Algorithms for the Evaluation of Forecasts, Filters and Smoothers from a State-Space Model with the Feature of Time Dependent Dimension

by

© **Wen Teng**

A thesis submitted to the School of Graduate Studies in partial fulfillment of the requirements for the degree of Master of Science.

Department of Mathematics and Statistics

Memorial University

September 2016

St. John's, Newfoundland and Labrador, Canada

Abstract

This thesis illustrates two approaches for the evaluation of forecasting, filtering and smoothing from a flexible state-space model. Parameters of this model can be time dependent and the dimension of its state or observed vectors can vary over time. The first approach consists of establishing an algorithm based on the Kalman filter and Kalman smoother as well as properties derived from the model. Another approach is to reconstruct the model. In addition, an extension of the model is proposed.

Acknowledgements

I am very grateful to have the chance of studying at Memorial University and should thank many people for their help.

At first, I would like to express my sincere gratitude to my supervisor Dr. J Concepción Loredo-Osti for his patience, motivation and continuous support during my master's study. His guidance helped a lot in writing of this thesis.

Besides my supervisor, I also would like to thank the faculty and staff at Department of Mathematics and Statistics. My deep thanks go to Dr. Zhaozhi Fan for his excellent course, help and brilliant advices regarding my study and career. Taking courses of Dr. Alwell Oyet and Dr. Hong Wang is also important for me, it leaves me a wonderful memory. Many other faculty or staff give me support and encouragement in different ways, thank you all so much.

At last, I want to thank those friends met in St. John's for participation in my life.

Table of contents

Chapter 1

Introduction

This thesis would mainly explore the relationship among estimators like forecasts, filters and smoothers of a state-space model. We will propose a flexible state-space model, in which any subset of parameters can be time dependent. In addition, the dimension of the state and observed vectors of our model could vary over time, which is only assumed in limited literature, for instance, the article of Mclauchlan and Murray [1996]. The relationship among estimators could be well described by their joint distribution. Based on the algorithms - Kalman filter and Kalman smoother, we adopt two methods to obtain the complete information of the distribution. One method is established by researching on some properties derived from the model, the other one can be accomplished by reconstructing the model, which is a classic approach in the study of state-space model (see Anderson and Moore [2012]). Besides these, inspired by the work of Qian [2014], we extend our state-space model into a new one, and propose a procedure to calculate forecasts, filters and smoothers of the new model through the model reconstruction method.

The rest of the thesis is organized as follows. Chapter 2 provides an introduction to our state-space model and preliminary concepts associated with it as well as a brief introduction to Kalman filter and Kalman smoother. The organization and content of the second chapter is referred to the books of Shumway and Stoffer [2010] as well as Hamilton [1994], however we should notice that our model differs from theirs due to the time-varying dimension of the state and observed vectors. Chapter 3 illustrates the main work of this thesis. In the first section of chapter 3, an algorithm has been established to assist us in studying the joint distribution, and in the second section of it, we examine the model reconstruction method. we also extend the model in the third section. Chapter 4 gives a summary and indicate our further work.

Chapter 2

Preliminary knowledge of the state-space model

Starting with the breakthrough papers of Kalman [1960] as well as Kalman and Bucy [1961], the state-space model has been widely applied in many fields such as statistics, economics, engineering and medicine. Harvey [1990], Hamilton [1994], Tsay [2010], Durbin and Koopman [2012] and Tsay [2014] present its theory and applications in time series analysis. The article of Basdevant [2003] contains an application on macroeconomics, Mergner [2009] addresses applications in the area of finance and Jones [1984] provides an example of applications on health science. In order to solve practical problems, building a good state-space model is essential. We should know how to cast a structural model into an appropriate state-space form. The representation is not unique, for one can enlarge the state vector but describe the same process.

The primary and standard tools for analyzing a state-space model are filtering and smoothing algorithms. Filtering provides an estimate of the state vector at a given time point conditional on the information observed up to that point. Smoothing enables us to estimate the state vector at any time given all the available observations.

The best known filtering and smoothing algorithms are Kalman filter and Kalman smoother, that could be found in a large amount of literature, for example, books of Shumway and Stoffer [2010], Anderson and Moore [2012], Brockwell and Davis [2013].

In the first section of this chapter, we describe a state-space model which is the focus of this thesis, and give an instance to illustrate the state-space representation. In the second section, we would discuss the forecasting, filtering and smoothing problems and introduce the solution - the Kalman filter and Kalman smoother algorithms.

2.1 The state-space model and representation

Many linear dynamic systems can be written in a state-space form. Before exploring the state-space representation, we consider a simple dynamic system: first-order autoregression

$$z_t = \phi z_{t-1} + \varepsilon_t, \tag{2.1}$$

where $\varepsilon_t \sim$ i.i.d. $N(0, \sigma^2)$. Since the future values $\{z_{t+1},\ z_{t+2},\ \cdots\}$ of this process only depend on the present value z_t, we can easily to analyze the dynamics and make forecasts for this process. For example, we solve equation (2.1) by recursive substitution:

$$\begin{aligned} z_{t+m} &= \phi z_{t+m-1} + \varepsilon_{t+m} \\ &= \phi(\phi z_{t+m-2} + \varepsilon_{t+m-1}) + \varepsilon_{t+m} \\ &= \phi^m z_t + \phi^{m-1}\varepsilon_{t+1} + \phi^{m-2}\varepsilon_{t+2} + \cdots + \phi\varepsilon_{t+m-1} + \varepsilon_{t+m} \end{aligned} \tag{2.2}$$

for $m = 1, 2, \cdots$, which implies an optimal (in a mean square error sense) m-ahead forecast

$$E(z_{t+m}|z_t, z_{t-1}, \cdots) = \phi^m z_t.$$

In order to represent a more complicated linear dynamic system by the state-space model, we should derive the dynamics of the system from an $(n_t \times 1)$ observation vector y_t. The system dynamics is expressed by the change of a possibly unobserved $(r_t \times 1)$ vector x_t known as the state vector of the system. The dynamics of a linear system can be illustrated as a generalization of (2.1):

$$x_t = g_t + F_t x_{t-1} + v_t, \tag{2.3}$$

where F_t denotes an $(r_t \times r_{t-1})$ matrix, g_t is an $(r_t \times 1)$ predetermined vector, the $(r_t \times 1)$ vector v_t is taken to be $N(0, Q_t)$ and $\{v_t\}_{t=1}^{\infty}$ is an independent sequence. State vector x_0 is assumed to be the initial value of the state, furthermore, we assume that it is $N(\gamma, O)$ and independent of $\{v_t\}_{t=1}^{\infty}$. We usually name (2.3) the state equation.

Note that 0 in the last paragraph symbolizes a zero matrix fitted its position, for instance, population mean 0 for vector v_t is an $(r_t \times 1)$ zero matrix. In the following context, symbol 0 would be used like this way for convenience. Similarly, Q_t and O in the last paragraph are covariance matrices fitted their positions.

Like (2.2), we can write

$$\begin{aligned} x_{t+m} =& \tilde{F}^{m}_{t+m} x_t + \tilde{F}^{m-1}_{t+m} g_{t+1} + \tilde{F}^{m-2}_{t+m} g_{t+2} + \cdots + \tilde{F}^{1}_{t+m} g_{t+m-1} + g_{t+m} \\ &+ \tilde{F}^{m-1}_{t+m} v_{t+1} + \tilde{F}^{m-2}_{t+m} v_{t+2} + \cdots + \tilde{F}^{1}_{t+m} v_{t+m-1} + v_{t+m} \end{aligned} \tag{2.4}$$

for $m = 1, 2, \cdots$, where

$$\tilde{F}^{n}_{t+m} = F_{t+m} \times F_{t+m-1} \times \cdots \times F_{t+m-(n-1)}$$

for $n = 1, 2, \cdots, m$. Thus, the optimal m-ahead forecast can be written as

$$
\begin{aligned}
&E(x_{t+m}|x_t, x_{t-1}, \cdots) \\
&= \tilde{F}_{t+m}^{m} x_t + \tilde{F}_{t+m}^{m-1} g_{t+1} + \tilde{F}_{t+m}^{m-2} g_{t+2} + \cdots + \tilde{F}_{t+m}^{1} g_{t+m-1} + g_{t+m}.
\end{aligned} \tag{2.5}
$$

Assume that the observation vectors are related to the state vectors through the equation

$$
y_t = a_t + H_t x_t + w_t, \tag{2.6}
$$

where y_t is an $(n_t \times 1)$ vector representing the observation of the system at time t, H_t is an $(n_t \times r_t)$ matrix of coefficients, and w_t is an $(n_t \times 1)$ vector which could be thought as measurement error; w_t is presumed to be $N(0, R_t)$ (R_t is covariance matrix fitted its position), $\{w_t\}_{t=1}^{\infty}$ is an independent sequence and also independent of $\{v_t\}_{t=1}^{\infty}$ as well as the initial value of state x_0. Equation (2.6) also includes a_t, an $(n_t \times 1)$ observed or predetermined vector. For example, a_t could include the information of lagged values of y. We usually call (2.6) the observation equation of the system.

