COMMENTATIONES MATHEMATICAE UNIVERSITATIS CAROLINAE
28,3 (1987)

STUDY ON KALMAN FILTER IN TIME SERIES ANALYSIS

T. CIPRA, I. MOTYKOVÁ

Abstract: Some procedures of estimation and prediction based on Kalman filter in multivariate time series models of the type ARMA are suggested in the paper. Cases of multivariate time series with missing observations and with components known till various time periods are also considered. Numerical simulations demonstrate some of the results.

Key words: ARMA model, estimation, Kalman filter, missing observation, multivariate time series, prediction.

Classification: 62M10, 62M20, 60G25, 60G35

I. Introduction. Kalman filtering [6],[7] can be used as a very practical instrument for the adaptive estimation and prediction of time series not only in technical applications but also for shorter (e.g. economic) time series (see e.g. [1],[11]). As far as Kalman filtering in time series analysis is concerned, some authors prefer to construct, by means of the Kalman filter, the (exact) likelihood function of the time series models (see [4],[8],[12] and others). However, in this paper the Kalman filter provides directly the recursive estimation and prediction formulas which are optimal in the sense of the least squares principle.

After the recapitulation of the classical results we suggest some estimation and prediction procedures for multivariate ARMA models. Moreover, it is shown how to modify these procedures when some components of observations are missing or when we know the particular components of the multivariate time series till various time periods.

We shall apply the Kalman filter in the context of the following discrete linear dynamic system

(1.1) $$x_{t+1}=\Phi_t x_t+\Gamma_t w_{t+1},$$

(1.2) $$y_t=M_t x_t+v_t,$$

where (1.1) is the state equation and (1.2) is the observation equation of the system. Here x_t is the (vector) state variable of the type $(m_t,1)$ at time t;

y_t is the (vector) observation of the type $(n_t,1)$ at t; Φ_t, Γ_t, M_t are matrices of the type (m_{t+1},m_t), (m_{t+1},q_{t+1}), (n_t,m_t), respectively (Φ_t is the state transition matrix or system matrix, Γ_t is the input matrix and M_t is the observation matrix at t); w_t and v_t are random vectors of the type $(q_t,1)$ and $(n_t,1)$ fulfilling

$$E(w_t)=0,\ E(v_t)=0,$$
$$\mathrm{var}(w_t)=Q_t,\ \mathrm{var}(v_t)=R_t,$$
$$\mathrm{cov}(w_s,w_t)=0,\ \mathrm{cov}(v_s,v_t)=0,\ s\neq t,$$
$$\mathrm{cov}(w_s,v_t)=0$$

with variance matrices Q_t and R_t of the type (q_t,q_t) and (n_t,n_t) at t. Moreover, the initial value x_0 of the state variable is assumed to fulfil

$$\mathrm{cov}(x_0,w_t)=0,\ \mathrm{cov}(x_0,v_t)=0.$$

If Y_t denotes the $(n_1+\ldots+n_t)$-dimensional Hilbert space spanned by the components of the random vectors $y_1,\ldots,y_t$ then the Kalman filter produces recursively the orthogonal projections $\hat{x}_t^t$ and $\hat{x}_{t+1}^t$ of x_t and x_{t+1} into Y_t together with the matrices

$$P_t^t=E(x_t-\hat{x}_t^t)(x_t-\hat{x}_t^t)',\quad P_{t+1}^t=E(x_{t+1}-\hat{x}_{t+1}^t)(x_{t+1}-\hat{x}_{t+1}^t)'.$$

The filter can be written in the form

$$\hat{x}_{t+1}^t=\Phi_t\hat{x}_t^t, \tag{1.3}$$

$$P_{t+1}^t=\Phi_t P_t^t\Phi_t'+\Gamma_t Q_{t+1}\Gamma_t', \tag{1.4}$$

$$\hat{x}_t^t=\hat{x}_t^{t-1}+K_t(y_t-M_t\hat{x}_t^{t-1}), \tag{1.5}$$

$$P_t^t=(I-K_tM_t)P_t^{t-1}, \tag{1.6}$$

where

$$K_t=P_t^{t-1}M_t'(M_tP_t^{t-1}M_t'+R_t)^{-1}=P_t^tM_t'R_t^{-1}. \tag{1.7}$$

The matrix $M_tP_t^{t-1}M_t'+R_t$ is supposed to be regular (i.e. positively definite) in the first expression (1.7) and the matrix R_t is supposed to be regular in the second expression (1.7) for K_t (moreover, the regularity of R_t also guarantees the regularity of $M_tP_t^{t-1}M_t'+R_t$). The derivation of (1.3) - (1.7) is given e.g. in [5, p. 201] or [10, p. 807] with the only exception that the dimensions of all vectors and matrices in the system (1.1) and (1.2) are considered to be constant in time. However, the proof can be extended to our case with changing dimensions (see [9]).

