A Flexible State Space Model and its Applications

By Hang Qian

Department of Economics, Iowa State University, Ames, IA, 50010

Email:matlabist@gmail.com

January 6, 2013

A Flexible State Space Model and its Applications

The standard state space model treats observations as imprecise measurement of the Markovian states. Our flexible model handles the states and observations symmetrically, which are simultaneously determined by past observations and up to first-lagged states. The only distinction between the states and observations is the observability. When it is applied to the ARMA model, the mixed frequency regression, the dynamic factor and the stochastic volatility models, the state space form is both parsimonious and intuitive, for low-dimension states are constructed simply by stacking all the relevant but unobserved components in the structural model.

Key words: Kalman Filter, ARMA, Mixed Frequency, Factor Model, Stochastic Volatility

JEL classification: C32, C51

I. Introduction

Starting with the path-breaking paper of Kalman (1960), the state space model (SSM) has been widely applied in engineering, statistics and economics. Harvey (1991), Hamilton (1994), Durbin and Koopman (2012) present its theory and applications in time series analysis. Basdevant (2003) surveys macroeconomic applications and Mergner (2009) reviews use cases in finance. For practitioners, the art consists in the model building, that is, to cast a structural model into its state space form. The representation is not unique, for one can enlarge the state vector but characterize the same process. The major concern is parsimony and intuitiveness. A parsimonious model with minimum length of the state vector avoids large matrix manipulations, and thus accelerates the Kalman filter. An intuitive form with interpretable states enhances its attractiveness, for predicted and smoothed states bear economic significances.

The states usually refer to unobserved variables with Markovian transition. The observations are imprecise measurement of the states. Since states never echo observations, it is a one-way dependence. Our argument is that the Kalman filter does not necessarily require Markovian states. The recursion is valid as long as no higher than first-lagged states are in the dynamic system, without restrictions on how lagged observations affect current states and observations. That motivates us to bring in more symmetry and two-way dynamics between the states and observations. This feature is most useful when it is combined with the time-varying dimension (TVD) of the state and measurement vectors. The idea of building a flexible SSM is to put all the relevant but unobserved components in the state vector. Since observability of a variable may change over time, the size of the state vector is also dynamic. This often invites lower-dimension states compared with a standard SSM. Furthermore, states in the flexible SSM can always be meaningfully interpreted, for they are simply the unobserved components in the original economic model.

The rest of the paper is organized as follows. Section II. sets up the flexible SSM and Section III. explains the filtering procedure. Section IV. illustrates the TVD feature by an ARMA model in which the state vector shrinks in the initial periods. Section V. and Section VI. consider the mixed frequency vector autoregression and the dynamic factor model with missing data. These two applications explore both features, namely, the TVD and observations in the transition equation. Our state space representations are distinct from those in the literature and have lower-dimension states. Section VII. discusses a stochastic volatility model in which asymmetric volatility is achieved by putting a non-linear function of past observation in the transition equation, in contrast with the traditional approach using correlated disturbances.

II. A FLEXIBLE STATE SPACE MODEL

First consider a standard SSM. Let ξ_t be a $m \times 1$ state vector and Y_t be a $n \times 1$ measurement vector. The dynamic system consists of a transition equation and a measurement equation

$$(1) \quad \begin{aligned} \xi_t &= c_t + F_t \xi_{t-1} + \varepsilon_t, \\ Y_t &= d_t + H_t \xi_t + u_t, \end{aligned}$$

where $\begin{pmatrix} \varepsilon_t \\ u_t \end{pmatrix} \sim N \left[0, \begin{pmatrix} Q_t & S_t \\ S_t' & R_t \end{pmatrix} \right]$. Coefficients $c_t, F_t, d_t, H_t, Q_t, R_t, S_t$ are time-varying but deterministic. The system starts from time 1 and runs through time T with the observations $Y_1^T \equiv \{Y_1, ..., Y_T\}$, which is the information set at time T. The initial state ξ_0 has a known distribution.

