Economics Letters 127 (2015) 10–13

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

A low dimensional Kalman filter for systems with lagged states in the measurement equation

Kristoffer P. Nimark

Economics Department, Cornell University, Ithaca, 14853, NY, United States

HIGHLIGHTS

- Presents a modified Kalman filter where the vector of observables can depend on lagged states.
- The modified filter does not require increasing the state dimension.
- The modified filter can be used together with the Kalman simulation smoother.

ARTICLE INFO

Article history:
Received 6 October 2014
Received in revised form
10 December 2014
Accepted 11 December 2014
Available online 19 December 2014

JEL classification:
C1

Keywords:
Kalman filter
Lagged observables
Kalman smoother
Simulation smoother

ABSTRACT

This note describes how the Kalman filter and the Kalman smoother can be modified to allow for the vector of observables to be a function of lagged state variables without increasing the dimension of the state vector in the filter.

This note describes how the Kalman filter can be modified to allow for the vector of observable variables in the measurement equation to be a function of lagged latent variables. The standard approach, which is to augment the state vector of the filter to include also lagged variables, works well in most applications. However, it also doubles the dimension of the state vector which is undesirable in some applications. The modified filter presented here avoids increasing the dimension of the state by exploiting that the innovation representation can be modified so as to make it unnecessary to augment the state vector with lagged variables.

In a related paper, Qian (2014) shows how the Kalman filter can be modified to allow for the current state and observation vectors to depend on both lagged observation and state vectors. The main difference between that paper and the present is thus that here the current observation vector may only depend on lagged latent states, rather than on lagged variables that are directly observable. This results in somewhat simpler derivations.

The derivation of the modified filter, which nests the standard filter as a special case, is presented in the next section. This is followed by a brief description of how to use the modified filter together with standard algorithms for the Kalman Smoother and Kalman Simulation Smoother. The last section concludes and references existing work where the modified filter has proved to be useful.

1. A filtering problem

Consider a standard state space system augmented to allow the measurement equation to depend on lagged states

$$X_t = AX_{t-1} + Cu_t : u_t \sim N(0, I) \tag{1.1}$$

$$Z_t = D_1X_t + D_2X_{t-1} + Ru_t \tag{1.2}$$

where X_t is the $n \times 1$ dimensional state vector, A is an $n \times n$ matrix, C is and $n \times m$ matrix. Z_t is a $p \times 1$ vector of observable variables and D_1 and D_2 are both $p \times n$ matrices and R is a $p \times m$ matrix.

Define the notation

$$X_{t|t-s} \equiv E\left[X_t \mid Z^{t-s}, X_{0|0}\right] \tag{1.3}$$

E-mail address: pkn8@cornell.edu.
URL: http://www.kris-nimark.net.

http://dx.doi.org/10.1016/j.econlet.2014.12.016
0165-1765/

$$P_{t|t-s} \equiv E\left[\left(X_t - X_{t|t-s}\right)\left(X_t - X_{t|t-s}\right)'\right] \tag{1.4}$$

where $X_{0|0}$ is the mean of the exogenously prior distribution of X_0 given by

$$X_0 \sim N\left(X_{0|0}, P_{0|0}\right). \tag{1.5}$$

We want to find the Kalman gain K_t in the recursive updating equation

$$X_{t|t} = AX_{t-1|t-1} + K_t\left[Z_t - (D_1A + D_2)X_{t-1|t-1}\right] \tag{1.6}$$

so that $X_{t|t}$ is the conditional minimum variance estimate of X_t.

1.1. The standard approach

The state space system (1.1)–(1.2) is standard apart from the fact that the vector of observables Z_t depends on both the current and the lagged state. A straightforward and common way to get around this problem is to redefine the state so as to include also lagged X_t to get

$$\overline{X}_t = \overline{AX}_{t-1} + \overline{C}u_t \tag{1.7}$$

$$Z_t = \overline{DX}_t + Ru_t \tag{1.8}$$

where

$$\overline{X}_t = \begin{bmatrix} X_t' & X_{t-1}' \end{bmatrix}', \qquad \overline{A} = \begin{bmatrix} A & \mathbf{0} \\ I & \mathbf{0} \end{bmatrix}$$

$$\overline{C} = \begin{bmatrix} C \\ \mathbf{0} \end{bmatrix}, \qquad \overline{D} = \begin{bmatrix} D_1 & D_2 \end{bmatrix}.$$

The standard filter can then be applied to the augmented system (1.7)–(1.8). In most application, this does not cause any complications. However, in some cases it is desirable to have a state of low dimensionality and redefining the state as above doubles the dimension of the state, i.e. $\overline{X}_t$ is a $2n \times 1$ vector. Below, a new filter is derived that solves the filtering problem while maintaining an n-dimensional state vector.

