posterior distribution corresponding to the matrix $\boldsymbol{U}_n$.

Denote

$$\boldsymbol{S}^{-1} = \boldsymbol{R} = \begin{Vmatrix} \boldsymbol{R}_{11} & \boldsymbol{R}_{12} \\ \boldsymbol{R}_{21} & \boldsymbol{R}_{22} \end{Vmatrix}.$$

where the block $\boldsymbol{R}_{22}$ is of the type (p, p). $\boldsymbol{S} > 0$ implies $\boldsymbol{R} > 0$. Thus we have $\boldsymbol{R}_{11} > 0$, $\boldsymbol{R}_{22} > 0$ and $\boldsymbol{R}_{11} - \boldsymbol{R}_{12}\boldsymbol{R}_{22}^{-1}\boldsymbol{R}_{21} > 0$ (the last assertion can be proved in the same way as in the proof of Theorem 5). Introduce matrices

$$\begin{aligned} \boldsymbol{P}_1 &= (\boldsymbol{R}_{11} - \boldsymbol{R}_{12}\boldsymbol{R}_{22}^{-1}\boldsymbol{R}_{21})^{-1/2}, \\ \boldsymbol{P}_2 &= -\boldsymbol{R}_{22}^{-1}\boldsymbol{R}_{21}(\boldsymbol{R}_{11} - \boldsymbol{R}_{12}\boldsymbol{R}_{22}^{-1}\boldsymbol{R}_{21})^{-1/2}, \\ \boldsymbol{P}_3 &= \boldsymbol{R}_{22}^{-1/2}, \end{aligned}$$

$$\boldsymbol{P} = \begin{Vmatrix} \boldsymbol{P}_1 & \boldsymbol{0} \\ \boldsymbol{P}_2 & \boldsymbol{P}_3 \end{Vmatrix}.$$

We can verify easily that $\boldsymbol{S} = \boldsymbol{P}\boldsymbol{P}'$. Consider the density $q(\boldsymbol{U} \mid \boldsymbol{x})$ given in (11) and write $\boldsymbol{U}^* = \|\boldsymbol{U}_1^*, \ldots, \boldsymbol{U}_n^*\|$, where $\boldsymbol{U}_1^*, \ldots, \boldsymbol{U}_n^*$ are the blocks of the type (p, p). Obviously

$$q(\boldsymbol{U} \mid \boldsymbol{x}) = c_3 \mid \boldsymbol{I} + (\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')^{-1/2} (\boldsymbol{U} - \boldsymbol{U}^*) \boldsymbol{P}\boldsymbol{P}'(\boldsymbol{U} - \boldsymbol{U}^*)' \,. \\ . (\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')^{-1/2}|^{-(N-n+p)/2}$$

holds, where c_3 is a positive constant. Make the linear substitution

$$(\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')^{-1/2} (\boldsymbol{U} - \boldsymbol{U}^*) \boldsymbol{P} = \boldsymbol{W}, \tag{12}$$

the Jacobian of which is a constant. Write $\boldsymbol{W}$ in the form

$$\boldsymbol{W} = \|\boldsymbol{W}_1, \boldsymbol{W}_2\|,$$

where $\boldsymbol{W}_1$ is of the type $(p, np - p)$ and W_2 is of the type (p, p). Note that (12) implies

$$\boldsymbol{W}_2 = (\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')^{-1/2} (\boldsymbol{U}_n - \boldsymbol{U}_n^*) \boldsymbol{R}_{22}^{-1/2}.$$

