Časopis pro pěstování matematiky, roč. 103 (1978), Praha

A REMARK ON MODELS OF BOX-JENKINS TYPE

HA HUY TOAN, Praha

(Received March 31, 1977)

1. INTRODUCTION

In the well-known book *Time series analysis, forecasting and control* the authors BOX and JENKINS introduced a special non-stationary model for discrete random processes. At the beginning, we shall describe shortly the Box-Jenkins approach. Further, we shall show that there is an inaccuracy in their theory, and a modified model will be described. Then we shall concentrate our attention to a problem arising when the best extrapolation is evaluated.

Let $\{Z_t\}$ be a discrete random process. Define an operator B by $BZ_t = Z_{t-1}$. Then B is called a backward operator. Suppose that $\{a_t\}$ is a white noise. Let $\{Z_t\}$ satisfy the condition

$$\varphi(B)\, Z_t = \Theta(B)\, a_t \tag{1}$$

where

$$\varphi(B) = 1 - \varphi_1 B - \ldots - \varphi_p B^p\,, \quad \Theta(B) = 1 - \Theta_1 B - \ldots - \Theta_q B^q\,.$$

If all the roots of the polynomial $\varphi(B)$ have their absolute values smaller than 1, then $\varphi(B)$ is called a stationary operator. In this case we say that $\{Z_t\}$ is an ARMA process (autoregressive moving average process).

Box and Jenkins investigated in [2] a model

$$\varphi(B)\,(1 - B)^d\, Z_t = \Theta(B)\, a_t \tag{2}$$

where $\varphi(B)$ is a stationary operator. In this case $\{Z_t\}$ is called an ARIMA process (autoregressive integrated moving average process).

STUDENTS' RESEARCH ACTIVITY AT THE FACULTY OF MATHEMATICS AND PHYSICS, CHARLES UNIVERSITY. Awarded the 4th prize in the National Students' Research Work Competition, section Statistics, in the year 1977. Scientific adviser: Professor J. ANDĚL.

Let us suppose for a moment that the random variables $a_t, a_{t-1}, \ldots$ are known. We are looking for the best estimate (or for the "best extrapolation") of the variable Z_{t+g}, when the estimate is based on $a_t, a_{t-1}, \ldots$. We restrict ourselves to linear estimates which have the form $\sum_{k=0}^{\infty} c_k a_{t-k}$. We say that $\hat{Z}_t(g)$ is the best linear extrapolation of Z_{t+g} if

$$E[\hat{Z}_t(g) - Z_{t+g}]^2 \leqq E[Z_t(g) - Z_{t+g}]^2$$

where both $Z_t(g)$ and $\hat{Z}_t(g)$ are of the type $\sum_{k=0}^{\infty} c_k a_{t-k}$.

Let us suppose that there exists such a polynomial $\Psi(B) = 1 + \Psi_1 B + \Psi_2 B^2 + \ldots$ that

(3) $$Z_t = \Psi(B)\, a_t = a_t + \sum_{i=1}^{\infty} a_{t-i} \Psi_i$$

holds for every integer t.

Under these assumptions, Box and Jenkins proved that the best linear extrapolation is

(4) $$\hat{Z}_t(g) = \Psi_g a_t + \Psi_{g+1} a_{t-1} + \ldots.$$

Formula (2) implies

$$\varphi(B)\,(1 - B)^d\, \Psi(B) = \Theta(B)$$

and it is easy to find the coefficients $\Psi_1, \Psi_2, \ldots$ of the polynomial $\Psi(B)$. Also the problem of extrapolation is solved in this way.

The basic problem, however, is that the series

$$a_t + \sum_{k=1}^{\infty} a_{t-k} \Psi_k$$

may not converge, as we can see from the following simple example. If $\varphi(B) = 1 = \Theta(B)$, $d = 1$, then

$$(1 - B)\, Z_t = a_t$$

and we have

$$Z_t = a_t + Z_{t-1} = \ldots = \sum_{i=1}^{\infty} a_{t-i}\,.$$

Since $a_t, a_{t-1}, \ldots$ are independent random variables with the same positive variance, the sum $\sum_{i=0}^{\infty} a_{t-i}$ does not exist.

2. A MODIFIED MODEL

To avoid the problems mentioned above we shall assume that the process $\{Z_t\}$ just starts at a fixed moment, say at $t = 1$.