The relations (1.3), (1.4) are the prediction steps and (1.5) - (1.7) are the correction steps of the Kalman filtering algorithm. If connecting these two steps one can obviously write

(1.8) $$\hat{x}_{t+1}=\Phi_t\hat{x}_t+K_{t+1}(y_{t+1}-M_{t+1}\Phi_t\hat{x}_t),$$

(1.9) $$P_{t+1}=(I-K_{t+1}M_{t+1})(\Phi_tP_t\Phi_t'+\Gamma_tQ_{t+1}\Gamma_t'),$$

(1.10) $$K_{t+1}=(\Phi_tP_t\Phi_t'+\Gamma_tQ_{t+1}\Gamma_t')M_{t+1}'[M_{t+1}(\Phi_tP_t\Phi_t'+\Gamma_tQ_{t+1}\Gamma_t')M_{t+1}'+R_{t+1}]^{-1}$$
$$=P_{t+1}M_{t+1}'R_{t+1}^{-1},$$

where we put for simplicity

(1.11) $$\hat{x}_t=\hat{x}_t^t,\ P_t=P_t^t.$$

If the system matrix Φ and the observation matrix M are constant in time then the orthogonal projection $\hat{x}_{t+k}^t$ of x_{t+k} into Y_t has the form

(1.12) $$\hat{x}_{t+k}^t=\Phi^k\hat{x}_t^t,\ k\geq 0$$

so that the prediction $\hat{y}_{t+k}^t$ of y_{t+k} at time t can be written as

(1.13) $$\hat{y}_{t+k}^t=M\hat{x}_{t+k}^t=M\Phi^k\hat{x}_t^t,\ k\geq 0.$$

2. Adaptive parameter estimation in time series models by Kalman filter. First we remind briefly the adaptive parameter estimation results (see e.g. [3, p. 61]) in the classical linear regression model of the form

(2.1) $$y_t=x_t'b+\varepsilon_t,$$

where $x_t=(x_{1,t},\dots,x_{r,t})'$ is the (r,1) vector of regressors at time t, $b=(b_1,\dots,b_r)'$ is the (r,1) vector of regression parameters and ε_t is the residual at time t such that $E(\varepsilon_t)=0$, $var(\varepsilon_t)=\sigma^2$, $cov(\varepsilon_s,\varepsilon_t)=0$ for $s\neq t$ ($\sigma^2>0$ is the further unknown parameter of the model).

In this case the state space representation (1.1) and (1.2) suitable for the adaptive estimation of the parameters b can be written as

(2.2) $$b_{t+1}=b_t,$$

(2.3) $$y_t=x_t'b_t+\varepsilon_t,$$

where $x_t=b_t$, $v_t=\varepsilon_t$, $\Phi_t=I$, $\Gamma_t=0$, $M_t=x_t'$, $Q_t=0$, $R_t=\sigma^2$. Now the relations (1.8) and (1.9) using (1.10) have the form

(2.4) $$\hat{b}_{t+1}=\hat{b}_t+\sigma^{-2}P_{t+1}x_{t+1}(y_{t+1}-x_{t+1}'\hat{b}_t),$$

(2.5) $$P_{t+1}=P_t-(x_{t+1}'P_tx_{t+1}+\sigma^2)^{-1}P_tx_{t+1}x_{t+1}'P_t.$$

Since the parameter σ^2 is unknown one can put

$$V_t = \sigma^{-2} P_t \tag{2.6}$$

and rewrite (2.4) and (2.5) to the form

$$\hat{b}_{t+1} = \hat{b}_t + V_{t+1} x_{t+1} (y_{t+1} - x'_{t+1} \hat{b}_t), \tag{2.7}$$

$$V_{t+1} = V_t - (x'_{t+1} V_t x_{t+1} + 1)^{-1} V_t x_{t+1} x'_{t+1} V_t, \tag{2.8}$$

which are the formulas of the recursive least squares method for the model (2.1). As far as the adaptive estimation of the parameter σ^2 is concerned one can obtain using the relation

$$\sigma_t^2 = \frac{1}{t-r} \sum_{i=1}^{t} (y_i - x'_i \hat{b}_t)^2$$

the adaptive formula (see [3, p. 78])

$$\hat{\sigma}^2_{t+1} = \frac{1}{t+1-r} [(t-r)\hat{\sigma}_t^2 + (x'_{t+1} V_t x_{t+1} + 1)^{-1} (y_{t+1} - x'_{t+1} \hat{b}_t)^2]. \tag{2.9}$$

The matrix

$$\hat{P}_{t+1} = \hat{\sigma}^2_{t+1} \; V_{t+1} \tag{2.10}$$

can be taken as the estimate of the variance matrix of $\hat{b}_{t+1}$. If there is no apriori information on the parameters one can choose the initial values of the estimates at time t_o, e.g., as

$$\hat{b}_{t_o} = 0, \; V_{t_o} = c^{-1} I, \; \hat{\sigma}^2_{t_o} = c, \tag{2.11}$$

where c is a small positive constant (see [3]).

In the following text adaptive estimation formulas are suggested for some multivariate time series models.