The state equation (2.3) and observation equation (2.6) constitute a state-space representation for the dynamic system of y. In this paper, we would focus on the system described by (2.3) and (2.6).

Note that the dimension of both state and observed vectors of this model can change over time. The time-varying dimension of observed vectors has been well understood and practiced. For instance, if some elements of an observation have missed, then its size is accordingly reduced. If the observation itself has missed, the updating step of the Kalman filter would be skipped (see Jones [1980], Harvey and Pierse [1984]). However the time-varying dimension of state vectors had not been enough appreciated until recently. Jungbacker et al. [2011] put common factors and idiosyncratic disturbances corresponding to missing data in the state vector when consider a factor model with missing data, then the state vector varies in dimension over time due to the variation in the amount of missing data. In our model, the dimension of state vectors is allowed to vary over time, not only for the generality of the model but also for the convenience to introduce the model reconstruction method in next chapter.

Because a_t is deterministic, the state vector x_t and measurement error w_t contain everything in the past which is relevant for the future values of y,

$$\begin{aligned} & E(y_{t+m}|x_t, x_{t-1}, \cdots, y_t, y_{t-1}, \cdots) \\ & = E(a_{t+m} + H_t x_{t+m} + w_{t+m}|x_t, x_{t-1}, \cdots, y_t, y_{t-1}, \cdots) \\ & = a_{t+m} + H_t E(x_{t+m}|x_t, x_{t-1}, \cdots, y_t, y_{t-1}, \cdots) \\ & = a_{t+m} + H_t E(x^{t+m}|x_t, x_{t-1}, \cdots), \end{aligned} \tag{2.7}$$

where $E(x^{t+m}|x_t, x_{t-1}, \cdots)$ can be obtained from (2.5)

We take a pth order autoregression as a simple example of a system which can be written in state-space form,

$$y_t - \mu = \phi_1(y_{t-1} - \mu) + \phi_2(y_{t-2} - \mu) + \cdots + \phi_p(y_{t-p} - \mu) + \varepsilon_t. \tag{2.8}$$

Then we write (2.8) as

$$\begin{bmatrix} y_t - \mu \\ y_{t-1} - \mu \\ \vdots \\ y_{t-p+1} - \mu \end{bmatrix} = \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} y_{t-1} - \mu \\ y_{t-2} - \mu \\ \vdots \\ y_{t-p} - \mu \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}. \tag{2.9}$$

The first row of (2.9) indicates (2.8) and other rows simply state the identity. We use the following notations:

$$x_t \stackrel{\text{def}}{=} (y_t - \mu, y_{t-1} - \mu, \cdots, y_{t-p+1} - \mu)',$$

$$v_t \stackrel{\text{def}}{=} (\varepsilon_t, 0, \cdots, 0)',$$

$$F \stackrel{\text{def}}{=} \begin{bmatrix} \phi_1 & \phi_2 & \cdots & \phi_{p-1} & \phi_p \\ 1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 \end{bmatrix},$$

$$H \stackrel{\text{def}}{=} (1, 0, \cdots, 0)'.$$

Hence equation (2.9) can be written as

$$x_t = Fx_{t-1} + v_t,$$

which is the state equation. The observation equation is

$$y_t = \mu + Hx_t.$$

2.2 Forecasting, filtering and smoothing

The relevant state-space model is taken as previously mentioned:

- state equation

$$x_t = g_t + F_t x_{t-1} + v_t$$
$$x_0 \sim N(\gamma, O) \tag{2.10}$$
$$v_t \sim N(0, Q_t),$$

- observation equation

$$y_t = a_t + H_t x_t + w_t$$
$$w_t \sim N(0, R_t), \tag{2.11}$$

for $t = 1, 2, \cdots$. Note that all the elements of g_t, a_t, F_t, H_t, γ, O, Q_t and R_t are known with certainty, and $\{v_t\}$, $\{w_t\}$ as well as x_0 are mutually independent.

In practice, the main aim for the analysis of a state-space model is to provide estimators for the underlying unobserved state x_t on the basis of observations $Y_s = \{y_s, y_{s-1}, \cdots, y_1\}$. When $s < t$, the problem is called forecasting; when $s = t$, it is called filtering; when $s > t$, it is called smoothing. Besides the estimators, we also want to measure their precision. The solution to these problems can be accomplished by Kalman filter and Kalman smoother.

For convenience, we will use the following notations:

$$x_t^s \stackrel{\text{def}}{=} E(x_t | Y_s),$$

$$P_{t_1,t_2}^{s_1,s_2} \stackrel{\text{def}}{=} E\{(x_{t_1} - x_{t_1}^{s_1})(x_{t_2} - x_{t_2}^{s_2})'\}$$

and

$$P_{t_1,t_2}^{s} \stackrel{\text{def}}{=} P_{t_1,t_2}^{s,s}, \quad P_t^{s_1,s_2} \stackrel{\text{def}}{=} P_{t,t}^{s_1,s_2}, \quad P_t^s \stackrel{\text{def}}{=} P_{t,t}^{s,s}.$$

Because $x_t - x_t^s$ and any vector from Y_s are uncorrelated, by the assumption of normality, $x_t - x_t^s$ is independent from Y_s, which implies

$$E\{(x_{t_1} - x_{t_1}^s)(x_{t_2} - x_{t_2}^s)'|Y_s\} = E\{(x_{t_1} - x_{t_1}^s)(x_{t_2} - x_{t_2}^s)'\} = P_{t_1,t_2}^s.$$

Before introducing Kalman filter and Kalman smoother, we should notice that, for our model, the dimension of state vectors is time-varying, which is different from the usual models normally with fixed size of state vectors. However this feature of our model does not affect the proof of following algorithms at all. Therefore Kalman filter and smoother do work as expected for the model.

2.2.1 Kalman filter

First, we present the Kalman filter. Kalman filter can be described as a recursive algorithm for calculating the one-ahead forecast and the filter of x_t through the information observed. The advantage of Kalman filter is that it specifies how to update the filter from x_{t-1}^{t-1} to x_t^t via the observation y_t, without having to reprocess the entire observations $\{y_t, \cdots, y_1\}$.

Now we begin to present Kalman filter. The algorithm is started by setting the initial condition:

$$x_0^0 = \gamma, \quad P_0^0 = O,$$

where γ and O are the mean and variance of the distribution of x_0 respectively. Then next step is to calculate the forecast of x_t conditional on Y_{t-1}. We have

$$x_t^{t-1} = g_t + F_t x_{t-1}^{t-1}, \tag{2.12}$$

and

$$P_t^{t-1} = F_t P_{t-1}^{t-1} F_t' + Q_t. \tag{2.13}$$

Therefore x_t^{t-1}, P_t^{t-1} could be obtained from x_{t-1}^{t-1}, P_{t-1}^{t-1} via (2.12) and (2.13). The final step is to get the filter of x_t conditional on Y_t. We have

$$x_t^t = x_t^{t-1} + K_t(y_t - a_t - H_t x_t^{t-1}), \tag{2.14}$$

where the Kalman gain, K_t, is defined as

$$K_t \stackrel{\text{def}}{=} P_t^{t-1} H_t' (H_t P_t^{t-1} H_t' + R_t)^{-1}.$$

P_t^t can be obtained as

$$P_t^t = (I_{r_t} - K_t H_t) P_t^{t-1}, \tag{2.15}$$

where I_{r_t} represents an $(r_t \times r_t)$ identity matrix. Therefore we obtain x_t^t, P_t^t from x_t^{t-1}, P_t^{t-1} through (2.14) and (2.15).

To summarize, the Kalman filter is an recursive algorithm that could be stated as

- Initial condition:

$$x_0^0 = \gamma, \quad P_0^0 = O.$$

- Forecast equation:

$$x_t^{t-1} = g_t + F_t x_{t-1}^{t-1},$$

$$P_t^{t-1} = F_t P_{t-1}^{t-1} F_t' + Q_t.$$

- Filter equation:

$$x_t^t = x_t^{t-1} + K_t(y_t - a_t - H_t x_t^{t-1}),$$

$$P_t^t = (I_{r_t} - K_t H_t) P_t^{t-1}.$$

$$(K_t = P_t^{t-1} H_t' (H_t P_t^{t-1} H_t' + R_t)^{-1})$$

We should notice that some or all parameters of the model could vary with time and the state or observation dimension could change with time as well. The variance P_t^{t-1} and P_t^t are not functions of the data and could be evaluated without calculating the forecast x_t^{t-1} and filter x_t^t.