The probability density of the matrix $\mathbf{W}$ is

$$\begin{aligned} q_2(\mathbf{W} \mid \mathbf{x}) &= c_4 |\mathbf{I} + \mathbf{W}\mathbf{W}'|^{-(N-n+p)/2} = \\ &= c_4 |\mathbf{I} + \mathbf{W}_1\mathbf{W}_1' + \mathbf{W}_2\mathbf{W}_2'|^{-(N-n+p)/2} = \\ &= c_4 |\mathbf{I} + \mathbf{W}_2\mathbf{W}_2'|^{-(N-n+p)/2} \, |\mathbf{I} + (\mathbf{I} + \mathbf{W}_2\mathbf{W}_2')^{-1/2} \, . \\ &\quad . \, \mathbf{W}_1\mathbf{W}_1'(\mathbf{I} + \mathbf{W}_2\mathbf{W}_2')^{-1/2}|^{-(N-n+p)/2} \, . \end{aligned}$$

In order to obtain the marginal distribution of the matrix $\mathbf{W}_2$, we make the substitution

$$(\mathbf{I} + \mathbf{W}_2\mathbf{W}_2')^{-1/2} \, \mathbf{W}_1 = \mathbf{V}_1 \, .$$

The Jacobian is

$$|\mathbf{I} + \mathbf{W}_2\mathbf{W}_2'|^{(np-p)/2} \, .$$

This leads to the marginal density

$$q_3(\mathbf{W}_2 \mid \mathbf{x}) = c_5 |\mathbf{I} + \mathbf{W}_2\mathbf{W}_2'|^{-(N-n-np+2p)/2}$$

and thus the posterior density of the matrix $\mathbf{V}_2 = (N - n - np + 2p)^{1/2} \, \mathbf{W}_2$ is given by the formula

$$q_4(\mathbf{V}_2 \mid \mathbf{x}) = c_6 |\mathbf{I} + (N - n - np + 2p)^{-1} \, \mathbf{V}_2\mathbf{V}_2'|^{-(N-n-np+2p)/2} \tag{13}$$

(c_4, c_5 and c_6 are constants). The result may be formulated as follows:

Theorem 7. *Suppose $\mathbf{S}_0 > 0$ and let the density $g(\mathbf{G}, \mathbf{U} \mid \mathbf{x})$ be given by the formula (6). Then the marginal posterior distribution of the matrix*

$$\mathbf{V}_2 = (N - n - np + 2p)^{1/2} \, (\mathbf{C}_0 - \mathbf{C}\mathbf{S}^{-1}\mathbf{C}')^{1/2} \, (\mathbf{U}_n - \mathbf{U}_n^*) \, \mathbf{R}_{22}^{-1/2}$$

has the density $q_4(\mathbf{V}_2 \mid \mathbf{x})$ given in the formula (13).

Thus the density of $\mathbf{V}_2$ is the same as $t_{p,p,N-n-np+2p}(\mathbf{V}_2)$. The testing of the hypothesis $\mathbf{U}_n = \mathbf{U}_n^0$ (where $\mathbf{U}_n^0$ is a given matrix) can be done quite similarly as in the previous case. The elements of the matrix $\mathbf{V}_2$ have an asymptotically normal distribution with vanishing means and the unit covariance matrix. Therefore, the sum of the squares of the elements of the matrix

$$\mathbf{V}_2^0 = (N - n - np + 2p)^{1/2} \, (\mathbf{C}_0 - \mathbf{C}\mathbf{S}^{-1}\mathbf{C}')^{-1/2} \, (\mathbf{U}_n^0 - \mathbf{U}_n^*) \, \mathbf{R}_{22}^{-1/2}$$

has an asymptotically χ^2-distribution with p^2 degrees of freedom. The sum is the same as

$$\operatorname{Tr} \mathbf{V}_2^0\mathbf{V}_2^{0\prime} = (N - n - np + 2p) \operatorname{Tr} \left[(\mathbf{U}_n^0 - \mathbf{U}_n^*) \, \mathbf{R}_{22}^{-1} (\mathbf{U}_n^0 - \mathbf{U}_n^*)' \, (\mathbf{C}_0 - \mathbf{C}\mathbf{S}^{-1}\mathbf{C}')^{-1} \right] .$$

If the value Tr $\mathbf{V}_2^0 \mathbf{V}_2^{0\prime}$ exceeds the α per cent critical value of the χ^2-distribution with p^2 degrees of freedom, we reject the hypothesis $\mathbf{U}_n = \mathbf{U}_n^0$ on the level α per cent. As we mentioned, the hypothesis $\mathbf{U}_n = \mathbf{0}$ occurs most often in the practice.