2. A modified filter

In this section, the modified filter is derived. The system is linear with Gaussian disturbances and the minimum variance estimate of the latent state then coincides with the orthogonal projection onto the set of conditioning variables (or signals). The filter is derived using the Gram–Schmidt approach of recursively orthogonalizing the time series of observable variables.[1] This approach exploits that the projection of a random variable onto a set of mutually orthogonal signals is equivalent to adding up the projections of the variable onto the individual signals. That is,

$$E(x \mid z, y) = E(x \mid z) + E(x \mid y) \tag{2.1}$$

if

$$E(zy') = 0 \tag{2.2}$$

and x, y and z are zero-mean Gaussian random variables.

It is the property (2.1)–(2.2) that will allow us to write down a recursive update equation for $X_{t|t}$. The first step is to find the projection of X_t onto the component of the period t signals that is orthogonal to information known in period $t-1$. This projection can then be added to the prior estimate $X_{t|t-1}$, i.e. the projection of X_t onto period $t-1$ information, to form a posterior estimate $X_{t|t}$. To this end, define the innovation $\widetilde{Z}_t$ as the component of Z_t that is orthogonal to period $t-1$ information

$$\widetilde{Z}_t \equiv Z_t - Z_{t|t-1} \tag{2.3}$$

so that the posterior estimate $X_{t|t}$ will be given by

$$X_{t|t} = X_{t|t-1} + E\left(X_t \mid \widetilde{Z}_t\right). \tag{2.4}$$

To solve the filtering problem we thus need to find an expression for $E\left(X_t \mid \widetilde{Z}_t\right)$. We will start by solving this problem for period 1. The resulting expressions are then straightforward to generalize to period t.

2.1. Projecting the state onto the innovation in the observable vector

In the initial period there are two pieces of information available: the exogenously given prior distribution (1.5) and the initial signal Z_1. By (2.4) the prior and the signal can be combined as

$$X_{1|1} = X_{1|0} + K_1\widetilde{Z}_1 \tag{2.5}$$

to form the posterior mean $X_{1|1}$ if $K_1\widetilde{Z}_1 = E\left(X_1 \mid \widetilde{Z}_1\right)$. From the projection theorem (e.g. Brockwell and Davis, 2006), the appropriate K_1 is given by the standard projection formula

$$K_1 = E\left(X_1\widetilde{Z}_1'\right)\left[E\left(\widetilde{Z}_1\widetilde{Z}_1'\right)\right]^{-1}. \tag{2.6}$$

To compute the Kalman gain K_1 we thus need to derive operational expressions for $E\left(X_1\widetilde{Z}_1'\right)$ and $E\left(\widetilde{Z}_1\widetilde{Z}_1'\right)$.

2.2. The covariance of the state and the innovation vector

To find the covariance $E\left(X_1\widetilde{Z}_1'\right)$, start by using the identities implied by (1.1)–(1.2) to rewrite the innovation as

$$\widetilde{Z}_t = (D_1A + D_2)\left(X_0 - X_{0|0}\right) + (D_1C + R)\,u_1. \tag{2.7}$$

It is helpful to define the posterior state estimation error $\widetilde{X}_t$ as

$$\widetilde{X}_t \equiv X_t - X_{t|t} \tag{2.8}$$

and use this together with (2.7) to express the covariance of the state and the innovation as

$$\begin{aligned} E\left(X_1\widetilde{Z}_1'\right) = {} & E\left[\left(A\left(\widetilde{X}_0 + X_{0|0}\right) + Cu_1\right)\right. \\ & \left.\times\left((D_1A + D_2)\widetilde{X}_0 + D_1Cu_1 + Ru_1\right)'\right]. \end{aligned} \tag{2.9}$$

Since $E\left(X_{0|0}\widetilde{X}_0'\right) = 0$ and $P_{0|0} \equiv E\left(\widetilde{X}_0\widetilde{X}_0'\right)$ Eq. (2.9) can be simplified to

$$E\left(X_1\widetilde{Z}_1'\right) = AP_{0|0}\left(D_1A + D_2\right)' + CC'D_1' + CR'. \tag{2.10}$$

We thus have the first term in the Kalman gain (2.6).

2.3. The covariance of the innovation vector

To find the covariance of the innovation vector $\widetilde{Z}_1$, simply use that (2.7) implies that

$$\begin{aligned} E\left(\widetilde{Z}_1\widetilde{Z}_1'\right) = {} & (D_1A + D_2)P_{0|0}(D_1A + D_2)' \\ & + (D_1C + R)(D_1C + R)' \end{aligned} \tag{2.11}$$

yielding the second term in the Kalman gain (2.6).

[1] A derivation of the standard filter along similar lines can be found in Anderson and Moore (1979).