We have had the one-ahead forecast from forecast equation of Kalman filter, an m-ahead forecast can be calculated by (2.4):

$$\begin{aligned} x_{t+m}^t &= E(x_{t+m}|Y_t) \\ &= \tilde{F}_{t+m}^m x_t^t + \tilde{F}_{t+m}^{m-1} g_{t+1} + \tilde{F}_{t+m}^{m-2} g_{t+2} + \cdots + \tilde{F}_{t+m}^1 g_{t+m-1} + g_{t+m}. \end{aligned} \tag{2.16}$$

Thus the error of this forecast can be obtained from (2.4) and (2.16),

$$x_{t+m} - x_{t+m}^t = \tilde{F}_{t+m}^m (x_t - x_t^t) + \tilde{F}_{t+m}^{m-1} v_{t+1} + \tilde{F}_{t+m}^{m-2} v_{t+2} + \cdots + \tilde{F}_{t+m}^1 v_{t+m-1} + v_{t+m},$$

where it follows that the mean squared error of the forecast, P_{t+m}^t, is

$$\begin{aligned} P_{t+m}^t &= E\{(x_{t+m} - x_{t+m}^t)(x_{t+m} - x_{t+m}^t)'\} \\ &= \tilde{F}_{t+m}^m P_t^t (\tilde{F}_{t+m}^m)' + \tilde{F}_{t+m}^{m-1} Q_{t+1} (\tilde{F}_{t+m}^{m-1})' + \cdots + \tilde{F}_{t+m}^1 Q_{t+m-1} (\tilde{F}_{t+m}^1)' + Q_{t+m}. \end{aligned} \tag{2.17}$$

2.2.2 Kalman smoother

Up to this point we have been concerned with the forecast and filter of the state vector, however in some applications the value of the state is of interest in its own right. It is desirable to use the information through the end of the sample to conduct the inference about the past values of the state. Such an inference is known as a smoothed estimate, the estimator is x_t^T for $t = 1, 2, \cdots, T-1$, and the corresponding mean squared error is P_t^T, where T denote the time for the last observation.

The smoothed estimates can be calculated by Kalman smoother, which is a recursive algorithm for obtaining the smoother x_t^T and its mean squared error P_t^T. Now we present the Kalman smoother as follows. First, we run the observed data through Kalman filter to obtain $\{P_t^{t-1}\}_{t=1}^T$, $\{P_t^t\}_{t=1}^T$ from (2.13), (2.15) respectively, and accordingly obtain $\{x_t^{t-1}\}_{t=1}^T$, $\{x_t^t\}_{t=1}^T$ from (2.12), (2.14). We set x_T^T, P_T^T as the initial value of this algorithm, therefore the sequence of smoothed estimates $\{x_t^T\}_{t=1}^T$ can be calculated in reverse order by iterating on

$$x_{t-1}^T = x_{t-1}^{t-1} + J_{t-1}(x_t^T - x_t^{t-1}), \tag{2.18}$$

where

$$J_{t-1} \overset{\text{def}}{=} P_{t-1}^{t-1} F' (P_t^{t-1})^{-1},$$

for $t = T-1, T-2, \cdots, 1$. The corresponding mean squared errors are similarly found by iterating on

$$P_{t-1}^T = P_{t-1}^{t-1} + J_{t-1}(P_t^T - P_t^{t-1}) J_{t-1}', \tag{2.19}$$

in reverse order for $t = T-1, T-2, \cdots, 1$.

To summarize, the Kalman smoother is a recursive algorithm which could be stated as

- Initial conditions:

  run Kalman filter to obtain

$$\{P_t^{t-1}\}_{t=1}^T, \quad \{P_t^t\}_{t=1}^T, \quad \{x_t^{t-1}\}_{t=1}^T, \quad \{x_t^t\}_{t=1}^T.$$

- Smoother equation:

  run equations below in reverse order

$$x_{t-1}^T = x_{t-1}^{t-1} + J_{t-1}(x_t^T - x_t^{t-1}),$$

$$P_{t-1}^T = P_{t-1}^{t-1} + J_{t-1}(P_t^T - P_t^{t-1})J_{t-1}'.$$

$$(J_{t-1} = P_{t-1}^{t-1}F'(P_t^{t-1})^{-1})$$

Chapter 3

Discussion of the joint distribution of estimators

As we discussed in the last chapter, from the algorithm named Kalman filter, we can obtain the filter x_t^t, m-ahead forecast x_{t+m}^t (by (2.16)) and corresponding mean squared errors P_t^t, P_{t+m}^t (by (2.17)). Furthermore, we could also obtain the smoother x_{t-m}^t ($m = 1, 2, \cdots, t-1$) and the corresponding mean squared error P_{t-m}^t from the algorithm called Kalman smoother. Up to this point, we have a procedure to compute these estimators and their mean squared errors, then naturally the next aim is to obtain the joint distribution of these estimators based on the given data or observations Y_T, which provides an evaluation of these estimators.

Based on the model which consists of (2.10) and (2.11), It is easy to notice that vectors $\{x_{T+m}, \cdots, x_1, y_T, \cdots, y_1\}$ have a joint normal distribution, which means that the joint conditional distribution of $\{x_{T+m}, \cdots, x_1\}$ given Y_T is normal. We can illustrate the result as:

$$\begin{bmatrix} x_{T+m} \\ \vdots \\ x_T \\ \vdots \\ x_1 \end{bmatrix} |Y_T \sim N(\begin{bmatrix} x_{T+m}^T \\ \vdots \\ x_T^T \\ \vdots \\ x_1^T \end{bmatrix}, \begin{bmatrix} P_{T+m}^T & \cdots & P_{T+m,T}^T & \cdots & P_{T+m,1}^T \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ P_{T,T+m}^T & \cdots & P_T^T & \cdots & P_{T,1}^T \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ P_{1,T+m}^T & \cdots & P_{1,T}^T & \cdots & P_1^T \end{bmatrix}).$$

We already know the method to obtain $x_{T+m}^T, \cdots, x_T^T, \cdots, x_1^T$ and $P_{T+m}^T, \cdots, P_T^T, \cdots, P_1^T$, thus next step is to develop an approach to obtain the conditional covariances $P_{s,t}^T$ for $s, t \in \{1, 2, \cdots, T+m\}$ and $s \neq t$. We could work out this problem by focusing on matrix

$$\begin{bmatrix} P_{T+m}^T & \cdots & P_{T+m,T}^T & \cdots & P_{T+m,1}^T \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ P_{T,T+m}^T & \cdots & P_T^T & \cdots & P_{T,1}^T \\ \vdots & \cdots & \vdots & \cdots & \vdots \\ P_{1,T+m}^T & \cdots & P_{1,T}^T & \cdots & P_1^T \end{bmatrix}. \tag{3.1}$$

We consider two approaches to address this problem. In the first section of this chapter, we propose an algorithm to calculate the related conditional covariances. Since this algorithm is recursive, we call it recursive method. On the other hand, in the second section, we will reconstruct the state-space model to propose another way to solve the problem. We call it model reconstruction method, which could be found in the book of Anderson and Moore [2012] (chapter 7). Besides giving a solution to this problem, by the model reconstruction method, we would also extend our state-space model into a relatively more general one and try to develop the corresponding method to do filtering, forecasting, smoothing of the new model. This extension would be introduced in the third section.

3.1 Recursive method

Based on the original model ((2.10) and (2.11)), we develop an algorithm to obtain the conditional covariance

$$P_{a,b}^{s,t} = E\{(x_a - x_a^s)(x_b - x_b^t)'\}.$$

Before deriving the algorithm, we should introduce a lemma.

Lemma. *For the model consisting of (2.10) and (2.11), the following equations hold:*

1. $(P_{a,b}^{s,t})' = P_{b,a}^{t,s}$.

2. $P_{a,b}^{s,t} = P_{a,b}^{t} = P_{a,b}^{t,s}$, *if* $s \le t$.

3. $P_{a,b}^{t} = F_a P_{a-1,b}^{t}$, *if* $a > \max(b,t)$.

4. $P_b^t = F_b P_{b-1}^t F_b' + Q_b$, *if* $b > t$.

5. $P_{t,b}^{t} = (I_{r_t} - K_t H_t) P_{t,b}^{t-1}$, *if* $t > b$.

6. $P_{a,b}^{t} = P_{a,b}^{a} + J_a(P_{a+1,b}^{b} - P_{a+1,b}^{a}) + J_a(P_{a+1,b+1}^{t} - P_{a+1,b+1}^{b})J_b'$, *if* $t > a > b$.

Proof.