2.1. **Estimation in multivariate AR process.** Let us consider an n-dimensional AR(p) model of the form

$$y_t = \Phi_1 y_{t-1} + \dots + \Phi_p y_{t-p} + \varepsilon_t, \tag{2.12}$$

where $\Phi_1, \dots, \Phi_p$ are (n,n) matrices of parameters and $\{\varepsilon_t\}$ is an n-dimensional white noise, i.e. $E(\varepsilon_t) = 0$, $\mathrm{var}(\varepsilon_t) = \Sigma > 0$, $\mathrm{cov}(\varepsilon_s, \varepsilon_t) = 0$ for $s \neq t$. Let us denote

$$b = \begin{pmatrix} \mathrm{vec}\,\Phi_1 \\ \mathrm{vec}\,\Phi_2 \\ \vdots \\ \mathrm{vec}\,\Phi_p \end{pmatrix}, \quad X_t = \left(\begin{matrix} y'_{t-1} & & 0 \\ & y'_{t-1} & \\ & & \ddots \\ 0 & & y'_{t-1} \end{matrix} \middle| \; \cdots \; \middle| \begin{matrix} y'_{t-p} & & 0 \\ & y'_{t-p} & \\ & & \ddots \\ 0 & & y'_{t-p} \end{matrix} \right),$$

where the operation vec Φ arranges the rows of a matrix Φ to a column vector. The dimension of the vector b is $(n^2p,1)$ and the matrix X_t is (n,n^2p).

The adaptive estimation formulas for the system

$$b_{t+1}=b_t, \tag{2.13}$$

$$y_t=X_tb_t+\varepsilon_t \tag{2.14}$$

analogous to (2.4), (2.5) and (2.9) are

$$\hat{b}_{t+1}=\hat{b}_t+P_{t+1}X_{t+1}'\hat{\Sigma}_t^{-1}(y_{t+1}-X_{t+1}\hat{b}_t), \tag{2.15}$$

$$P_{t+1}=P_t-P_tX_{t+1}'(X_{t+1}P_tX_{t+1}'+\hat{\Sigma}_t)^{-1}X_{t+1}P_t, \tag{2.16}$$

$$\hat{\Sigma}_{t+1}=\frac{1}{t+1-n^2p}[(t-n^2p)\hat{\Sigma}_t+(y_{t+1}-X_{t+1}\hat{b}_{t+1})(y_{t+1}-X_{t+1}\hat{b}_{t+1})']. \tag{2.17}$$

Remark 1. In the formulas (2.15) and (2.16) for $\hat{b}_{t+1}$ and P_{t+1} the estimate $\hat{\Sigma}_t$ from time t is used. One can improve this procedure calculating at time t+1 auxiliary values

$$b^*_{t+1}=\hat{b}_t+P^*_{t+1}X_{t+1}'\hat{\Sigma}_t^{-1}(y_{t+1}-X_{t+1}\hat{b}_t),$$

$$P^*_{t+1}=P_t-P_tX_{t+1}'(X_{t+1}P_tX_{t+1}'+\hat{\Sigma}_t)^{-1}X_{t+1}P_t$$

and then the final values for time t+1

$$\hat{b}_{t+1}=\hat{b}_t+P_{t+1}X_{t+1}'\hat{\Sigma}_{t+1}^{-1}(y_{t+1}-X_{t+1}\hat{b}_t),$$

$$P_{t+1}=P_t-P_tX_{t+1}'(X_{t+1}P_tX_{t+1}'+\hat{\Sigma}_{t+1})^{-1}X_{t+1}P_t,$$

$$\hat{\Sigma}_{t+1}=\frac{1}{t+1-n^2p}[(t-n^2p)\hat{\Sigma}_t+(y_{t+1}-X_{t+1}\hat{b}^*_{t+1})(y_{t+1}-X_{t+1}\hat{b}^*_{t+1})'].$$

In the case of a multivariate AR(p) model we can also proceed in the following way using the state space representation for particular components of the process. The model (2.12) can be rewritten to the form

$$\Phi_0y_t=\Phi_1y_{t-1}+\dots+\Phi_py_{t-p}+\varepsilon_t, \tag{2.18}$$

where Φ_0 is a lower triangular matrix with unities on the main diagonal and an n-dimensional white noise $\{\varepsilon_t\}$ has a diagonal variance matrix

$$\text{var}(\varepsilon_t)=\begin{pmatrix}\sigma_1^2 & & 0\\ & \ddots & \\ 0 & & \sigma_n^2\end{pmatrix}$$

Then it holds for particular components of (2.18)

$$
(2.19)\quad
\begin{aligned}
y_{1,t} &= && \varphi_{11}^{(1)} y_{1,t-1} + \dots + \varphi_{1n}^{(p)} y_{n,t-p} + \varepsilon_{1,t}\\
y_{2,t} &= -\varphi_{21}^{(o)} y_{1,t} && + \varphi_{21}^{(1)} y_{1,t-1} + \dots + \varphi_{2n}^{(p)} y_{n,t-p} + \varepsilon_{2,t}\\
&\vdots\\
y_{n,t} &= -\varphi_{n1}^{(o)} y_{1,t} - \dots - \varphi_{n,n-1}^{(o)} y_{n-1,t} && + \varphi_{n1}^{(1)} y_{1,t-1} + \dots + \varphi_{nn}^{(p)} y_{n,t-p} + \varepsilon_{n,t},
\end{aligned}
$$

where $\Phi_i = (\varphi_{jk}^{(i)})$, $j,k=1,\dots,n$; $i=0,1,\dots,p$.