5. MODEL WITH EXOGENOUS AND ENDOGENOUS VARIABLES

The model (1) may be regarded as a pure autoregressive model. Sometimes a more general model is used.

Let $\mathbf{X}_1, \ldots, \mathbf{X}_n$ be (p-dimensional) random vectors with vanishing mean values and finite second moments. Let $\mathbf{Y}_{n+1}, \ldots, \mathbf{Y}_N$ be uncorrelated random vectors such that

$$E\mathbf{Y}_t = \mathbf{0}\,, \quad E\mathbf{Y}_t\mathbf{Y}_t' = \mathbf{I}\,, \quad E\mathbf{X}_s\mathbf{Y}_t' = \mathbf{0}$$

for $1 \leqq s \leqq n < t \leqq N$. Further let $\boldsymbol{\varphi}_{n-m+1}, \boldsymbol{\varphi}_{n-m+2}, \ldots, \boldsymbol{\varphi}_N$ be given (generally non-random) vectors, where m is a non-negative integer. Define random vectors $\mathbf{X}_{n+1}, \ldots, \mathbf{X}_N$ by the formula

$$\mathbf{X}_t = \sum_{j=1}^{n} \mathbf{U}_j \mathbf{X}_{t-j} + \sum_{j=0}^{m} \mathbf{V}_j \boldsymbol{\varphi}_{t-j} + \mathbf{A}_0^{-1}\mathbf{Y}_t\,, \quad n < t \leqq N\,, \tag{14}$$

where $\mathbf{U}_j$, $\mathbf{V}_j$ and $\mathbf{A}_0$ are matrices of the type (p, p) with real elements. The elements of the vectors $\mathbf{X}_t$ are called endogenous variables and those of $\boldsymbol{\varphi}_t$ exogenous ones.

The main problem is to find estimates for the matrices $\mathbf{U}_j$, $\mathbf{V}_j$, $\mathbf{A}_0$. This problem sometimes is complicated with respect to the fact that some elements of these matrices are dependent on the others or that they have given values. We shall not consider such conditions and we deal with the simplest case.

Let the random vectors $\mathbf{X}_1, \ldots, \mathbf{X}_n, \mathbf{Y}_{n+1}, \ldots, \mathbf{Y}_N$ have a simultaneous normal distribution. Then given $\mathbf{X}_1 = \mathbf{x}_1, \ldots, \mathbf{X}_n = \mathbf{x}_n$, the conditional density of $\mathbf{X}_{n+1}, \ldots$ $\ldots, \mathbf{X}_N$ is

$$f(\mathbf{x}_{n+1}, \ldots, \mathbf{x}_N \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) = (2\pi)^{-(N-n)p/2}\, |\mathbf{G}|^{(N-n)/2}\,.$$

$$.\exp\Big\{-\tfrac{1}{2} \sum_{t=n+1}^{N} \big(\mathbf{x}_t - \sum_{j=1}^{n} \mathbf{U}_j \mathbf{x}_{t-j} - \sum_{j=0}^{m} \mathbf{V}_j \boldsymbol{\varphi}_{t-j}\big)' \mathbf{G} \big(\mathbf{x}_t - \sum_{j=1}^{n} \mathbf{U}_j \mathbf{x}_{t-j} - \sum_{j=0}^{m} \mathbf{V}_j \boldsymbol{\varphi}_{t-j}\big)\Big\}\,,$$

where $\mathbf{G} = \mathbf{A}_0'\mathbf{A}_0$. Introduce a matrix $\mathbf{U} = \|\mathbf{U}_1, \ldots, \mathbf{U}_n, \mathbf{V}_0, \mathbf{V}_1, \ldots, \mathbf{V}_m\|$, vectors