1. Based on the definition of $P_{a,b}^{s,t}$, we have

$$(P_{a,b}^{s,t})' = E\{(x_a - x_a^s)(x_b - x_b^t)'\}' = E\{(x_b - x_b^t)(x_a - x_a^s)'\} = P_{b,a}^{t,s}.$$

2. Denote $\mathcal{G} \stackrel{\text{def}}{=} \sigma(Y_t)$, the σ-algebra generated by $\{y_t, y_{t-1}, \cdots, y_1\}$. If vector

function g is $\mathcal{G}$-measurable, thus

$$\begin{aligned} E\{(x_a - x_a^t)g'\} &= E(x_a g') - E(x_a^t g') \\ &= E\{E(x_a g'|Y_t)\} - E(x_a^t g') \\ &= E(x_a^t g') - E(x_a^t g') \\ &= 0. \end{aligned}$$

Note that x_b^t and x_b^s $(s \leq t)$ are $\mathcal{G}$-measurable, thus we have

$$E\{(x_a - x_a^t)(x_b^t)'\} = 0 \text{ and } E\{(x_a - x_a^t)(x_b^s)'\} = 0,$$

then

$$\begin{aligned} P_{a,b}^t &= E\{(x_a - x_a^t)(x_b - x_b^t)'\} \\ &= E\{(x_a - x_a^t)x_b'\} \\ &= E\{(x_a - x_a^t)(x_b - x_b^s)'\} \\ &= P_{a,b}^{t,s}. \end{aligned}$$

By the property 1 of this lemma, we have

$$P_{b,a}^{t,s} = P_{b,a}^t \;\Rightarrow\; P_{a,b}^{s,t} = P_{a,b}^t.$$

Therefore when $s \leq t$, $P_{a,b}^{s,t} = P_{a,b}^t = P_{a,b}^{t,s}$ holds.

3. Since $a > t$, by (2.10), we have

$$x_a^t = g_a + F_a x_{a-1}^t,$$

hence

$$x_a - x_a^t = g_a + F_a x_{a-1} + v_a - (g_a + F_a x_{a-1}^t)$$
$$= F_a(x_{a-1} - x_{a-1}^t) + v_a.$$

Therefore

$$P_{a,b}^t = E\{(x_a - x_a^t)(x_b - x_b^t)'\}$$
$$= E\{[F_a(x_{a-1} - x_{a-1}^t) + v_a](x_b - x_b^t)'\}$$
$$= F_a E\{(x_{a-1} - x_{a-1}^t)(x_b - x_b^t)'\}$$
$$= F_a P_{a-1,b}^t,$$

where the third equation holds as $E\{(v_a)(x_b - x_b^t)'\} = 0$ when $a > \max(b, t)$.

4. We already know that

$$x_b - x_b^t = F_b(x_{b-1} - x_{b-1}^t) + v_b,$$

thus

$$P_b^t = E\{(x_b - x_b^t)(x_b - x_b^t)'\}$$
$$= E\{[F_b(x_{b-1} - x_{b-1}^t) + v_b][F_b(x_{b-1} - x_{b-1}^t) + v_b]'\}$$
$$= E\{[F_b(x_{b-1} - x_{b-1}^t)(x_{b-1} - x_{b-1}^t)' F_b']\} + E(v_b v_b')$$
$$= F_b P_{b-1}^t F_b' + Q_b,$$

where the third equation holds since

$$E\{[F_b(x_{b-1} - x_{b-1}^t)]v_b'\} = 0 \text{ and } E\{v_b[F_b(x_{b-1} - x_{b-1}^t)]'\} = 0$$

when $b > t$.

5. From the filter equation (2.14) of Kalman filter, we obtain

$$\begin{aligned} x_t^t &= x_t^{t-1} + K_t(y_t - a_t - H_t x_t^{t-1}) \\ &= x_t^{t-1} + K_t[(a_t + H_t x_t + w_t) - a_t - H_t x_t^{t-1}] \\ &= x_t^{t-1} + K_t[H_t(x_t - x_t^{t-1}) + w_t] \\ &= x_t^{t-1} + K_t H_t(x_t - x_t^{t-1}) + K_t w_t. \end{aligned}$$

Thus we have

$$\begin{aligned} P_{t,b}^t &= P_{t,b}^{t,b} \\ &= E\{(x_t - x_t^t)(x_b - x_b^b)'\} \\ &= E\{(x_t - x_t^{t-1} - K_t H_t(x_t - x_t^{t-1}) - K_t w_t)(x_b - x_b^b)'\} \\ &= E\{[(I_{r_t} - K_t H_t)(x_t - x_t^{t-1}) - K_t w_t](x_b - x_b^b)'\} \\ &= (I_{r_t} - K_t H_t)E\{(x_t - x_t^{t-1})(x_b - x_b^b)'\} \\ &= (I_{r_t} - K_t H_t)P_{t,b}^{t-1,b} \\ &= (I_{r_t} - K_t H_t)P_{t,b}^{t-1}, \end{aligned}$$

where the first and last equations hold because of property 2 and the fact $t > b$, and the fifth equation holds because

$$K_t E\{w_t(x_b - x_b^b)'\} = 0,$$

when $t > b$.

6. From (2.18) of Kalman smoother, we have

$$x_a^t = x_a^a + J_a(x_{a+1}^t - x_{a+1}^a),$$

thus

$$x_a - x_a^t = x_a - x_a^a - J_a(x_{a+1}^t - x_{a+1}^a),$$

then

$$x_a - x_a^t + J_a x_{a+1}^t = x_a - x_a^a + J_a x_{a+1}^a. \qquad (3.2)$$

Similarly, the following holds:

$$x_b - x_b^t + J_b x_{b+1}^t = x_b - x_b^b + J_b x_{b+1}^b. \qquad (3.3)$$

Next, multiply the left side of (3.2) by the transpose of the left hand side of (3.3), and equate this to the corresponding result of the right hand sides of (3.2) and (3.3). Then taking expectation of both sides, we arrive to

$$\begin{aligned} &P_{a,b}^t + E\{(x_a - x_a^t)(x_{b+1}^t)'\}J_b' + J_a E\{x_{a+1}^t(x_b - x_b^t)'\} + J_a E\{x_{a+1}^t(x_{b+1}^t)'\}J_b' \\ &= P_{a,b}^{a,b} + E\{(x_a - x_a^a)(x_{b+1}^b)'\}J_b' + J_a E\{x_{a+1}^a(x_b - x_b^b)'\} + J_a E\{x_{a+1}^a(x_{b+1}^b)'\}J_b'. \end{aligned} \qquad (3.4)$$

Here vectors x_{b+1}^t and x_{a+1}^t are $\sigma(Y_t)$-measurable, then

$$E\{(x_a - x_a^t)(x_{b+1}^t)'\} = 0 \text{ and } E\{x_{a+1}^t(x_b - x_b^t)'\} = 0. \qquad (3.5)$$

Since $b < a \;\Rightarrow\; \sigma(Y_b) \subseteq \sigma(Y_a)$ and x_{b+1}^b is $\sigma(Y_b)$-measurable, x_{b+1}^b should be

$\sigma(Y_a)$-measurable, which implies

$$E\{(x_a - x_a^a)(x_{b+1}^b)'\} = 0. \tag{3.6}$$

We have

$$\begin{aligned}
P_{a+1,b+1}^t =& E\{(x_{a+1} - x_{a+1}^t)(x_{b+1} - x_{b+1}^t)'\} \\
=& E(x_{a+1}x_{b+1}') - E\{E[x_{a+1}(x_{b+1}^t)'|Y_t]\} - E\{E(x_{a+1}^t x_{b+1}'|Y_t)\} \\
& + E\{x_{a+1}^t(x_{b+1}^t)'\} \\
=& E(x_{a+1}x_{b+1}') - E\{x_{a+1}^t(x_{b+1}^t)'\},
\end{aligned}$$

thus

$$E\{x_{a+1}^t(x_{b+1}^t)'\} = E(x_{a+1}x_{b+1}') - P_{a+1,b+1}^t. \tag{3.7}$$

Because $b < a \Rightarrow \sigma(Y_b) \subseteq \sigma(Y_a)$, we have

$$\begin{aligned}
E(x_{a+1}^a|Y_b) &= E\{E(x_{a+1}|Y_a)|Y_b\} \\
&= E(x_{a+1}|Y_b) \\
&= x_{a+1}^b,
\end{aligned}$$

therefore

$$\begin{aligned}
E\{x_{a+1}^a(x_{b+1}^b)'\} &= E\{E(x_{a+1}^a(x_{b+1}^b)'|Y_b)\} \\
&= E\{E(x_{a+1}^a|Y_b)(x_{b+1}^b)'\} \\
&= E\{x_{a+1}^b(x_{b+1}^b)'\}
\end{aligned}$$

Then, as in the proof of (3.7), we could establish

$$E\{x_{a+1}^a(x_{b+1}^b)'\} = E(x_{a+1}x_{b+1}') - P_{a+1,b+1}^b. \tag{3.8}$$

We have

$$E\{x_{a+1}^a(x_b - x_b^b)'\} = E\{x_{a+1}(x_b - x_b^b)'\} - E\{(x_{a+1} - x_{a+1}^a)(x_b - x_b^b)'\}.$$

Here

$$E\{(x_{a+1} - x_{a+1}^a)(x_b - x_b^b)'\} = P_{a+1,b}^{a,b} = P_{a+1,b}^a,$$

where the last equation holds because of property 2 as well as $b < a$; and

$$E\{x_{a+1}(x_b - x_b^b)'\} = E\{(x_{a+1} - x_{a+1}^b)(x_b - x_b^b)'\} = P_{a+1,b}^b,$$

where the first equation holds as x_{a+1}^b is $\sigma(Y_b)$-measurable, which implies that

$$E\{x_{a+1}^b(x_b - x_b^b)'\} = 0.$$

Therefore

$$E\{x_{a+1}^a(x_b - x_b^b)'\} = P_{a+1,b}^b - P_{a+1,b}^a. \tag{3.9}$$

From (3.5), (3.6), (3.7), (3.8), (3.9) and $P_{a,b}^{a,b} = P_{a,b}^a$ (by property 2), (3.4) reduces to

$$P_{a,b}^t = P_{a,b}^a + J_a(P_{a+1,b}^b - P_{a+1,b}^a) + J_a(P_{a+1,b+1}^t - P_{a+1,b+1}^b)J_b'.$$

Up to this point, we have proven the lemma, then what we shall do next is to state the algorithm. This algorithm is in essence a recursive one like Kalman filter as well as Kalman smoother and it is established on the basis of these two algorithms. We could calculate any conditional covariance $P_{a,b}^{s,t}$ by this algorithm in terms of the model we introduced ((2.10) and (2.11)).