Let the random variables $Z_1, \ldots, Z_p$ and $a_{p-q+1}, \ldots, a_p, \ldots$ be given. Then it is possible to define Z_t for $t > p$ by a formula

$$\varphi(B)\, Z_t = \Theta(B)\, a_t \tag{5}$$

recurrently. The symbols $\{a_t\}$, $\{Z_t\}$, $\varphi(B)$, $\Theta(B)$ are the same as in (1), but this time we put no condition on the polynomial $\varphi(B)$.

Inserting $t + 1$ instead of t we get

$$Z_{t+1} = \varphi_1 Z_t + \varphi_2 Z_{t-1} + \ldots + \varphi_p Z_{t-p+1} + \Theta(B)\, a_t\,. \tag{6}$$

This formula enables us to express Z_{p+g} in terms of $Z_1, \ldots, Z_p$ and a_t. First, we can see that

$$Z_{p+g} = \varphi_1 Z_{p+g-1} + \ldots + \varphi_p Z_g + \Theta(B)\, a_{p+g}\,.$$

After g steps we have the wanted formula.

If we denote (after k steps)

$$Z_{p+g} = \varphi_k^k Z_{p+g-k} + \ldots + \varphi_{k+p-1}^k Z_{g-k+1} + \Theta(B)\, F^k(B)\, a_{p+g}$$

then after inserting for $\varphi_k^k Z_{p+g-k}$ from (6) we get

$$Z_{p+g} = (\varphi_k^k \varphi_1 + \varphi_{k+1}^k)\, Z_{p+g-k-1} + \ldots + (\varphi_k^k \varphi_{p-1} + \varphi_{p+k-1}^k)\, Z_{g-k+1} + \\ + \varphi_k^k \varphi_p Z_{p-k} + \Theta(B)\,[\varphi_k^k B^k + F^k(B)]\, a_{p+g}\,.$$

Put $\varphi_i^k = 0$ for $i < k$ and for $i - k \geqq p$.

Comparing these formulas we see that

$$\varphi_i^{k+1} = \varphi_k^k \varphi_{i-k} + \varphi_i^k \tag{7}$$

for $i = k + 1, \ldots, k + p$ and

$$F^{k+1}(B) = \varphi_k^k B^k + F^k(B) = \ldots = \sum_{i=0}^{k} \varphi_i^i B^i \tag{8}$$

where $\varphi_0^0 = 1$, $\varphi_i^1 = \varphi_i$.

After the g-th step we come to

$$Z_{p+g} = \varphi_g^g Z_p + \ldots + \varphi_{g+p-1}^g Z_1 + \Theta(B)\, F^g(B)\, a_{p+g} \tag{9}$$

where φ_i^g and $F^g(B)$ are defined recurrently in (7) and (8).

Now, we shall investigate how to express Z_{p+N+T} in the form

$$Z_{p+N+T} = L_1(Z_1, \ldots, Z_{p+N}) + L_2(a_{p-q+1}, \ldots, a_{p-1}, a_p) + \\ + L_3(a_{p+N+1}, \ldots, a_{p+N+T})\,, \tag{10}$$

where L_1, L_2, L_3 denote linear combinations of the members written in the brackets.

Inserting $p + N$ and T instead of p and q we get

$$Z_{p+N+T} = \varphi_T^T Z_{p+N} + \ldots + \varphi_{T+p-1}^T Z_{N+1} + S(B)\, a_{p+N+T} \tag{11}$$

where

$$\Theta(B) \sum_{i=0}^{T-1} \varphi_i^i B^i = S(B) = 1 - s_1 B - \ldots - s_{q+T-1} B^{q+T-1} .$$

If we denote

$$A = s_T a_{p+N} + s_{T+1} a_{p+N-1} + \ldots + s_{q+T-1} a_{p+N-q+1} \tag{12}$$

then

$$S(B)\, a_{p+N+T} = a_{p+N+T} - \ldots - s_{T-1} a_{p+N+1} - A .$$

To get Z_{p+N+T} in the form (10) we must decompose $a_{p+N}, \ldots, a_{p+1}$ (i.e., the expressions in A for $N \geqq q$) into terms containing $Z_1, \ldots, Z_{p+N}$ and $a_{p-q+1}, \ldots, a_p$.