The particular relationships of (2.19) have the following state space representation for $i=1,\dots,n$

$$(2.20)\qquad b_{nt+i+1} = b_{nt+i},$$

$$(2.21)\qquad z_{nt+i} = x'_{nt+i} b_{nt+i} + v_{nt+i},$$

where

$$b = (\varphi_{11}^{(1)}, \dots, \varphi_{1n}^{(1)}, \dots, \varphi_{11}^{(p)}, \dots, \varphi_{1n}^{(p)}, \varphi_{21}^{(o)}, \varphi_{21}^{(1)}, \dots, \varphi_{2n}^{(1)}, \dots, \varphi_{21}^{(p)}, \dots,$$
$$\varphi_{2n}^{(p)}, \dots, \varphi_{n1}^{(o)}, \dots, \varphi_{n,n-1}^{(o)}, \varphi_{n1}^{(1)}, \dots, \varphi_{nn}^{(1)}, \dots, \varphi_{n1}^{(p)}, \dots, \varphi_{nn}^{(p)})',$$

$$z_{nt+i} = y_{i,t},\quad v_{nt+i} = \varepsilon_{i,t},$$

$$x_{nt+i} = (0, \dots, 0, -y_{1,t}, \dots, -y_{i-1,t}, y_{1,t-1}, \dots, y_{n,t-1}, \dots, y_{1,t-p}, \dots, y_{n,t-p}, 0, \dots$$
$$\dots, 0)',$$

$$\operatorname{var}(v_{nt+i}) = \sigma_i^2$$

(the number of zero components in the vector x_{nt+i} follows from the i-th relationship of (2.19)). The adaptive estimation formulas are for $i=1,\dots,n$

$$(2.22)\qquad \hat{b}_{nt+i} = \hat{b}_{nt+i-1} + P_{nt+i} x_{nt+i}\, \sigma^{-2} (z_{nt+i} - x'_{nt+i} \hat{b}_{nt+i-1}),$$

$$(2.23)\qquad P_{nt+1} = P_{nt+i-1} - P_{nt+i-1} x_{nt+i} [x'_{nt+i} P_{nt+i-1} x_{nt+i} + \hat{\sigma}^2_{n(t-1)+i}]^{-1} x'_{nt+i} P_{nt+i-1},$$

$$(2.24)\qquad \hat{\sigma}^2_{nt+i} = \frac{1}{t-np-i} [(t-np-i-1)\hat{\sigma}^2_{n(t-1)+i} + (z_{nt+i} - x'_{nt+i}\hat{b}_{nt+i-1})^2].$$

2.2. Estimation in multivariate ARMA process. Let us consider an n-dimensional ARMA (p,q) model of the form

$$(2.25)\qquad y_t = \Phi_1 y_{t-1} + \dots + \Phi_p y_{t-p} + \varepsilon_t + \Theta_1 \varepsilon_{t-1} + \dots + \Theta_q \varepsilon_{t-q},$$

where $\Phi_1, \dots, \Phi_p, \Theta_1, \dots, \Theta_q$ are (n,n) matrices of parameters and $\{\varepsilon_t\}$ has the same form as in (2.12). Let us denote

$$b=\begin{pmatrix} \mathrm{vec}\,\Phi_1 \\ \vdots \\ \mathrm{vec}\,\Phi_p \\ \mathrm{vec}\,\Theta_1 \\ \vdots \\ \mathrm{vec}\,\Theta_q \end{pmatrix},$$

$$\hat{X}_t=\left(\begin{matrix} y'_{t-1}, & & 0 \\ & y'_{t-1} & \\ 0 & & \ddots\, y'_{t-1} \end{matrix}\middle|\cdots\middle|\begin{matrix} y'_{t-p}, & & 0 \\ & y'_{t-p} & \\ 0 & & \ddots\, y'_{t-p} \end{matrix}\middle|\begin{matrix} \hat{\varepsilon}'_{t-1} & & 0 \\ & \hat{\varepsilon}'_{t-1} & \\ 0 & & \ddots\, \hat{\varepsilon}'_{t-1} \end{matrix}\middle|\cdots\right.$$

$$\left.\cdots\middle|\begin{matrix} \hat{\varepsilon}'_{t-q} & & 0 \\ & \hat{\varepsilon}'_{t-q} & \\ 0 & & \ddots\, \hat{\varepsilon}'_{t-q} \end{matrix}\right),$$

$$\hat{\varepsilon}_t=y_t-\hat{X}_t\hat{b}_t.$$

The dimension of the vector b is $(n^2(p+q),1)$ and the matrix $\hat{X}_t$ is $(n,n^2(p+q))$.

