



Conditional forecasts on SVAR models using the Kalman filter

Gonzalo Camba-Mendez*

European Central Bank, Kaiserstrasse 29, D-60311, Frankfurt am Main, Germany

ARTICLE INFO

Article history:

Received 8 November 2010

Received in revised form

11 November 2011

Accepted 19 December 2011

Available online 3 January 2012

JEL classification:

C32

C53

Keywords:

Conditional forecasting

Vector autoregression

Kalman filter

ABSTRACT

This note shows how conditional forecasts from identified VAR models can be computed using Kalman filtering techniques. These techniques are nowadays routine for applied macroeconomists, and hence the computation of conditional forecasts using these methods are simple to implement.

© 2012 Elsevier B.V. All rights reserved.

1. Introduction

Ever since the publication of Sims (1980), the use of vector autoregressive (VAR) models for applied economic research and policy analysis has been widespread. Impulse response analysis from well identified VAR systems, the so called 'structural' VAR models (SVAR), has been routinely used to ascertain the impact of a certain shock on macroeconomic variables. Policy institutions are often interested not only on the impact of a single shock but rather on the forecast of a macroeconomic variable conditioned on certain assumptions on the future path of certain macroeconomic variables or shocks. This would allow for example the computation of forecasts of GDP growth under alternative assumptions on the future path of monetary policy, say gradual tightening, or loosening of the monetary policy stance. Waggoner and Zha (1999) provided a satisfactory method to solve this problem. The aim of this note is to illustrate how conditional forecasts from SVAR models can be also simply computed using Kalman filtering techniques. The use of Kalman filtering techniques is nowadays routine for applied macroeconomists, and hence the computation of conditional forecasts and associated error bands using these methods should be straightforward to implement. The method relies on the formulation of a state space representation for the SVAR model. This state space representation of the SVAR model

enables to easily formulate conditional forecasts which rely on imposing alternative assumptions on either the value of future endogenous variables or the value of future well identified shocks, or on a combination of both.

Whether the computation of conditional forecasts by means of the Bayesian techniques put forward in Waggoner and Zha (1999) and extended in Jarocinski (2010) is computationally more efficient than the techniques presented in this note remains an issue for further research. It is our view that this issue will be very much dependent on computer architecture structure and on the actual implementation of the programming code. For example, many of the matrix algebra operations in the Kalman filter iterations reported in this note can be very much simplified by means of sparse matrix methods. Of course, the techniques reported in Waggoner and Zha (1999) are more general and can deal with 'range' conditional forecasts, i.e., the projected path of a certain endogenous variable is constrained to lie within a certain range; something not dealt with in this note.

2. Conditional forecasts for SVAR models

2.1. State space representation of SVAR models

Assume that the dynamics of a $k \times 1$ macroeconomic series \mathbf{x}_t is well described by the following SVAR model:

$$\mathbf{x}_t = \mathbf{B}\mathbf{x}_{t-1} + \mathbf{R}\boldsymbol{\varepsilon}_t, \quad (1)$$

where \mathbf{B} and \mathbf{R} are $k \times k$ matrices of parameters and $\boldsymbol{\varepsilon}_t$ is a $k \times 1$ zero mean Gaussian random vector with identity covariance matrix.

* Tel.: +49 69 1344 6481; fax: +49 69 1344 6514.

E-mail addresses: gonzalo.camba-mendez@ecb.europa.eu, gonzalo.camba-mendez@ecb.int.

For the sake of clarity only one lag of \mathbf{x}_t is included on the right hand side of (1), extension to more lags follows easily. If we define $\mathbf{y}_t = (\mathbf{x}'_t, \xi'_t)'$, where the new artificial variable ξ_t is nothing but \mathbf{e}_{t+1} in disguise, the SVAR model can be written in state space form as:

$$\mathbf{y}_t = \mathbf{C}\mathbf{s}_t$$

$$\mathbf{s}_t = \mathbf{A}\mathbf{s}_{t-1} + \mathbf{u}_t, \tag{2}$$

where

$$\mathbf{C} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}; \quad \mathbf{A} = \begin{pmatrix} \mathbf{B} & \mathbf{R} \\ \mathbf{0} & \mathbf{0} \end{pmatrix};$$

$$\mathbf{u}_t = \begin{pmatrix} \mathbf{0} \\ \mathbf{e}_{t+1} \end{pmatrix}; \quad \Sigma_{\mathbf{u}} = E\mathbf{u}_t\mathbf{u}'_t = \begin{pmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$$