The fundamental formula (5) yields easily

$$a_{t+1} = \Theta_1 a_t + \ldots + \Theta_q a_{t-q+1} + \varphi(B)\, Z_{t+1} .$$

Put $\Theta_0^0 = 1$ and define Θ_i^j recurrently by

$$\Theta_i^{n+1} = \Theta_n^n \Theta_{i-n} + \Theta_i^n \tag{13}$$

where $\Theta_i^1 = \Theta_i$. Then for any positive integer g we have

$$a_{p+g} = \Theta_g^g a_p + \Theta_{g+1}^g a_{p-1} + \ldots + \Theta_{g+q-1}^g a_{p-q+1} + \varphi(B)\, L^g(B)\, Z_{p+g} \tag{14}$$

where $L^g(B) = \sum_{i=0}^{g-1} \Theta_i^i B^i$.

Inserting (14) into (12) gives

$$A = \sum_{i=0}^{q-1} s_{T+i} [\Theta_{N-i}^{N-i} a_p + \ldots + \Theta_{N-i+q-1}^{N-i} a_{p-q+1} + \varphi(B)\, L^{N-i}(B)\, B^i Z_{p+N}] .$$

Putting

$$T_j^N = \sum_{i=0}^{q-1} s_{T+i} \Theta_{N-i+j}^{N-i} , \tag{15}$$

$$\begin{aligned} D^N(B) &= d_0^N + d_1^N B + \ldots + d_{p+N-1}^N B^{p+N-1} \\ &= \sum_{i=0}^{q-1} s_{T+i}\, \varphi(B)\, L^{N-i}(B)\, B^i , \end{aligned} \tag{16}$$

we can write

$$A = T_0^N a_p + T_1^N a_{p-1} + \ldots + T_{q-1}^N a_{p-q+1} + D^N(B)\, Z_{p+N} . \tag{17}$$

According to (12), this implies

$$Z_{p+N+T} = (\varphi_T^T + d_0^N) Z_{p+N} + \ldots + (\varphi_{T+p-1}^T + d_{p-1}^N) Z_{N+1} + \tag{18}$$
$$+ d_p^N Z_N + \ldots + d_{p+N-1}^N Z_1 - T_0^N a_p - T_1^N a_{p-1} - \ldots - T_{q-1}^N a_{p-q+1} +$$
$$+ a_{p+N+T} - s_1 a_{p+N+T-1} - \ldots - s_{T-1} a_{p+N+1} .$$

3. EXTRAPOLATION IN THE MODIFIED MODEL

Suppose that the variables $Z_1, \ldots, Z_{p+n}$ and $a_{p-q+1}, \ldots, a_p$ are known. Let a_j be independent of Z_i for $j > i$. For a given positive integer T we are going to construct an extrapolation of Z_{p+N+T} based on the values $Z_1, \ldots, Z_{p+N}$ and $a_{p-q+1}, \ldots, a_p$. Denote

$$\hat{Z}_{p+N}(T) = z_0^N Z_{p+N} + z_1^N Z_{p+N-1} + \ldots + z_{p+N-1}^N z_1 - \tag{19}$$
$$- T_0^N a_p - \ldots - T_{q-1}^N a_{p-q+1} ,$$

$$\varepsilon_{p+N}(T) = a_{p+N+T} - s_1 a_{p+N+T-1} - \ldots - s_{T-1} a_{p+N+1} , \tag{20}$$

where

$$z_i^N = \varphi_{T+i}^T + d_i^N \quad \text{for} \quad i = 0, \ldots, p-1 ,$$
$$z_i^N = d_i^N \quad \text{for} \quad i = p, \ldots, p+N-1 ,$$

and φ_i^T, s_i, T_i^N, D_i^N are defined respectively in (7), (11), (15) and (16).

Theorem 1. *The variable $\hat{Z}_{p+N}(T)$ is the best linear extrapolation for Z_{p+N+T} based on the given Z_t and a_t.*

Proof. Denote

$$Z_{p+N}(T) = z_0 Z_{p+N} + \ldots + z_{p+N-1} Z_1 - t_0 a_p - \ldots - t_{q-1} a_{p-q+1}$$

and put

$$L = (z_0^N - z_0) Z_{p+N} + \ldots + (z_{p+N-1}^N - z_{p+N-1}) Z_1 -$$
$$- (T_0^N - t_0) a_p - \ldots - (T_{q-1}^N - t_{q-1}) a_{p-q+1} .$$

The variables L and $\varepsilon_{p+N}(T)$ are independent because a_j and Z_i are independent for $j > i$ and a_j and a_i are independent for $j \neq i$. It implies

$$E[Z_{p+N+T} - \hat{Z}_{p+N}(T)]^2 = E[\varepsilon_{p+N}(T)]^2 ,$$
$$E[Z_{p+N+T} - Z_{p+N}(T)]^2 = E[L + \varepsilon_{p+N}(T)]^2 = EL^2 + E[\varepsilon_{p+N}(T)]^2 .$$

This completes the proof.