The adaptive estimation formulas for the system

(2.26) $\quad b_{t+1}=b_t,$

(2.27) $\quad y_t=\hat{X}_t b_t+\varepsilon_t$

have the following form

(2.28) $\quad \hat{b}_{t+1}=\hat{b}_t+P_{t+1}\hat{X}'_{t+1}\hat{\Sigma}_t^{-1}(y_{t+1}-\hat{X}_{t+1}\hat{b}_t),$

(2.29) $\quad P_{t+1}=P_t-P_t\hat{X}'_{t+1}(\hat{X}_{t+1}P_t\hat{X}'_{t+1}+\hat{\Sigma}_t)^{-1}\hat{X}_{t+1}P_t,$

(2.30) $\hat{\Sigma}_{t+1}=\dfrac{1}{t+1-n^2(p+q)}\{[t-n^2(p+q)]\hat{\Sigma}_t+(y_{t+1}-\hat{X}_{t+1}\hat{b}_{t+1})(y_{t+1}-\hat{X}_{t+1}\hat{b}_{t+1})'\},$

(2.31) $\quad \hat{\varepsilon}_{t+1}=y_{t+1}-\hat{X}_{t+1}\hat{b}_{t+1}.$

Remark 2. The same improvement as in Remark 1 or the treatment of particular components from Section 2.1 is also possible for the model ARMA.

2.3. Estimation in multivariate AR process with missing observations. Let us consider the model (2.12) and let only the components $i_1,\dots,i_d$ $(1\leq i_1<\dots<i_d\leq n)$ be at our disposal at time t. The indices $i_1,\dots,i_d$ may change in time so that one should write $i_1(t),\dots,i_{d(t)}(t)$. If we preserve the denotation from Section 2.1 and, moreover, denote S_t the matrix of the type (d,n) which has unities in the positions $(1,i_1),\dots,(d,i_d)$ and zeroes in the remaining positions then one can modify the state space representation (2.13) and (2.14) to the form

(2.32) $$b_{t+1}=b_t,$$

(2.33) $$z_t=\hat{M}_t b_t+v_t,$$

where

$$z_t=S_t y_t,\quad v_t=S_t\,\varepsilon_t,\quad \hat{M}_t=S_t\hat{X}_t,$$

$$X_t=\left(\begin{array}{ccc|c|ccc} \hat{y}'_{t-1} & & 0 & & \hat{y}'_{t-p} & & 0\\ & \hat{y}'_{t-1} & & \cdots & & \hat{y}'_{t-p} & \\ 0 & & \ddots\;\hat{y}'_{t-1} & & 0 & & \ddots\;\hat{y}'_{t-p} \end{array}\right),$$

$$\hat{y}_t=\hat{X}_t\hat{b}_t,$$

$$\mathrm{var}(v_t)=S_t\,\Sigma\,S'_t.$$

Since the vector z_t is completely observable at time t, one can use the following adaptive estimation formulas

(2.34) $$\hat{b}_{t+1}=\hat{b}_t+P_{t+1}\hat{M}^{*\prime}_{t+1}(S_{t+1}\hat{\Sigma}_t\,S'_{t+1})^{-1}(z_{t+1}-\hat{M}^*_{t+1}\hat{b}_t),$$

(2.35) $$P_{t+1}=P_t-P_t\hat{M}^{*\prime}_{t+1}(\hat{M}^*_{t+1}P_t\hat{M}^{*\prime}_{t+1}+S_{t+1}\hat{\Sigma}_t\,S'_{t+1})^{-1}\hat{M}^*_{t+1}P_t,$$

(2.36) $$\hat{\Sigma}_{t+1}=\frac{1}{t+1-n^2(p+q)}\{[t-n^2(p+q)]\,\hat{\Sigma}_t+(\hat{y}^*_{t+1}-\hat{X}^*_{t+1}\hat{b}_{t+1})(\hat{y}^*_{t+1}-\hat{X}^*_{t+1}\hat{b}_{t+1})'\},$$

(2.37) $$\hat{y}_{t+1}=\hat{X}^*_{t+1}\hat{b}_{t+1},$$

(2.38) $$\hat{M}^*_{t+1}=S_{t+1}\hat{X}^*_{t+1}.$$

The vector $\hat{y}^*_t$ and the matrix $\hat{X}^*_t$ originate from the vector $\hat{y}_t$ and the matrix $\hat{X}_t$ using the known components of y_t, y_{t-1}, $y_{t-2},\dots$ in $\hat{y}_t,\hat{y}_{t-1},\hat{y}_{t-2},\dots$.

3. Prediction in time series models by Kalman filter. Besides its use in the adaptive estimation the Kalman filter is convenient for the construction of prediction in time series models.

3.1. Prediction in multivariate ARMA process. Let us consider an n-dimensional model of the form

$$y_t = \Phi_1 y_{t-1} + \dots + \Phi_p y_{t-p} + \varepsilon_t + \Theta_1 \varepsilon_{t-1} + \dots + \Theta_{p-1} \varepsilon_{t-p-1}, \tag{3.1}$$

where $\Phi_1, \dots, \Phi_p$, $\Theta_1, \dots, \Theta_{p-1}$ are known (n,n) matrices and $\{\varepsilon_t\}$ has the same form as in (2.12) with a known variance matrix Σ. A more general ARMA (p^*,q^*) model can be transformed to the ARMA (p,p-1) in (3.1) by introducing zero parameter matrices if it is necessary. As far as the known parameter matrices are concerned, e.g., they could be estimated from the observations which we have at our disposal for the construction of prediction in the given process.