and \mathbf{I} denotes a k -dimensional identity matrix, and the dimensions of the other matrices are defined accordingly. When the parameter matrices of the state space model (2) are a function of a known parameter vector θ_0 , for known starting values of \mathbf{s}_0 and its covariance matrix \mathbf{P}_0 , the fixed interval smoother of Ansley and Kohn (1982) provides an optimal estimator, in the mean square error sense, of the state \mathbf{s}_t conditional on available information. Importantly, in the implementation of the fixed interval smoother shown below, missing observations in \mathbf{y}_t can be properly handled. When the length of the \mathbf{y}_t series is T , then this estimator is denoted as $\mathbf{s}_{t|T} = E\{\mathbf{s}_t | \mathbf{y}_1, \dots, \mathbf{y}_T, \theta_0\}$.

Conditional forecasts in the context of our state space model can thus be simply retrieved from this estimator. The macroeconomic series \mathbf{x}_t are observed from $t = 1$ to T , while over the projection horizon $t = T + 1$ to $T + P$ some of them are treated as known, while others, those for which the conditional forecast will be computed, are treated as missing. As for the shocks ξ_t , these are treated as missing from $t = 1$ to $T - 1$ while from $t = T$ to $T + P - 1$ we may choose to treat some as known by imposing a certain value on those observations. Conditional forecasts are then computed from $\mathbf{s}_{t|T+P}$. Table 1 illustrates different types of conditional forecasts which may be computed by imposing a variety of restrictions on a three-dimensional macroeconomic series and corresponding shocks.

- *Unconditional projection (CF1)*. This is equivalent to running standard forecasts with a standard VAR model.¹
- *'Impulse response' conditional projection (CF2)*. The output provided would show the projected series assuming a unit shock to one of the innovations. Contrary to what is usually reported when displaying standard impulse response functions, the conditional forecasts provide the path for the series, and these depend on the starting value \mathbf{x}_T .
- *'Path shock' conditional projection (CF3)*. The output provided would show the projected series under the assumption that we know the future path of certain shocks. This would allow for example the possibility of studying the impact of gradual monetary policy tightening on growth.
- *'Path shock-series' conditional projection (CF4)*. This would provide the conditional forecasts under certain assumptions on the future path of certain regressors and certain shocks.
- *'Unbalanced path shock-series' conditional projection (CF5)*. This would provide the conditional forecasts under certain assumptions on the future path of certain regressors and certain shocks at some given points in time which may not necessarily coincide.

¹ A few remarks on the implementation of this method are in order. When ξ_t is assumed as 'missing' or 'unobserved', the model collapses to the standard SVAR model, and for a correct initialization of \mathbf{x}_0 the estimate of ξ_t computed by means of the Kalman filtering techniques described below in this section, will provide nothing but the same estimated residuals $\hat{\mathbf{e}}_t$ computed in a standard manner from the SVAR model. As for the projected values, the estimated value for the future residuals, i.e., ξ_{T+i} for $i = 0, 1, \dots, P$ would then be zero, as these are assumed zero mean Gaussian iid shocks.

Table 1
Alternative database structures to conduct conditional forecasts.^a

Type	t	Regressors \mathbf{x}_t			Shocks ξ_t		
		$X_{1,t}$	$X_{2,t}$	$X_{3,t}$	$\mathbf{e}_{1,t+1}$	$\mathbf{e}_{2,t+1}$	$\mathbf{e}_{3,t+1}$
CF1	\mathbf{y}'_{T-1}	$X_{1,T-1}$	$X_{2,T-1}$	$X_{3,T-1}$	$\hat{\mathbf{e}}_{1,T}$	$\hat{\mathbf{e}}_{2,T}$	$\hat{\mathbf{e}}_{3,T}$
	\mathbf{y}'_T	$X_{1,T}$	$X_{2,T}$	$X_{3,T}$.	.	.
	\mathbf{y}'_{T+1}

CF2	\mathbf{y}'_{T-1}	$X_{1,T-1}$	$X_{2,T-1}$	$X_{3,T-1}$	$\hat{\mathbf{e}}_{1,T}$	$\hat{\mathbf{e}}_{2,T}$	$\hat{\mathbf{e}}_{3,T}$
	\mathbf{y}'_T	$X_{1,T}$	$X_{2,T}$	$X_{3,T}$	1	.	.
	\mathbf{y}'_{T+1}

CF3	\mathbf{y}'_{T-1}	$X_{1,T-1}$	$X_{2,T-1}$	$X_{3,T-1}$	$\hat{\mathbf{e}}_{1,T}$	$\hat{\mathbf{e}}_{2,T}$	$\hat{\mathbf{e}}_{3,T}$
	\mathbf{y}'_T	$X_{1,T}$	$X_{2,T}$	$X_{3,T}$	$\mathbf{e}_{1,T+1}$	$\mathbf{e}_{2,T+1}$.
	\mathbf{y}'_{T+1}	.	.	.	$\mathbf{e}_{1,T+2}$	$\mathbf{e}_{2,T+2}$.