$${}^{\circ}\mathbf{x}_t = \begin{Vmatrix} \mathbf{x}_{t-1} \\ \vdots \\ \mathbf{x}_{t-n} \end{Vmatrix}, \quad {}^{\circ}\boldsymbol{\varphi}_t = \begin{Vmatrix} \boldsymbol{\varphi}_t \\ \boldsymbol{\varphi}_{t-1} \\ \vdots \\ \boldsymbol{\varphi}_{t-m} \end{Vmatrix}, \quad \mathbf{z}_t = \begin{Vmatrix} {}^{\circ}\mathbf{x}_t \\ {}^{\circ}\boldsymbol{\varphi}_t \end{Vmatrix}, \quad n < t \leqq N\,,$$

and matrices

$$\mathbf{C}_0 = \sum_{t=n+1}^{N} \mathbf{x}_t \mathbf{x}_t'\,, \quad \mathbf{C} = \sum_{t=n+1}^{N} \mathbf{x}_t \mathbf{z}_t'\,, \quad \mathbf{S} = \sum_{t=n+1}^{N} \mathbf{z}_t \mathbf{z}_t'\,,$$

$$\mathbf{S}_0 = \begin{Vmatrix} \mathbf{C}_0 & \mathbf{C} \\ \mathbf{C}' & \mathbf{S} \end{Vmatrix}.$$

Using this notation we have

$$f(\boldsymbol{x}_{n+1}, \ldots, \boldsymbol{x}_N \mid \boldsymbol{x}_1, \ldots, \boldsymbol{x}_n) = (2\pi)^{-(N-n)p/2} |\boldsymbol{G}|^{(N-n)/2} \,.$$
$$. \exp\{-\tfrac{1}{2} \sum_{t=n+1}^{N} (\boldsymbol{x}_t - \boldsymbol{U}\boldsymbol{z}_t)' \boldsymbol{G}(\boldsymbol{x}_t - \boldsymbol{U}\boldsymbol{z}_t)\} \,,$$

which is analogous to the formula (5). But in our case the matrix $\boldsymbol{U}$ is of the type $(p, np + mp + p)$ and the vectors $\boldsymbol{z}_t$ are $(np + mp + p)$-dimensional. Taking into account these differences, we can proceed analogously as in the previous Section and construct the estimates by the Bayes method.

Theorem 8. *Suppose that the elements of* $\boldsymbol{G}$ *and* $\boldsymbol{U}$ *are random variables with the prior density proportional to* $|\boldsymbol{G}|^{-1/2}$ *for* $\boldsymbol{G} > 0$ *independently of* $\boldsymbol{X}_1, \ldots, \boldsymbol{X}_n$. *Let* $\boldsymbol{S}_0 > 0$. *Then under the assumption of normality the modus of the posterior density is* $\boldsymbol{U}^* = \boldsymbol{C}\boldsymbol{S}^{-1}$, $\boldsymbol{G}^* = (N - n - 1)(\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')^{-1}$. *If* $N \geqq np + mp + 2p$, *then the posterior distribution of* $\boldsymbol{G}$ *is the Wishart one* $W_p(N - n - np - mp, \boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')$. *The posterior density* $q(\boldsymbol{U} \mid \boldsymbol{x})$ *of the matrix* $\boldsymbol{U}$ *is given by the formula*

$$q(\boldsymbol{U} \mid \boldsymbol{x}) = c|\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}' + (\boldsymbol{U} - \boldsymbol{U}^*)\,\boldsymbol{S}(\boldsymbol{U} - \boldsymbol{U}^*)'|^{-(N-n+p)/2}$$

and the elements of the matrix

$$\boldsymbol{V} = (N - n + p)^{1/2} (\boldsymbol{C}_0 - \boldsymbol{C}\boldsymbol{S}^{-1}\boldsymbol{C}')^{-1/2} (\boldsymbol{U} - \boldsymbol{U}^*)\,\boldsymbol{S}^{1/2}$$

have for $N \to \infty$ *an asymptotically normal distribution with vanishing expectation and the unit covariance matrix.*

Proof is analogous to those of Theorem 5 and Theorem 6.