The variable $\hat{Z}_{p+N}(T)$ contains $a_p, a_{p-1}, \ldots, a_{p-q+1}$. However, the variables $a_p, a_{p-1}, \ldots$ are usually not known and, moreover, it is even hardly possible to derive any good estimates fof them. Under some assumptions concerning the roots of the polynomial $\Theta(B)$ we prove that the influence of those variables rapidly decreases if the length of the realization grows. Usually, the variables $a_{p-q+1}, \ldots, a_p$ are neglected and we insert zeros instead of them. We shall show that this substitution does not influence the result too much if the number of the known variables Z_t is sufficiently large.

Theorem 2. *If all the roots of the polynomial $\Theta(B)$ are outside the unit circle, then $\Theta_N^N \to 0$ for $N \to \infty$, where Θ_N^N is defined in* (13).

Proof. According to (13) we have

$$\Theta_N^N = \Theta_{N-1}^{N-1}\Theta_1 + \Theta_N^{N-1}.$$

Since $\Theta_j^I = 0$ for $I > j$ and for $j - I \geqq q$, we obtain

$$\Theta_N^N = \sum_{i=1}^{q} \Theta_i \Theta_{N-i}^{N-i}. \tag{21}$$

Denote

$$\mathbf{M} = \begin{pmatrix} 0, & 1, & 0, & \ldots, & 0 \\ 0, & 0, & 1, & \ldots, & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ \Theta_q, & \Theta_{q-1}, & \Theta_{q-2}, & \ldots, & \Theta_1 \end{pmatrix}$$

and introduce

$$\Gamma_n = (\Theta_n^n, \Theta_{n+1}^{n+1}, \ldots, \Theta_{n+q-1}^{n+q-1})'.$$

Formula (21) gives

$$\Gamma_{n+1} = \mathbf{M}\Gamma_n$$

so that

$$\Gamma_{n+1} = \mathbf{M}^n \Gamma_1. \tag{22}$$

It is well known that

$$|\mathbf{M} - \lambda \mathbf{I}| = (-1)^{q+2} \lambda^q \Theta(1/\lambda) \tag{23}$$

and we see that all the roots of the matrix $\mathbf{M}$ are smaller then 1 in absolute value.

According to Perron's formula we get

$$\|\mathbf{M}^n\| \leq \sum_{k=1}^{s} \left[(\lambda_k)^n \|\mathbf{Z}_{k_1}\| + \ldots + \frac{n!}{(n - m_k + 1)!} (\lambda_k)^{n-m_k+1} \|\mathbf{Z}_{k_{m_k}}\| \right] \tag{24}$$

where $\|\mathbf{A}\|$ is a norm of the matrix $\mathbf{A}$ defined by

$$\|\mathbf{A}\|^2 = \sum_{i,j} a_{ij}^2.$$

Denote

$$\lambda = \max_{1 \leqq k \leqq s} |\lambda_k| ,$$

$$m = \max_{1 \leqq k \leqq s} (m_k - 1) ,$$

(25) $$\|\mathbf{Z}\| = \max_{1 \leqq k \leqq s} (\|\mathbf{Z}_{k_1}\| + \ldots + \|\mathbf{Z}_{k_{m_k}}\|) .$$

Then from (24) we obtain

(26) $$\|\mathbf{M}^n\| \leqq \frac{s\|\mathbf{Z}\| \, n! \, \lambda^{n-m}}{(n - m)!} .$$

Put

$$a_n = \frac{s\|\mathbf{Z}\| \, n! \, \lambda^{n-m}}{(n - m)!} .$$

We see from (25) that $0 \leqq \lambda < 1$. Then for any sufficiently small $\varepsilon > 0$ we have

$$\lambda(1 + \varepsilon) < 1$$

and for any $n \geqq n_0$ we have also

$$\frac{m}{n + 1 - m} < \varepsilon .$$

This yields

(27) $$\left|\frac{a_{n+1}}{a_n}\right| < \lambda(1 + \varepsilon) < 1 .$$

Denote

$$(1 + \varepsilon) \lambda = \lambda^* .$$

Then we see that

(28) $$|a_{n_0+r}| < |a_{n_0}| \, (\lambda^*)^r .$$

Formulas (26), (28), (22) give

$$\|\mathbf{M}^n\| < a_n \leqq |a_n| < \frac{|a_{n_0}|}{(\lambda^*)^{n_0}} (\lambda^*)^n , \quad \mathbf{\Gamma}_{n+1} = \mathbf{M}^n \mathbf{\Gamma}_1 .$$

We see that

(29) $$|\Theta_{j+n}^{j+n}| \leqq \|\mathbf{\Gamma}_{n+1}\| \leqq \|\mathbf{M}^n\| \, \|\mathbf{\Gamma}_1\| < \frac{\|\mathbf{\Gamma}_1\| \, |a_{n_0}|}{(\lambda^*)^{n_0}} (\lambda^*)^n$$

for $j = 0, \ldots, q - 1$.