CF4	\mathbf{y}'_{T-1}	$X_{1,T-1}$	$X_{2,T-1}$	$X_{3,T-1}$	$\hat{\mathbf{e}}_{1,T}$	$\hat{\mathbf{e}}_{2,T}$	$\hat{\mathbf{e}}_{3,T}$
	\mathbf{y}'_T	$X_{1,T}$	$X_{2,T}$	$X_{3,T}$.	$\mathbf{e}_{2,T+1}$.
	\mathbf{y}'_{T+1}	$X_{1,T+1}$.	.	.	$\mathbf{e}_{2,T+2}$.

CF5	\mathbf{y}'_{T-1}	$X_{1,T-1}$	$X_{2,T-1}$	$X_{3,T-1}$	$\hat{\mathbf{e}}_{1,T}$	$\hat{\mathbf{e}}_{2,T}$	$\hat{\mathbf{e}}_{3,T}$
	\mathbf{y}'_T	$X_{1,T}$	$X_{2,T}$	$X_{3,T}$.	.	.
	\mathbf{y}'_{T+1}	$X_{1,T+1}$

CF5	\mathbf{y}'_{T+2}	$\mathbf{e}_{2,T+3}$.
	\mathbf{y}'_{T+P-1}	$\mathbf{e}_{2,T+P}$.
	\mathbf{y}'_{T+P}	$X_{1,T+P}$

^a The symbol '.' is used to indicate missing data, or data which will be estimated. $\hat{\mathbf{e}}_{i,T}$ is used to denote the estimated and identified i -th residual at time T . CF1, ..., CF5 are defined in the text.

2.2. On well-defined conditional forecasts

A few remarks are in order on the different types of conditional forecasts reported in Table 1. In particular, we may very well encounter the problem that we are imposing more restrictions that we are in truth entitled to. For example, it is obvious that fixing the estimated path of all shocks is equivalent to fixing the estimated path of all series. Additionally, the structure of \mathbf{R} and \mathbf{B} may potentially render certain conditional forecasts as simply flawed or ill-posed. For example, if the structure of \mathbf{R} and \mathbf{B} suggests that one series depends solely on one shock, then fixing that shock over the projection horizon is equivalent to fixing the projected path for that series. Were we, under such structure, to fix the values of both that shock and this series, such conditional projection would be ill-posed. Formal conditions for the conditional projection to be well defined and not ill-posed are given below. To do so, system (1) projected P steps ahead can then be written as:

$$\tilde{\mathbf{x}}_T = \mathbf{F}\mathbf{x}_T + \mathbf{M}\tilde{\mathbf{e}}_T, \tag{3}$$

where:

$$\tilde{\mathbf{x}}_T = \begin{pmatrix} \mathbf{x}_{T+1} \\ \mathbf{x}_{T+2} \\ \dots \\ \mathbf{x}_{T+P} \end{pmatrix} \quad \tilde{\mathbf{e}}_T = \begin{pmatrix} \mathbf{e}_{T+1} \\ \mathbf{e}_{T+2} \\ \dots \\ \mathbf{e}_{T+P} \end{pmatrix}$$

$$\mathbf{F} = \begin{pmatrix} \mathbf{B} \\ \mathbf{B}^2 \\ \dots \\ \mathbf{B}^P \end{pmatrix} \quad \mathbf{M} = \begin{pmatrix} \mathbf{R} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{BR} & \mathbf{R} & \dots & \mathbf{0} \\ \dots & \dots & \dots & \mathbf{0} \\ \mathbf{B}^{P-1}\mathbf{R} & \mathbf{B}^{P-2}\mathbf{R} & \dots & \mathbf{R} \end{pmatrix}.$$