Theorem 8 gives the estimates and enables testing hypotheses in the same manner as in the case of the pure autoregressive model. Similarly it is possible to derive the asymptotic test of the hypothesis $\boldsymbol{V}_m = \boldsymbol{0}$ or that of $\boldsymbol{U}_n = \boldsymbol{0}$. The procedure is obvious and we do not present it here.

References

[1] *J. Anděl:* The Bayes approach in multiple autoregressive series. Aplikace matematiky 16 (1971), 220—228.
[2] *T. W. Anderson:* An introduction to multivariate statistical analysis. Wiley, New York 1958.
[3] *D. G. Champernowne:* Sampling theory applied to autoregressive sequences. J. Roy. Stat. Soc. ser. B, 10 (1948), 204—231.
[4] *H. Cramér:* Mathematical methods of statistics. Princeton Univ. Press 1946.
[5] *J. Hájek, J. Anděl:* Stacionárni procesy. (Lecture notes.) SPN 1969.
[6] *M. H. Quenouille:* The analysis of multiple time-series. Griffin, London 1957.
[7] *C. R. Rao:* Linear statistical inference and its applications. Wiley, New York 1965.

Souhrn

O BAYESOVĚ PŘÍSTUPU V OBECNÝCH MNOHOROZMĚRNÝCH AUTOREGRESNÍCH POSLOUPNOSTECH

JIŘÍ ANDĚL

Nechť $\boldsymbol{X}_1, \boldsymbol{X}_2, \ldots, \boldsymbol{X}_N$ je konečná část p-rozměrné normální autoregresní posloupnosti dané relací

$$\sum_{k=0}^{n} \boldsymbol{A}_k \boldsymbol{X}_{t-k} = \boldsymbol{Y}_t ,$$

kde $\boldsymbol{Y}_t$ jsou nekorelované náhodné vektory s nulovými středními hodnotami a s jednotkovými kovariančními maticemi. Za předpokladu, že $\boldsymbol{A}_0$ je regulární, se na základě podmíněné hustoty vektorů $\boldsymbol{X}_{n+1}, \ldots, \boldsymbol{X}_N$ při daných $\boldsymbol{X}_1 = \boldsymbol{x}_1, \ldots, \boldsymbol{X}_n = \boldsymbol{x}_n$ při obvyklé volbě apriorní hustoty odvozuje aposteriorní hustota autoregresních parametrů. Modus aposteriorní hustoty se volí za bodový odhad autoregresních parametrů. Jsou vypočteny i některé marginální aposteriorní hustoty a pomocí nich jsou odvozeny testy hypotéz o parametrech. Uvedené testy jsou asymptotické a mohou být použity zejména při testování hypotéz o řádu autoregrese. V závěru práce je tato teorie také aplikována na model generovaný vztahem

$$\boldsymbol{X}_t = \sum_{j=1}^{n} \boldsymbol{U}_j \boldsymbol{X}_{t-j} + \sum_{j=0}^{m} \boldsymbol{V}_j \varphi_{t-j} + \mathrm{A}_0^{-1} \boldsymbol{Y}_t ,$$

kde φ_t jsou tzv. exogenní proměnné.

Uvedený postup je zcela obecný. Není třeba předpokládat ani stacionaritu posloupnosti a kromě regularity $\boldsymbol{A}_0$ se nemusí klást žádné další požadavky na autoregresní matice $\boldsymbol{A}_0, \ldots, \boldsymbol{A}_n$.

Author's address: RNDr. *Jiří Anděl,* CSc., Matematicko-fyzikální fakulta Karlovy university, Sokolovská 83, 186 00 Praha 8 - Karlín.