We can choose such a small ε that $\lambda^* = \lambda(1 + \varepsilon)$ is also smaller than 1. We see from (29) that Θ_N^N exponentially converges to zero.

Theorem 3. *If all the roots of $\Theta(B)$ are outside the unit circle then T_j^N converges exponentially to zero for $N \to \infty$.*

Proof. From recurrent formula (13) we easily get

$$\Theta_{N+\alpha}^{N} = \sum_{i=\alpha+1}^{q} \Theta_{N+\alpha-i}^{N+\alpha-i} \Theta_i . \tag{30}$$

Since $\Theta_{N+\alpha-i}^{N+\alpha-i}$ converges exponentially to zero, we see that for any fixed α the coefficients $\Theta_{N+\alpha}^{N}$ also converge exponentially to zero. As

$$T_j^N = \sum_{i=0}^{q-1} s_{T+i} \Theta_{N-i+j}^{N-i} ,$$

it is clear that T_j^N also converges exponentially to zero. The proof is complete.

4. MULTIDIMENSIONAL MODEL

We have used a modified Box-Jenkins model for the scalar case. In this section we generalize the results to the vector case.

Suppose that $\boldsymbol{a}_{p-q+1}, \ldots, \boldsymbol{a}_p, \ldots$ and $\boldsymbol{Z}_1, \ldots$ are k-dimensional random vectors. We shall assume that

$$E\boldsymbol{a}_i = 0 , \quad \operatorname{var} \boldsymbol{a}_i = \boldsymbol{A} ,$$

$$\operatorname{cov}(\boldsymbol{a}_i, \boldsymbol{a}_j) = \boldsymbol{0} \quad \text{for} \quad i \neq j ,$$

$$\operatorname{cov}(\boldsymbol{a}_i, \boldsymbol{Z}_j) = \boldsymbol{0} \quad \text{for} \quad i > j .$$

Define $\boldsymbol{Z}_t (t \geqq p)$ recurrently by

$$\boldsymbol{\varphi}(B)\, \boldsymbol{Z}_t = \boldsymbol{\Theta}(B)\, \boldsymbol{a}_t$$

where

$$\boldsymbol{\varphi}(B) = \boldsymbol{I} - \boldsymbol{\varphi}_1 B - \ldots - \boldsymbol{\varphi}_p B^p , \quad \boldsymbol{\Theta}(B) = \boldsymbol{I} - \boldsymbol{\Theta}_1 B - \ldots - \boldsymbol{\Theta}_q B^q ,$$

and $\boldsymbol{\varphi}_i$, $\boldsymbol{\Theta}_i$ are matrices of the type (k, k).

Let $\mathscr{M}$ be the set of the random vectors

$$\{z_0 \boldsymbol{Z}_{p+N} + \ldots + z_{p+N-1} \boldsymbol{Z}_1 - t_0 \boldsymbol{a}_p - \ldots - t_{q-1} \boldsymbol{a}_{p-q+1}\}$$

where z_k and t_j are arbitrary real numbers. Let $\hat{\boldsymbol{Z}}_{p+N}(T)$ be a fixed element from $\mathscr{M}$. If the difference

$$\operatorname{var}[\boldsymbol{Z}_{p+N}(T) - \boldsymbol{Z}_{p+N+T}] - \operatorname{var}[\hat{\boldsymbol{Z}}_{p+N}(T) - \boldsymbol{Z}_{p+N+T}]$$

is a positive semidefinite matrix for all $\boldsymbol{Z}_{p+N}(T) \in \mathscr{M}$, we say that $\hat{\boldsymbol{Z}}_{p+N}(T)$ is the best linear extrapolation for $\boldsymbol{Z}_{p+N+T}$.

It is easy to see that $\hat{\boldsymbol{Z}}_{p+N}(T)$ defined quite analogously as in (19) is the best linear extrapolation for $\boldsymbol{Z}_{p+N+T}$.