Conditional forecasts amount to imposing restrictions on the value of certain elements of $\tilde{\mathbf{x}}_T$ or $\tilde{\mathbf{e}}_T$. Such restrictions will be ill-posed when as a result of them Eq. (3) cannot hold. Note that

by construction \mathbf{M} is a non-singular matrix, and $\mathbf{F}\mathbf{x}_T$ is a constant vector. Therefore, were we not to impose any restrictions on $\tilde{\mathbf{x}}_T$, we would be free to impose any number of restrictions on the value of the elements of $\tilde{\mathbf{e}}_T$. Of course, were we to fix all elements of $\tilde{\mathbf{e}}_T$, the elements of $\tilde{\mathbf{x}}_T$ would be automatically fixed from (3). Were we to impose restrictions in both $\tilde{\mathbf{x}}_T$ and $\tilde{\mathbf{e}}_T$ more caution is warranted. Imagine that we impose r restrictions on the elements of $\tilde{\mathbf{x}}_T$ given by the index vector \mathbf{i}_r , and s restrictions on the elements of $\tilde{\mathbf{e}}_T$ given by the index vector \mathbf{j}_s . Then conditions are ill-posed when one of the following conditions is not met: (i) the number of restrictions $r + s$ is strictly less than the dimension of $\tilde{\mathbf{x}}_T$; (ii) the submatrix of \mathbf{M} formed with those rows not included in \mathbf{i}_r and those columns not included in \mathbf{j}_s is of rank at least $q - r - s$; (iii) the submatrix of \mathbf{M} formed with those rows included in \mathbf{i}_r and those columns not included in \mathbf{j}_s is of full row rank.

2.3. Computation of conditional forecasts

If we further define the so called ‘filtered’ estimate of the state \mathbf{s}_t as $\mathbf{s}_{t|t} = E\{\mathbf{s}_t | \mathbf{y}_1, \dots, \mathbf{y}_t, \boldsymbol{\theta}_0\}$ and its associated covariance matrix as $\mathbf{P}_{t|t}$, the conditional forecasts $\mathbf{s}_{t|T+P}$ and their associated covariance matrix $\mathbf{P}_{t|T+P}$ can be computed using the following implementation of the fixed interval smoothing algorithm of Ansley and Kohn (1982)².

Step 1: Initialization. Set $\mathbf{s}_{T-1|T-1} = \mathbf{y}_{T-1}$ in line with the relevant setting as shown in Table 1; also set $\mathbf{P}_{T-1|T-1} = \mathbf{0}$. Further define for $i = 0, 1, \dots, P$, the $2k \times 2k$ diagonal matrices \mathbf{J}_{t+i} with the j -th diagonal element equal to 1 if the j -th observation of \mathbf{y}_{t+i} is available and zero otherwise.

Step 2: Kalman filtering ‘forward’ iterations. Compute recursively the following equations from $t = T$ until $t = T + P$.

$$\begin{aligned} \mathbf{P}_{t|t-1} &= \mathbf{A}\mathbf{P}_{t-1|t-1}\mathbf{A}' + \boldsymbol{\Sigma}_u \\ \mathbf{K}_t &= \mathbf{P}_{t|t-1}\mathbf{C}'(\mathbf{J}_t\mathbf{C}\mathbf{P}_{t-1|t-1}\mathbf{C}'\mathbf{J}_t)^+ \\ \mathbf{s}_{t|t} &= \mathbf{A}\mathbf{s}_{t-1|t-1} + \mathbf{K}_t(\mathbf{y}_t - \mathbf{C}\mathbf{A}\mathbf{s}_{t-1|t-1}) \\ \mathbf{P}_{t|t} &= \mathbf{P}_{t|t-1} - \mathbf{K}_t\mathbf{C}\mathbf{P}_{t-1|t-1} \end{aligned}$$

and where $^+$ is used to denote the Moore–Penrose inverse.

Step 3: Smoothing ‘backward’ iterations. Compute recursively ‘backwards’ from $t = T + P - 1$ until $t = T$ the following equations.

$$\begin{aligned} \mathbf{s}_{t|T+P} &= \mathbf{s}_{t|t} + \mathbf{H}_t(\mathbf{s}_{t+1|T+P} - \mathbf{A}\mathbf{s}_{t|t}) \\ \mathbf{P}_{t|T+P} &= \mathbf{P}_{t|t} - \mathbf{H}_t(\mathbf{P}_{t+1|t} - \mathbf{P}_{t+1|T+P})\mathbf{H}_t', \end{aligned}$$

where $\mathbf{H}_t = \mathbf{P}_{t|t}\mathbf{A}'\mathbf{P}_{t+1|t}^{-1}$, and $\mathbf{s}_{t|t}$, $\mathbf{P}_{t|t}$ and $\mathbf{P}_{t+1|t}$ are obtained from the Kalman filter iterations of step 2.