Algorithm.

step 1 For any conditional covariance $P^{s,t}_{a,b}$ of the model consisting of (2.10) and (2.11), assume i is a variable , let

$$\begin{cases} P^{s,t}_{a,b} = P^{s,t}_{a,b},\ i = 1 & \text{if } a \geq b, \\ P^{s,t}_{a,b} = (P^{s,t}_{a,b})',\ i = 0 & \text{otherwise.} \end{cases}$$

step 2 Due to the property 2, assume $n = \max(s,t)$, we have $P^{n}_{a,b} = P^{s,t}_{a,b}$.

step 3 Consider $P^{n}_{a,b}$ in several conditions (note that $a \geq b$ always holds due to step 1):

$$\begin{cases} \text{go to step 4.1} & \text{if } a > n, b = a, \\ \text{go to step 4.2} & \text{if } a > n, b < a, b > n, \\ \text{go to step 4.3} & \text{if } a > n, b = n, \\ \text{go to step 4.4} & \text{if } a > n, b < n, \\ \text{go to step 4.5} & \text{if } a = n, b = a \\ \text{go to step 4.6} & \text{if } a = n, b < a, \\ \text{go to step 4.7} & \text{if } a < n, b = a, \\ \text{go to step 4.8} & \text{if } a < n, b < a. \end{cases}$$

step 4.1 In this condition, $P^{n}_{a,b} = P^{n}_{a}$ $(a > n)$, by property 4 of lemma, this can be calculated by iterating on

$$P^{n}_{a} = F_a P^{n}_{a-1} F'_a + Q_a \quad (a > n) \tag{3.10}$$

until P_n^n, and P_n^n can be obtained by Kalman filter. After obtaining the value of $P_{a,b}^n$, go to step 5.

step 4.2 By property 3 of lemma, $P_{a,b}^n$ can be calculated by iterating on

$$P_{a,b}^n = F_a P_{a-1,b}^n \quad (a > \max(b, n)) \tag{3.11}$$

until P_b^n $(b > n)$ which can be calculated by iterating on (3.10) until P_n^n, then P_n^n can be obtained by Kalman filter. After obtaining the value of $P_{a,b}^n$, go to step 5.

step 4.3 For this condition, $P_{a,b}^n = P_{a,n}^n$ $(a > n)$, and it can be calculated by iterating on equation (3.11) until P_n^n, then P_n^n can be obtained by Kalman filter. After obtaining the value of $P_{a,b}^n$, go to step 5.

step 4.4 By property 3 and 5 of lemma, $P_{a,b}^n$ can be calculated by iterating on (3.11) until $P_{n,b}^n$, and $P_{n,b}^n$ can be calculated by iterating on

$$P_{n,b}^n = (I_n - K_n H_n) P_{n,b}^{n-1} = (I_n - K_n H_n) F_n P_{n-1,b}^{n-1} \quad (n > b) \tag{3.12}$$

until P_b^b which can be obtained by Kalman filter. After obtaining the value of $P_{a,b}^n$, go to step 5.

step 4.5 In this condition, $P_{a,b}^n = P_n^n$, it can be obtained by Kalman filter. After obtaining the value of $P_{a,b}^n$, go to step 5.

step 4.6 Under this condition, $P_{a,b}^n = P_{n,b}^n$ $(b < n)$, it can be calculated by iterating on (3.12) until P_b^b, and P_b^b can be obtained by Kalman filter. After obtaining the value of $P_{a,b}^n$, go to step 5.

step 4.7 For this condition, $P^n_{a,b} = P^n_a$ $(a < n)$, it can be obtained by Kalman smoother. After obtaining the value of $P^n_{a,b}$, go to step 5.

step 4.8 By property 6 of lemma, $P^n_{a,b}$ can be calculated by iterating on

$$P^n_{a,b} = P^a_{a,b} + J_a(P^b_{a+1,b} - P^a_{a+1,b}) + J_a(P^n_{a+1,b+1} - P^b_{a+1,b+1})J'_b \quad (n > a > b)$$

until $P^n_{n,n-a+b}$, where $P^a_{a,b}$ can be calculated by the procedure of step 4.6, $P^b_{a+1,b}$ can be calculated by the procedure of step 4.3, $P^a_{a+1,b}$ can be calculated by the procedure of step 4.4 and $P^b_{a+1,b+1}$ can be calculated by the procedure of step 4.2. As for $P^n_{n,n-a+b}$, since $n-a+b < n$, it could be calculated by the procedure of step 4.6. After obtaining the value of $P^n_{a,b}$, go to step 5.

step 5 Considering the procedure of step 1, by the property 1 of lemma, we have

$$\begin{cases} P^{s,t}_{a,b} = P^n_{a,b} & \text{if } i = 1, \\ P^{s,t}_{a,b} = (P^n_{a,b})' & \text{if } i = 0. \end{cases}$$

Then the calculation of $P^{s,t}_{a,b}$ is done.

By this algorithm, we could calculate any conditional covariance of the given state-space model ((2.10) and (2.11)), which implies that we can compute the value of every parameter of matrix (3.1), thus the original problem has been solved.

To derive this algorithm, we have developed several properties for the model. The algorithm is established on the basis of Kalman filter and Kalman smoother, which are repeatedly called to carry on the calculation.

This method is easy to apply, and it is especially suitable when we are only interested in some specific elements of the covariance matrix (3.1), because the algorithm is oriented to calculate the single conditional covariance. Certainly, this algorithm can be revised to make it more efficient if we were interested in values of almost all these parameters of matrix (3.1). However, for this situation, the method to be introduced in the next section would be more efficient. Therefore, the advantage of this method is to be able to calculate a single value.

3.2 Model reconstruction method

The key point of the model reconstruction method is to find an alternative expression of the original state-space model consisting of (2.10) and (2.11), then apply Kalman filter. With this process we can obtain enough information to calculate the related conditional covariances, thus solve the original problem.

At first, let us consider the state equation (2.3), for $t = 2, 3, \cdots$, which can be written as

$$\begin{bmatrix} x_t \\ x_{t-1} \\ x_{t-2} \\ \vdots \\ x_1 \end{bmatrix} = \begin{bmatrix} g_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} F_t & 0 & \cdots & 0 \\ I_{r_{t-1}} & 0 & \cdots & 0 \\ 0 & I_{r_{t-2}} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & I_{r_1} \end{bmatrix} \begin{bmatrix} x_{t-1} \\ x_{t-2} \\ \vdots \\ x_1 \end{bmatrix} + \begin{bmatrix} v_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \tag{3.13}$$

where I_{r_t} represents an $(r_t \times r_t)$ identity matrix and, as we have mentioned before, 0 symbolizes a zero matrix fitted its position. The first r_t rows of (3.13) indicate the state equation and other rows simply state the identity.

By using the following notation:

$$p_t \stackrel{\text{def}}{=} \begin{cases} r_t + r_{t-1} + \cdots + r_1 & t = 1, 2, \cdots, \\ r_0 & t = 0, \end{cases}$$

$$\underset{(p_t\times 1)}{Z_t} \overset{\text{def}}{=} \begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_1 \end{bmatrix}, \quad \underset{(p_t\times 1)}{G_t} \overset{\text{def}}{=} \begin{bmatrix} g_t \\ 0 \\ \vdots \\ 0 \end{bmatrix}, \quad \underset{(p_t\times 1)}{V_t} \overset{\text{def}}{=} \begin{bmatrix} v_t \\ 0 \\ \vdots \\ 0 \end{bmatrix},$$

$$\underset{(p_t\times p_{t-1})}{\Phi_t} \overset{\text{def}}{=} \begin{bmatrix} F_t & 0 & \cdots & 0 \\ I_{r_{t-1}} & 0 & \cdots & 0 \\ 0 & I_{r_{t-2}} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & I_{r_1} \end{bmatrix},$$

(3.13) can be written as

$$Z_t = G_t + \Phi_t Z_{t-1} + V_t. \tag{3.14}$$

We define

$$\underset{(p_t\times p_t)}{Q_t^*} \overset{\text{def}}{=} \begin{bmatrix} Q_t & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix},$$

hence the vector V_t follows $N(0, Q_t^*)$ and $\{V_t\}_{t=1}^{\infty}$ is an independent sequence. We denote $Z_0 \overset{\text{def}}{=} x_0$ and set Z_0 as the initial value of sequence $\{Z_t\}_{t=1}^{\infty}$, obviously Z_0 is independent of $\{V_t\}_{t=1}^{\infty}$. By the definition of Z_0, equation (3.14) holds from $t = 1$. Therefore, if we set Z_t as a state vector, (3.14) would be a state equation.