The state space representation of the model (3.1) for the purpose of prediction can be written in the form (1.1) and (1.2), where

$$\begin{pmatrix} x_{1,t+1} \\ x_{2,t+1} \\ \vdots \\ x_{p,t+1} \end{pmatrix} = \begin{pmatrix} 0 & 0 \dots 0 & \Phi_p \\ I & 0 \dots 0 & \Phi_{p-1} \\ \vdots & & \\ 0 & 0 \dots I & \Phi_1 \end{pmatrix} \begin{pmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{p,t} \end{pmatrix} + \begin{pmatrix} \Theta_{p-1} \\ \Theta_{p-2} \\ \vdots \\ I \end{pmatrix} \varepsilon_{t+1}, \tag{3.2}$$

$$y_t = (0 \dots 0 \; I)(x_{1,t}', \dots, x_{p,t}')', \tag{3.3}$$

i.e.

$$x_t = \begin{pmatrix} x_{1,t} \\ x_{2,t} \\ \vdots \\ x_{p,t} \end{pmatrix}, \quad \Phi_t = \Phi = \begin{pmatrix} 0 & 0 \dots 0 & \Phi_p \\ I & 0 \dots 0 & \Phi_{p-1} \\ \vdots & & \\ 0 & 0 \dots I & \Phi_1 \end{pmatrix}, \quad \Gamma_t = \Gamma = \begin{pmatrix} \Theta_{p-1} \\ \Theta_{p-2} \\ \vdots \\ I \end{pmatrix},$$

$$M_t = M = (0 \dots 0 \; I), \; w_t = \varepsilon_t, \; Q_t = \Sigma, \; R_t = 0.$$

The Kalman filter (1.3) - (1.7) has the form

$$\hat{x}_{t+1}^t = \Phi \hat{x}_t^t, \tag{3.4}$$

$$P_{t+1}^t = \Phi P_t^t \Phi' + \Gamma \Sigma \Gamma', \tag{3.5}$$

$$\hat{x}_t^t = \hat{x}_t^{t-1} + K_t (y_t - M \hat{x}_t^{t-1}), \tag{3.6}$$

(3.7) $$P_t^t=(I-K_tM)P_t^{t-1},$$

(3.8) $$K_t=P_t^{t-1}M'(MP_t^{t-1}M')^{-1}.$$

Then the prediction $\hat{y}_{t+k}^t$ of the process value y_{t+k} constructed at time t for k steps ahead is given according to the formula (1.13).

3.2. Prediction in multivariate ARMA process with missing observations. If we use the matrix S_t from Section 2.3 reflecting the position of the missing components of the observations at time t for the model (3.1) then the vector z_t which we observe at t fulfils

$$z_t=S_ty_t$$

and we must replace the relationships (3.6) - (3.8) in the filter (3.4) - (3.8) by the following ones

(3.9) $$\hat{x}_t^t=\hat{x}_t^{t-1}+K_t(z_t-S_tM\hat{x}_t^{t-1}),$$

(3.10) $$P_t^t=(I-K_tS_tM)P_t^{t-1},$$

(3.11) $$K_t=P_t^{t-1}M'S_t'(S_tMP_t^{t-1}M'S_t')^{-1}.$$

The formula (1.13) can be used not only for prediction (if $k>0$) but also for the estimation of the missing components of y_t (if $k=0$).

3.3. Prediction in multivariate ARMA process when components are known till various time periods. Let the particular components of an n-dimensional ARMA process $\{y_t\} = \{(y_{1,t},\dots,y_{n,t})'\}$ are known till time periods $t_1,\dots,t_n$ with the following ordering

$$t_{i_1}\leq t_{i_2}\leq\dots\leq t_{i_n}.$$

Such situation is frequent in various practical applications (see also [2]). Then, e.g., the value

$$\hat{x}_{t_{i_n}}^{t_{i_n}(t_1,\dots,t_n)}$$

produced by the Kalman filter at time t_{i_n} can be used for the construction of the following prediction

(3.12) $$\hat{y}_{t_{i_n}+k}^{t_{i_n}(t_1,\dots,t_n)}=M\Phi^k\hat{x}_{t_{i_n}}^{t_{i_n}(t_1,\dots,t_n)}.$$

The upper indices of the values $\hat{y}$ and $\hat{x}$ in (3.12) denote that these values are constructed at time t_{i_n} using $y_{1,t}$ for $t \le t_1, \ldots, y_{n,t}$ for $t \le t_n$.

Let us consider the previous situation and let the value y_{2,t_2+1} be newly supplied. If it is e.g. $t_2 \ge t_1$, $t_2 \ge t_3$, $t_2 \le t_4, \ldots, t_2 \le t_n$ then one can start with the value

$$\hat{x}_{t_2}^{t_2(t_1,t_2,t_3,t_2,\ldots,t_2)}$$

and calculate gradually by Kalman filtering the values

$$\hat{x}_{t_2+1}^{t_2+1(t_1,t_2+1,t_3,t_2+1,\ldots,t_2+1)},\ \hat{x}_{t_2+2}^{t_2+2(t_1,t_2+1,t_3,t_2+2,\ldots,t_2+2)},\ldots,\ x_{t_{i_n}}^{t_{i_n}(t_1,t_2+1,t_3,t_4,\ldots,t_n)}.$$