2.4. Asymptotic error bands for conditional forecasts

The conditional forecasts $\mathbf{s}_{T+i|T+P}$ and corresponding covariance matrices $\mathbf{P}_{T+i|T+P}$ shown above have been constructed under the assumption that the parameter vector $\boldsymbol{\theta}_0$ is known. This is clearly not satisfactory because the SVAR model is usually estimated. Let us adopt now the notation, which for the sake of

clarity was avoided in the previous sections, that makes the dependence on the parameter vector $\boldsymbol{\theta}$ explicit by using $\mathbf{P}_{T+i|T+P}(\boldsymbol{\theta})$ and $\mathbf{s}_{T+i|T+P}(\boldsymbol{\theta})$. To compute estimates which are not conditioned on the parameter vector being known, Hamilton (1986) proposed the adoption of a Bayesian approach and assumed that rather than being known or fixed, the parameter vector $\boldsymbol{\theta}$ should be best thought as random with known posterior probability distribution $\mathbf{f}(\boldsymbol{\theta}, \mathbf{y}_1, \dots, \mathbf{y}_T)$. Conditional forecasts and error bound estimators which take on board the random nature of $\boldsymbol{\theta}$, i.e., are conditioned exclusively on the data could thus be computed as follows:

$$\begin{aligned} \hat{\mathbf{s}}_{T+i|T+P} &= \int \mathbf{s}_{T+i|T+P}(\boldsymbol{\theta})\mathbf{f}(\boldsymbol{\theta}, \mathbf{y}_1, \dots, \mathbf{y}_T) d\boldsymbol{\theta} \\ \hat{\mathbf{P}}_{T+i|T+P} &= \int [\mathbf{P}_{T+i|T+P}(\boldsymbol{\theta}) + \mathbf{Q}_{T+i|T+P}(\boldsymbol{\theta}) \\ &\quad \times \mathbf{f}(\boldsymbol{\theta}, \mathbf{y}_1, \dots, \mathbf{y}_T) d\boldsymbol{\theta} \end{aligned}$$

and where:

$$\begin{aligned} \mathbf{Q}_{T+i|T+P}(\boldsymbol{\theta}) &= (\mathbf{s}_{T+i|T+P}(\boldsymbol{\theta}) - \hat{\mathbf{s}}_{T+i|T+P}) \\ &\quad \times (\mathbf{s}_{T+i|T+P}(\boldsymbol{\theta}) - \hat{\mathbf{s}}_{T+i|T+P})'. \end{aligned}$$

Rather than computing $\hat{\mathbf{s}}_{T+i|T+P}$ by means of simulations, use of $\mathbf{s}_{T+i|T+P}(\hat{\boldsymbol{\theta}})$, where $\hat{\boldsymbol{\theta}}$ is a maximum likelihood estimator, is a valid approximation. Computation of $\hat{\mathbf{P}}_{T+i|T+P}$ can be done by means of the following algorithm proposed by Hamilton (1986).

Step 1: Simulate the series $\{\boldsymbol{\theta}_1, \boldsymbol{\theta}_2, \dots, \boldsymbol{\theta}_N\}$ where N is a large number.

Step 2: For every $\boldsymbol{\theta}_j$ compute $\mathbf{s}_{T+i|T+P}(\boldsymbol{\theta}_j)$, $\mathbf{P}_{T+i|T+P}(\boldsymbol{\theta}_j)$ and $\mathbf{Q}_{T+i|T+P}(\boldsymbol{\theta}_j)$.

Step 3: The covariance matrix for the conditional forecasts is then computed as:

$$\hat{\mathbf{P}}_{T+i|T+P} = \frac{1}{N} \sum_{j=1}^N [\mathbf{P}_{T+i|T+P}(\boldsymbol{\theta}_j) + \mathbf{Q}_{T+i|T+P}(\boldsymbol{\theta}_j)].$$

Acknowledgments

The author would like to thank Marek Jarocinski for helpful comments in the preparation of this note. Remaining errors remain the responsibility of the author.

References

Ansley, C., Kohn, R., 1982. A geometrical derivation of the fixed interval smoothing algorithm. *Biometrika* 69 (2), 486–487.
 Hamilton, J.D., 1986. A standard error for the estimated state vector of a state-space model. *Journal of Econometrics* 33 (3), 387–397.
 Harvey, A.C., 1993. *Time Series Models*, second ed. Harvester and Wheatsheaf, London.
 Jarocinski, M., 2010. Conditional forecasts and uncertainty about forecast revisions in vector autoregressions. *Economics Letters* 108 (3), 257–259.
 Sims, C., 1980. Macroeconomics and reality. *Econometrica* 48 (1), 1–48.
 Waggoner, D.F., Zha, T., 1999. Conditional Forecasts in dynamic multivariate models. *Review of Economics and Statistics* 81 (4), 639–651.

² For a textbook treatment of the Kalman filter see Harvey (1993).