Similarly, for observation equation (2.6), it can be written as

$$y_t = a_t + \begin{bmatrix} H_t & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} x_t \\ x_{t-1} \\ \vdots \\ x_1 \end{bmatrix} + w_t. \tag{3.15}$$

By using the following notation:

$$\underset{(1\times p_t)}{\Lambda_t} \stackrel{\text{def}}{=} \begin{bmatrix} H_t & 0 & \cdots & 0 \end{bmatrix},$$

(3.15) can be expressed as

$$y_t = a_t + \Lambda_t Z_t + w_t. \tag{3.16}$$

It is easy to notice that (3.16) can be seen as an observation equation of (3.14), which means they can constitute an alternative representation for the original state-space model:

- state equation

$$Z_t = G_t + \Phi_t Z_{t-1} + V_t$$
$$Z_0 \sim N(\gamma, O)$$
$$V_t \sim N(0, Q_t^*),$$

- observation equation

$$y_t = a_t + \Lambda_t Z_t + w_t$$
$$w_t \sim N(0, R_t),$$

for $t = 1, 2, \cdots$.

To distinguish from notations used for the original model, we adopt the following notations:

$$Z_t^s \stackrel{\text{def}}{=} E(Z_t|Y_s),$$

$$C_{t_1,t_2}^{s_1,s_2} \stackrel{\text{def}}{=} E\{(Z_{t_1} - Z_{t_1}^{s_1})(Z_{t_2} - Z_{t_2}^{s_2})'\}$$

and

$$C_{t_1,t_2}^{s} \stackrel{\text{def}}{=} C_{t_1,t_2}^{s,s}, \quad C_t^{s_1,s_2} \stackrel{\text{def}}{=} C_{t,t}^{s_1,s_2}, \quad C_t^s \stackrel{\text{def}}{=} C_{t,t}^{s,s}.$$

Now we apply Kalman filter to this model. First, the initial condition becomes

$$Z_0^0 = \gamma, \quad C_0^0 = O,$$

where γ and O are the mean and variance of the distribution of Z_0. Next step is to calculate the forecast of Z_t conditional on Y_{t-1}. By forecast equations (2.12) and (2.13), we have

$$Z_t^{t-1} = G_t + \Phi_t Z_{t-1}^{t-1}, \tag{3.17}$$

and

$$C_t^{t-1} = \Phi_t C_{t-1}^{t-1} \Phi_t' + Q_t^*. \tag{3.18}$$

The final step is to obtain the filter of Z_t conditional on Y_t. By filter equations (2.14) and (2.15), we have

$$Z_t^t = Z_t^{t-1} + K_t^*(y_t - a_t - \Lambda_t Z_t^{t-1}), \tag{3.19}$$

where

$$K_t^* \stackrel{\text{def}}{=} C_t^{t-1}\Lambda_t'(\Lambda_t C_t^{t-1}\Lambda_t' + R_t)^{-1}.$$

C_t^t can be accomplished as

$$C_t^t = (I_{p_t} - K_t^*\Lambda_t)C_t^{t-1}, \tag{3.20}$$

where I_{p_t} represents an $(p_t \times p_t)$ identity matrix.

By the definition of Z_t^{t-1}, Z_t^t, C_t^{t-1} and C_t^t,

$$Z_t^{t-1} = \begin{bmatrix} x_t^{t-1} \\ x_{t-1}^{t-1} \\ \vdots \\ x_1^{t-1} \end{bmatrix}, \qquad Z_t^t = \begin{bmatrix} x_t^t \\ x_{t-1}^t \\ \vdots \\ x_1^t \end{bmatrix},$$

$$C_t^{t-1} = \begin{bmatrix} P_t^{t-1} & P_{t,t-1}^{t-1} & \cdots & P_{t,1}^{t-1} \\ P_{t-1,t}^{t-1} & P_{t-1}^{t-1} & \cdots & P_{t-1,1}^{t-1} \\ \vdots & \vdots & \cdots & \vdots \\ P_{1,t}^{t-1} & P_{1,t-1}^{t-1} & \cdots & P_1^{t-1} \end{bmatrix},$$

$$C_t^t = \begin{bmatrix} P_t^t & P_{t,t-1}^t & \cdots & P_{t,1}^t \\ P_{t-1,t}^t & P_{t-1}^t & \cdots & P_{t-1,1}^t \\ \vdots & \vdots & \cdots & \vdots \\ P_{1,t}^t & P_{1,t-1}^t & \cdots & P_1^t \end{bmatrix}.$$

Note that C_T^T is the lower right part of the conditional covariance matrix (3.1) and, if we look back on the algorithm of recursive method, calculation of elements from C_T^T are most laborious, therefore C_T^T provides a partial but important solution to the original problem. C_T^T can be calculated through the Kalman filter for the rewritten model. The rest of the matrix (3.1) can be easily calculated as follows. Assume that we have already computed the value of C_T^T, then by the property 3 of lemma,

$$P_{T+i,j}^T = F_{T+i} P_{T+i-1,j}^T$$

holds for $i = 1, 2, \cdots, m$ and $j = 1, 2, \cdots, T$. Therefore the upper right part of matrix (3.1)

$$\begin{bmatrix} P_{T+m,T}^T & \cdots & P_{T+m,1}^T \\ \vdots & \cdots & \vdots \\ P_{T+1,T}^T & \cdots & P_{T+1,1}^T \end{bmatrix}$$

can be obtained from C_T^T. In addition, by the property 4 of lemma,

$$P_{T+i}^T = F_{T+i} P_{T+i-1}^T F'_{T+i} + Q_{T+i}$$

holds for $i = 1, 2, \cdots, m$. Thus, we can obtain all the values of elements on the main diagonal of matrix (3.1) from C_T^T. As for $P_{T+i,T+j}^T$ ($i = 1, 2, \cdots, m$, $j = 1, 2, \cdots, m$ and $i > j$), by the property 3 of lemma, we have

$$P_{T+i,T+j}^T = F_{T+i} P_{T+i-1,T+j}^T,$$

which implies $P_{T+i,T+j}^T$ can be obtained from the P_{T+j}^T on the main diagonal. Thus, by this procedure, all the values from the upper triangular part of matrix (3.1) can be computed. Since the covariance matrix is symmetric, we can compute the whole covariance matrix (3.1) accordingly.

In summary, to compute matrix (3.1), we start by calculation C_T^T via Kalman filter. Then obtain the values of the upper triangular part from C_T^T by some easy procedures, thus the whole matrix by symmetry.

Although we have solved the original problem, a close look into (3.17), (3.18), (3.19) and (3.20) reveals more information. (3.17) and (3.18) can be written as

$$\begin{bmatrix} x_t^{t-1} \\ x_{t-1}^{t-1} \\ x_{t-2}^{t-1} \\ \vdots \\ x_1^{t-1} \end{bmatrix} = \begin{bmatrix} g_t \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} + \begin{bmatrix} F_t & 0 & \cdots & 0 \\ I_{r_{t-1}} & 0 & \cdots & 0 \\ 0 & I_{r_{t-2}} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & I_{r_1} \end{bmatrix} \begin{bmatrix} x_{t-1}^{t-1} \\ x_{t-2}^{t-1} \\ \vdots \\ x_1^{t-1} \end{bmatrix} \tag{3.21}$$

and

$$\begin{aligned} &\left[\begin{array}{c:ccc} P_t^{t-1} & P_{t,t-1}^{t-1} & \cdots & P_{t,1}^{t-1} \\ \hdashline P_{t-1,t}^{t-1} & P_{t-1}^{t-1} & \cdots & P_{t-1,1}^{t-1} \\ \vdots & \vdots & \cdots & \vdots \\ P_{1,t}^{t-1} & P_{1,t-1}^{t-1} & \cdots & P_1^{t-1} \end{array}\right] \\ &= \Phi_t \begin{bmatrix} P_{t-1}^{t-1} & P_{t-1,t-2}^{t-1} & \cdots & P_{t-1,1}^{t-1} \\ P_{t-2,t-1}^{t-1} & P_{t-2}^{t-1} & \cdots & P_{t-2,1}^{t-1} \\ \vdots & \vdots & \cdots & \vdots \\ P_{1,t-1}^{t-1} & P_{1,t-2}^{t-1} & \cdots & P_1^{t-1} \end{bmatrix} \Phi_t' + Q_t^* \\ &= \left[\begin{array}{c:ccc} F_t P_{t-1}^{t-1} F_t' + Q_t & F_t P_{t-1}^{t-1} & \cdots & F_t P_{t-1,1}^{t-1} \\ \hdashline P_{t-1}^{t-1} F_t' & P_{t-1}^{t-1} & \cdots & P_{t-1,1}^{t-1} \\ \vdots & \vdots & \cdots & \vdots \\ P_{1,t-1}^{t-1} F_t' & P_{1,t-1}^{t-1} & \cdots & P_1^{t-1} \end{array}\right]. \end{aligned} \tag{3.22}$$

The r_t rows from the top of (3.21) state

$$x_t^{t-1} = g_t + F_t x_{t-1}^{t-1},$$

which is the forecast equation (2.12) of original model. Considering the equation (3.22), we have

$$P_t^{t-1} = F_t P_{t-1}^{t-1} F_t' + Q_t,$$

which is the forecast error (2.13), and we also have

$$P_{t,i}^{t-1} = F_t P_{t-1,i}^{t-1},$$

for $i = 1, 2, \cdots, t-1$, it is just a instance of property 3 of lemma.