According to (3.12) the last value can be used for the construction of the updated prediction

$$\hat{y}_{t_{i_n}+k}^{t_{i_n}(t_1,t_2+1,t_3,\ldots,t_n)} = M\Phi^k \hat{x}_{t_{i_n}}^{t_{i_n}(t_1,t_2+1,t_3,\ldots,t_n)}$$

4. Numerical simulations. Let us demonstrate numerically some of the suggested procedures for the simulated two-dimensional AR (1) process

$$y_t = \Phi_1 y_{t-1} + \varepsilon_t \tag{4.1}$$

with

$$\Phi_1 = \begin{pmatrix} 0.5 & 0.1 \\ -0.2 & 0.8 \end{pmatrix},\ \varepsilon_t \sim \text{iid } N_2(0,\Sigma),\ \Sigma = \begin{pmatrix} 1 & -1 \\ -1 & 5 \end{pmatrix}. \tag{4.2}$$

Example 1. The adaptive formulas (2.15) - (2.17) from Section 2.1 were applied with the initial values

$$t_0 = 5,\ \hat{b}_{t_0} = 0_{(4,1)},\ P_{t_0} = I_{(4,4)},\ \hat{\Sigma}_{t_0} = I_{(2,2)}.$$

Table 1 presents the estimates $\hat{b}_t$ of $(\varphi_{11}^{(1)}, \varphi_{12}^{(1)}, \varphi_{21}^{(1)}, \varphi_{22}^{(1)})'$ for chosen time periods t in a simulation of (4.1), (4.2).

Example 2. The transformed model (2.18) has in our case the form

$$\Phi_0 y_t = \Phi_1 y_{t-1} + \varepsilon_t$$

with

$$\Phi_0 = \begin{pmatrix} 1 & 0 \\ 1 & 1 \end{pmatrix},\ \Phi_1 = \begin{pmatrix} 0.5 & 0.1 \\ 0.3 & 0.9 \end{pmatrix},\ \varepsilon_t \sim \text{iid } N_2(0,\Sigma),\ \Sigma = \begin{pmatrix} 1 & 0 \\ 0 & 4 \end{pmatrix}.$$

The adaptive formulas (2.22) - (2.24) from Section 2.1 were applied producing the estimates $\hat{b}_{2t+1}$ of $(\varphi_{11}^{(1)}, \varphi_{12}^{(1)}, 0,0,0)'$ and $\hat{b}_{2t+2}$ of $(0,0,\varphi_{21}^{(o)}, \varphi_{21}^{(1)}, \varphi_{22}^{(1)})'$ with the initial values

$$t_o=5,\ \hat{b}_{2t_o+1}=\hat{b}_{2t_o+2}=0_{(5,1)},\ P_{2t_o+1}=P_{2t_o+2}=I_{(5,5)},\ \hat{\sigma}^2_{2t_o+1}=\hat{\sigma}^2_{2t_o+2}=1.$$

Table 2 presents the estimates of $\hat{b}_{2t+1}$ and $\hat{b}_{2t+2}$ for chosen time periods t in a simulation of (4.1), (4.2).

Example 3. The adaptive estimation formulas (2.34) - (2.38) from Section 2.3 were applied in a simulation (4.1), (4.2) with $y_{1,t}$ missing for the time periods t=15,25,35,... and $y_{2,t}$ missing for the time periods t=10,20,30,..., i.e.

$$S_t=(0,1) \text{ for } t=15,25,35,\dots,$$
$$=(1,0) \text{ for } t=10,20,30,\dots .$$

The initial values were taken as

$$t_o=5,\ \hat{b}_{t_o}=0_{(4,1)},\ P_{t_o}=I_{(4,4)},\ \hat{\Sigma}_{t_o}=I_{(2,2)} .$$

Table 3 presents the estimates $\hat{b}_t$ of $(\varphi_{11}^{(1)}, \varphi_{12}^{(1)}, \varphi_{21}^{(1)}, \varphi_{22}^{(1)})'$ for chosen time periods t.

Example 4. The prediction procedure from Section 3.2 was applied in a simulation of (4.1), (4.2) with $y_{1,t}$ missing for the even time periods t and $y_{2,t}$ missing for the odd time periods t, i.e.

$$S_t=(0,1) \text{ for even } t,$$
$$=(1,0) \text{ for odd } t.$$

The initial values were taken as

$$\hat{x}_o^o=y_o\ ,\ P_o^o=\Sigma=\begin{pmatrix}1 & -1\\ -1 & 5\end{pmatrix}.$$

In Table 4 the actual values $y_t=(y_{1,t},y_{2,t})'$ are compared with the estimates $\hat{y}_t^t=(\hat{y}_{1,t}^t,\hat{y}_{2,t}^t)'$ for chosen time periods t. The missing components of the vectors y_t for the Kalmar filter are in parentheses.