Then consider (3.19), which could be written as

$$\begin{aligned}
\begin{bmatrix} x_t^t \\ x_{t-1}^t \\ \vdots \\ x_1^t \end{bmatrix} &= \begin{bmatrix} x_t^{t-1} \\ x_{t-1}^{t-1} \\ \vdots \\ x_1^{t-1} \end{bmatrix} + C_t^{t-1} \Lambda_t' (\Lambda_t C_t^{t-1} \Lambda_t' + R_t)^{-1} (y_t - a_t - \Lambda_t Z_t^{t-1}) \\
&= \begin{bmatrix} x_t^{t-1} \\ x_{t-1}^{t-1} \\ \vdots \\ x_1^{t-1} \end{bmatrix} + C_t^{t-1} \begin{bmatrix} H_t' \\ 0 \\ \vdots \\ 0 \end{bmatrix} \left(\begin{bmatrix} H_t & \hat{0} & \cdots & \hat{0} \end{bmatrix} C_t^{t-1} \begin{bmatrix} H_t' \\ 0 \\ \vdots \\ 0 \end{bmatrix} + R_t \right)^{-1} (y_t - a_t - H_t x_t^{t-1}) \\
&= \begin{bmatrix} x_t^{t-1} \\ x_{t-1}^{t-1} \\ \vdots \\ x_1^{t-1} \end{bmatrix} + \begin{bmatrix} P_t^{t-1} H_t' \\ P_{t-1,t}^{t-1} H_t' \\ \vdots \\ P_{1,t}^{t-1} H_t' \end{bmatrix} (H_t P_t^{t-1} H_t' + R_t)^{-1} (y_t - a_t - H_t x_t^{t-1}).
\end{aligned} \tag{3.23}$$

By (3.23), we have

$$x_i^t = x_i^{t-1} + P_{i,t}^{t-1} H_t' (H_t P_t^{t-1} H_t' + R_t)^{-1} (y_t - a_t - H_t x_t^{t-1}), \tag{3.24}$$

for $i = 1, 2, \cdots, t$. Particularly, when $i = t$, (3.24) becomes

$$x_t^t = x_t^{t-1} + P_t^{t-1} H_t' (H_t P_t^{t-1} H_t' + R_t)^{-1} (y_t - a_t - H_t x_t^{t-1}), \tag{3.25}$$

which is the filter equation (2.14). In addition, if we consider (3.24) together with (3.25), the following holds:

$$\begin{aligned} x_i^t &= x_i^{t-1} + P_{i,t}^{t-1} (P_t^{t-1})^{-1} (x_t^t - x_t^{t-1}) \\ &= x_i^{t-1} + P_{i,t-1}^{t-1} F_t' (P_t^{t-1})^{-1} (x_t^t - x_t^{t-1}), \end{aligned} \tag{3.26}$$

for $i = 1, 2, \cdots, t-1$. We should notice that (3.26) also provides another way to calculate smoothed estimators x_i^t $(i < t)$. Assume we already know the values of sequence $\{P_{a,b}^s\}$ $(a, b \le s \le t)$ and have run Kalman filter, then we can obtain x_i^t $(i < t)$ from x_i^i by iterating on (3.26). As for the calculation of $P_{a,b}^s$ $(a, b \le s \le t)$, this can be accomplished as follows.

Let us consider (3.20), that can be written as

$$\begin{bmatrix} P_t^t & P_{t,t-1}^t & \cdots & P_{t,1}^t \\ P_{t-1,t}^t & P_{t-1}^t & \cdots & P_{t-1,1}^t \\ \vdots & \vdots & \cdots & \vdots \\ P_{1,t}^t & P_{1,t-1}^t & \cdots & P_1^t \end{bmatrix}$$

$$= C_t^{t-1} - C_t^{t-1}\Lambda_t'(\Lambda_t C_t^{t-1}\Lambda_t' + R_t)^{-1}\Lambda_t C_t^{t-1}$$

$$= \begin{bmatrix} P_t^{t-1} & P_{t,t-1}^{t-1} & \cdots & P_{t,1}^{t-1} \\ P_{t-1,t}^{t-1} & P_{t-1}^{t-1} & \cdots & P_{t-1,1}^{t-1} \\ \vdots & \vdots & \cdots & \vdots \\ P_{1,t}^{t-1} & P_{1,t-1}^{t-1} & \cdots & P_1^{t-1} \end{bmatrix} - \begin{bmatrix} P_t^{t-1}H_t' \\ P_{t-1,t}^{t-1}H_t' \\ \vdots \\ P_{1,t}^{t-1}H_t' \end{bmatrix} (H_t P_t^{t-1} H_t' + R_t)^{-1} \begin{bmatrix} P_t^{t-1}H_t' \\ P_{t-1,t}^{t-1}H_t' \\ \vdots \\ P_{1,t}^{t-1}H_t' \end{bmatrix}'. \tag{3.27}$$

(3.27) indicates that

$$P_{i,j}^t = P_{i,j}^{t-1} - P_{i,t}^{t-1}H_t'(H_t P_t^{t-1} H_t' + R_t)^{-1} H_t P_{t,j}^{t-1}, \tag{3.28}$$

for $i = 1, 2, \cdots, t$ and $j = 1, 2, \cdots, t$. When $i = j = t$, equation (3.28) becomes the error of filter (2.15). Like (3.26), equation (3.28) also gives an alternative approach to calculate the covariance of smoother of original model $P_{i,j}^t$ $(i, j < t)$. Equation (3.28) can be written as

$$P_{i,j}^t = \begin{cases} P_t^t & i = j = t, \\ F_t P_{t-1,j}^{t-1} - P_t^{t-1}H_t'(H_t P_t^{t-1}H_t' + R_t)^{-1}H_t F_t P_{t-1,j}^{t-1} & i = t, j < t, \\ P_{i,t-1}^{t-1}F_t' - P_{i,t-1}^{t-1}F_t'H_t'(H_t P_t^{t-1}H_t' + R_t)^{-1}H_t P_t^{t-1} & i < t, j = t, \\ P_{i,j}^{t-1} - P_{i,t-1}^{t-1}F_t'H_t'(H_t P_t^{t-1}H_t' + R_t)^{-1}H_t F_t P_{t-1,j}^{t-1} & i < t, j < t. \end{cases} \tag{3.29}$$

Therefore if we have run Kalman filter, by (3.29), the value of $P^t_{i,j}$ $(i, j \leq t)$ could be obtained by iterating on equation (3.28). Since t could be any positive integer, based on Kalman filter, we have already developed a method to calculate $P^s_{a,b}$ $(a, b \leq s \leq t)$. Since equation (3.28) can also be used to calculate the conditional covariances of smoothers like the property 6 of lemma, it could substitute for the property 6 in the recursive method. It is easy to notice that equations (3.26) and (3.28) could constitute an alternative method to calculate the smoothed estimators and their covariance. The calculation through this procedure is not as simple as for the Kalman smoother, however it could provide a way to obtain covariances for smoothers, not only variances.

3.3 Extension of the state-space model

The state-space model we adopted ((2.10) and (2.11)) actually is a rather general one, we can clearly notice that parameters of this model can vary with time and the state or observation dimension can vary over time too. However we can still extend our state-space model to make it more versatile.

Consider a new model:

- state equation

$$x_t = \begin{cases} g_t + F_{t,t-1}x_{t-1} + \cdots + F_{t,1}x_1 + v_t & t = 2, 3, \cdots \\ g_1 + F_{1,0}x_0 + v_1 & t = 1 \end{cases}$$
$$x_0 \sim N(\gamma, O)$$
$$v_t \sim N(0, Q_t), \tag{3.30}$$

- observation equation

$$\begin{aligned} & y_t = a_t + H_{t,t}x_t + \cdots + H_{t,1}x_1 + w_t \quad (t = 1, 2, \cdots) \\ & w_t \sim N(0, R_t), \end{aligned} \tag{3.31}$$

Like before, assume all the elements of g_t, a_t, $F_{t,i}$ $(0 < i < t)$, $H_{t,j}$ $(0 < j \leq t)$, γ, O, Q_t and R_t are known with certainty, and $\{v_t\}$, $\{w_t\}$ as well as x_0 are mutually independent. For convenience, we call the new model G-model and the original one N-model.