Table 1. Simulation results for Example 1 (estimation in two-dimensional AR (1) process)

	$\varphi_{11}^{(1)}=0.5$	$\varphi_{12}^{(1)}=0.1$	$\varphi_{21}^{(1)}=-0.2$	$\varphi_{22}^{(1)}=0.8$
t	$\hat{\varphi}_{11,t}^{(1)}$	$\hat{\varphi}_{12,t}^{(1)}$	$\hat{\varphi}_{21,t}^{(1)}$	$\hat{\varphi}_{22,t}^{(1)}$
10	1.353	0.249	-1.588	-1.989
20	0.918	0.188	-0.871	0.588
30	0.850	0.183	-0.768	0.634
200	0.573	0.083	-0.353	0.784
210	0.590	0.086	-0.339	0.780
220	0.566	0.086	-0.316	0.780
230	0.577	0.087	-0.331	0.779

Table 2. Simulation results for Example 2 (estimation in two-dimensional AR (1) process with state space representation for particular components)

	$\varphi_{21}^{(o)}=1$	$\varphi_{11}^{(1)}=0.5$	$\varphi_{12}^{(1)}=0.1$	$\varphi_{21}^{(1)}=0.3$	$\varphi_{22}^{(1)}=0.9$
t	$\hat{\varphi}_{21,t}^{(o)}$	$\hat{\varphi}_{11,t}^{(1)}$	$\hat{\varphi}_{12,t}^{(1)}$	$\hat{\varphi}_{21,t}^{(1)}$	$\hat{\varphi}_{22,t}^{(1)}$
10	0.804	0.910	0.141	0.142	1.002
25	1.029	0.750	0.108	0.551	0.962
50	0.892	0.506	0.047	0.358	0.927
100	0.860	0.477	0.045	0.307	0.871
150	0.950	0.342	0.068	0.474	0.905
200	0.994	0.416	0.087	0.466	0.868

Table 3. Simulation results for Example 3 (estimation in two-dimensional AR (1) process with missing observations)

	$\varphi_{11}^{(1)}=0.5$	$\varphi_{12}^{(1)}=0.1$	$\varphi_{21}^{(1)}=-0.2$	$\varphi_{22}^{(1)}=0.8$
t	$\hat{\varphi}_{11,t}^{(1)}$	$\hat{\varphi}_{12,t}^{(1)}$	$\hat{\varphi}_{21,t}^{(1)}$	$\hat{\varphi}_{22,t}^{(1)}$
10	0.291	0.233	-0.195	0.647
25	0.462	0.302	-0.252	0.771
50	0.417	0.294	-0.154	0.765
100	0.383	0.290	-0.193	0.720

150	0.438	0.319	-0.182	0.701
200	0.399	0.307	-0.127	0.705

Table 4. Simulation results for Example 4 (prediction in two-dimensional AR (1) process with missing observations).

t	$y_{1,t}$	$y_{2,t}$	$\hat{y}_{1,t}$	$\hat{y}_{2,t}$
10	(2.097)	-3.485	0.710	-3.485
19	-0.729	(3.818)	-0.729	2.723
50	(-2.923)	3.324	-0.086	3.324
79	1.456	(-3.608)	1.456	-2.825
90	(1.207)	-5.225	1.091	-5.225
99	-1.245	(-3.254)	-1.245	-2.744
120	(-0.347)	5.472	-0.730	5.472
149	-0.769	(-4.793)	-0.769	-6.388
180	(0.503)	4.256	0.700	4.256

References

[1] AOKI M.: Notes on Economic Time Series Analysis: System Theoretic Perspectives, Springer, Berlin 1983.

[2] CIPRA T.: On improvement of prediction in ARMA processes, Math. Operationsforsch. Statist., Ser. Statistics 12(1981), 567-580.

[3] FAHRMEIR L.: Rekursive Algorithmen für Zeitreihemodelle, Vandenhoeck und Ruprecht, Göttingen 1981.

[4] GARDNER G., HARVEY A.C., PHILLIPS G.D.A.: An algorithm for exact likelihood estimation of autoregressive-moving average models by means of Kalman filtering, Applied Statistics 29(1980), 311-322.

[5] JAZWINSKI A.H.: Stochastic Processes and Filtering Theory, Academic Press, New York 1970.

[6] KALMAN R.E.: A new approach to linear filtering and prediction problems, Trans. ASME, Ser. D, J. Basic Eng. 82(1960), 35-45.

[7] KALMAN R.E., BUCY R.S.: New results in linear filtering and prediction theory, J. Basic Eng. 83(1961), 95-108.

[8] KOHN R.,ANSLEY C.F.: Estimation, prediction, and interpolation for ARIMA models with missing data, JASA 81(1986), 751-761.

[9] MOTYKOVÁ I.: Kalman Filter in Time Series, Diploma Work, Charles University, Prague 1987 (in Czech).

[10] PRIESTLEY M.B.: Spectral Analysis and Time Series (vol. 2: Multivariate series, Prediction and Control), Academic Press, London 1981.

[11] SCHNEIDER W.: Der Kalmanfilter als Instrument zur Diagnose und Schätzung variabler Parameter in ökonometrischen Modellen, Physica Verlag, Heidelberg 1986.

[12] SHEA B.L.: Maximum likelihood estimation of multivariate ARMA processes via the Kalman filter, in Time Series Analysis (O.D. Anderson ed.), Elsevier Science Publishers, Amsterdam 1984, 91-101.

Dept. of Statistics, Charles University, Sokolovská 83, 18600 Praha 8, Czechoslovakia

(Oblatum 4.5. 1987)