Compared to N-model, the present observation and state of G-model may depend on some or all state vectors in the past. In fact, lagged variables in the observation equation (3.31) are useful, the article of Qian [2014] gives us several cases for the applications of this. However allowing more than one lagged state vector in the state equation (3.30) is rarely seen in the literature, since it would make state vectors lose Markov property. In practice, we usually set $F_{t,t-2} = \cdots = F_{t,1} = 0$, for now we assume they are given functions. These changes in model can broaden the scope of application of model, however they also make it more difficult to obtain forecasts, filters and smoothers. The original Kalman filter and smoother become inapplicable for this model, furthermore it is hard to propose a new edition of Kalman filter and smoother for G-model based on the original idea invovled in the proof for N-model. Fortunately it is straightforward to calculate these estimators for G-model through the model reconstruction procedure introduced in last section. Next, we would illustrate these results.

We denote

$$\underset{(p_t \times p_{t-1})}{\Phi_t^\bullet} \stackrel{\text{def}}{=} \begin{bmatrix} F_{t,t-1} & F_{t,t-2} & \cdots & F_{t,1} \\ I_{r_{t-1}} & 0 & \cdots & 0 \\ 0 & I_{r_{t-2}} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & I_{r_1} \end{bmatrix}$$

and

$$\underset{(1 \times p_t)}{\Lambda_t^\bullet} \stackrel{\text{def}}{=} \begin{bmatrix} H_{t,t} & H_{t,t-1} & \cdots & H_{t,1} \end{bmatrix}.$$

By using notations Z_t, G_t, V_t in the last section, the state equation (3.30) can be written as

$$Z_t = G_t + \Phi_t^\bullet Z_{t-1} + V_t, \tag{3.32}$$

which holds from $t = 1$. Similarly, the observation equation (3.31) can be written as

$$y_t = a_t + \Lambda_t^\bullet Z_t + w_t. \tag{3.33}$$

If we treat Z_t as a state vector, (3.32) can be seen as the state equation for Z_t and (3.33) can be the corresponding observation equation. Therefore the equations (3.32) and (3.33) constitute a N-model for state Z_t. Considering the relationship between x_t of the G-model and Z_t of the N-model, we should expect to obtain forecasts, filters and smoothers for G-model by applying Kalman filter to N-model ((3.32) and (3.33)).

Now we apply Kalman filter. At first, we set the initial condition:

$$Z_0^0 = \gamma, \quad C_0^0 = O.$$

By the forecast equation (2.12) and (2.13), we have

$$Z_t^{t-1} = G_t + \Phi_t^\bullet Z_{t-1}^{t-1} \tag{3.34}$$

and

$$C_t^{t-1} = \Phi_t^\bullet C_{t-1}^{t-1}(\Phi_t^\bullet)' + Q_t^*. \tag{3.35}$$

By the filter equation (2.14) and (2.15), we have

$$Z_t^t = Z_t^{t-1} + C_t^{t-1}(\Lambda_t^\bullet)'(\Lambda_t^\bullet C_t^{t-1}(\Lambda_t^\bullet)' + R_t)^{-1}(y_t - a_t - \Lambda_t^\bullet Z_t^{t-1}) \tag{3.36}$$

and

$$C_t^t = C_t^{t-1} - C_t^{t-1}(\Lambda_t^\bullet)'(\Lambda_t^\bullet C_t^{t-1}(\Lambda_t^\bullet)' + R_t)^{-1}\Lambda_t^\bullet C_t^{t-1}. \tag{3.37}$$

Due to the complex structure of $\Phi_t^\bullet$ and $\Lambda_t^\bullet$, it is difficult to derive good results such as (3.26) or (3.28) from equation (3.34) to (3.37). For example, even if only one element of C_t^t is needed, we still have to compute all the elements of C_t^{t-1}. Hence we prefer to calculate these data by using (3.34) to (3.37) directly. Since

$$Z_t^t = \begin{bmatrix} x_t^t \\ x_{t-1}^t \\ \vdots \\ x_1^t \end{bmatrix},$$

we can obtain the filters and smoothers of G-model from the filters of corresponding

N-model. Based on the state equation (3.30),

$$x^T_{T+m} = g_{T+m} + F_{T+m,T+m-1}x^T_{T+m-1} + \cdots + F_{T+m,T}x^T_T + \cdots + F_{T+m,1}x^T_1 \qquad (3.38)$$

hold for $m = 1, 2, \cdots$ and $T = 1, 2, \cdots$. As we already know the filters and smoothers, the forecast x^T_{T+m} can be obtained by iterating on equation (3.38). Therefore, by the procedure introduced in last section, we can get forecasts, filters and smoothers of the new model.

Chapter 4

Summary and future work

The calculation of estimators of the underlying unobserved states on the basis of observations is a classic and important problem of state-space models. This has been accomplished by Kalman filter and Kalman smoother. Even though we can also have errors of these estimators through Kalman filter and smoother, the statistical relationships between these estimators can not be accordingly obtained. Several articles, reports or books have given lag-one covariances for smoothers, enlightened by this, at first we propose an algorithm to calculate any covariance for these estimators, which provides a good solution to the problem. In fact, based on the same observation, these estimators follow a normal distribution. All the parameters of the distribution can be calculated by our new algorithm and Kalman filter and smoother. In addition, we also try another method to solve this problem. We reconstruct the model and apply Kalman filter to the new model. Then we can obtain a partial solution, by which the complete solution could be achieved through some easy procedures. We continue to explore the result and derive some interesting properties, which provide another view to establish the algorithm for the calculation of covariances. In the end, we extend the definition of our state-space model, and obtain a more general new model. By the model reconstruction method, we can obtain an approach to calculating those estimators as well.

There are still some problems for further research. The algorithm established in the first section of chapter 3 is designed for the calculation of single specific conditional covariance, which means the output of this algorithm is just one covariance. If we want large numbers of covariances, we have to fulfil the algorithm several times, there must be large redundancies in the process. It persuades us to reconstruct the algorithm for those cases, although this work should be easy. Another problem is that we want an easier and clearer way to get estimators like forecasts, filters and smoothers for the new model in third section of chapter 3, even though the application of model reconstruction method could get them successfully. The structure of the new model is so complicated that we could not expect an easy algorithm like Kalman filter or Kalman smoother, however we do hope to get an easy algorithm like what we have established in this article, at least for some special cases of the new model. These problems provide the aim for our future work.

Bibliography

Brian DO Anderson and John B Moore. *Optimal filtering*. Courier Corporation, 2012.

Olivier Basdevant. On applications of state-space modelling in macroeconomics. *Reserve Bank of New Zealand Discussion Paper No. DP2003/02*, 2003.

Peter J Brockwell and Richard A Davis. *Time series: theory and methods*. Springer Science & Business Media, 2013.

James Durbin and Siem Jan Koopman. *Time series analysis by state space methods*. Number 38. Oxford University Press, 2012.

James Douglas Hamilton. *Time series analysis*, volume 2. Princeton university press Princeton, 1994.

Andrew C Harvey. *Forecasting, structural time series models and the Kalman filter*. Cambridge university press, 1990.

Andrew C Harvey and Richard G Pierse. Estimating missing observations in economic time series. *Journal of the American Statistical Association*, 79(385):125–131, 1984.

Richard H Jones. Maximum likelihood fitting of arma models to time series with missing observations. *Technometrics*, 22(3):389–395, 1980.

Richard H Jones. Fitting multivariate models to unequally spaced data. In *Time series analysis of irregularly observed data*, pages 158–188. Springer, 1984.

Borus Jungbacker, Siem Jan Koopman, and Michel Van der Wel. Maximum likelihood estimation for dynamic factor models with missing data. *Journal of Economic Dynamics and Control*, 35(8):1358–1368, 2011.

Rudolph E Kalman and Richard S Bucy. New results in linear filtering and prediction theory. *Journal of basic engineering*, 83(1):95–108, 1961.

Rudolph Emil Kalman. A new approach to linear filtering and prediction problems. *Journal of basic Engineering*, 82(1):35–45, 1960.

Philip F Mclauchlan and David W Murray. Active camera calibration for a head-eye platform using the variable state-dimension filter. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 18(1):15–22, 1996.

Sascha Mergner. Applications of state space models in finance. 2009.

Hang Qian. A flexible state space model and its applications. *Journal of Time Series Analysis*, 35(2):79–88, 2014.

Robert H Shumway and David S Stoffer. *Time series analysis and its applications: with R examples*. Springer Science & Business Media, 2010.

Ruey S Tsay. Analysis of financial time series. 2010.

Ruey S Tsay. Multivariate time series analysis : with r and financial applications. 2014.