The flexible SSM is a moderate generalization of the standard model. Let ξ_t be a $m_t \times 1$ state vector and Y_t be a $n_t \times 1$ measurement vector. They are simultaneously determined by lagged observations and up to first-lagged states:

$$(2) \quad \xi_t = f_t\left(Y_1^{t-1}\right) + F_t \xi_{t-1} + \varepsilon_t,$$
$$Y_t = g_t\left(Y_1^{t-1}\right) + H_t \xi_t + J_t \xi_{t-1} + u_t.$$

where $f_t(\cdot), g_t(\cdot)$ are two linear or non-linear functions that maps the information set of time $t-1$ into $\mathbb{R}^{m_t}$ and $\mathbb{R}^{n_t}$ respectively. In some applications of the flexible SSM, the contemporaneous correlation between ε_t and u_t is essential.

The flexible model has two features.

First, both the state and measurement vectors can change size over time. The TVD of Y_t is well understood and implemented in practice. For example, if some elements of Y_t are missing, the size of the measurement vector is effectively reduced at time t. If all data are missing in that period, the updating step of the Kalman filter is skipped (see Jones, 1980; Harvey and Pierse, 1984). The TVD of ξ_t had been under-appreciated in the literature until recently. Jungbacker et al. (2011) consider a dynamic factor model with missing data. Common factors and idiosyncratic disturbances corresponding to missing data are put in the state vector. Since the amount of missing data varies over time, the state vector is dynamic in size. Chan et al. (2011) explore TVD in a different setting. The model switches to a more parsimonious representation at random dates controlled by hidden Markov-switching regimes. This is a dynamic mixture model with stochastic TVD. Our paper is closer to Jungbacker et al. (2011) in that the dimension changes at known dates.

Second, historical observations $Y_{t-1},\ldots,Y_1$ can affect both Y_t and ξ_t. The dependence of Y_t on $Y_{t-1},\ldots,Y_1$ is well understood. The setup of the SSM in Hamilton (1994, p.372 - 373) includes an $A'x_t$ term in the measurement equation. Hamilton mentions "x_t could include lagged values of y ...", though no application of such feature is provided in the book. In fact, lagged variables in the measurement equation are most useful when they are used together with the TVD feature. Suppose we write $g_t(\cdot)$ as a function of p lagged values $Y_{t-1},\ldots,Y_{t-p}$, we will encounter a problem handling the presample since $Y_0,\ldots,Y_{-p+1}$ are not observed. With the TVD feature, unobserved lagged variables can be temporarily put in the state vector and then removed when data become available.

Allowing lagged observations in the transition equation is rarely seen in the literature. Some may argue that the modeling philosophy of the SSM is to keep the state vector Markovian – summarizing the entire history into the states of yesterday. This argument is not entirely relevant for our model, for we never introduce high-order lagged states $\xi_{t-2}, \xi_{t-3}, \ldots$ in the system, but only allow past observations in the transition equation. In the Kalman filter, ξ_t is predicted and updated conditional on Y_1^{t-1}. Technically, introducing $f_t\left(Y_1^{t-1}\right)$ does not change the filter since it is treated as a constant conditional on Y_1^{t-1}. However, this feature substantially enriches the dependence structure of the SSM. In the standard SSM, ξ_t evolves regardless of Y_t. If we cast a time series model into Eqs. (1), we must ensure the state vector can evolve in a self-sufficient manner. This often entails a larger state vector by including variables that we do observe. However, in the flexible SSM the state vector may temporarily disappear, but reappear later relying on $f_t\left(Y_1^{t-1}\right)$.

III. THE FILTERING PROCEDURE

The forward recursion consists of the prediction step and the update step. Assume $\xi_0 \sim N(\mu_0, \Sigma_0)$. Since Y_1^0 is empty, $\xi_0 \left| Y_1^0 \sim N\left(\widehat{\xi}_{0|0}, P_{0|0}\right)\right.$, where $\widehat{\xi}_{0|0} = \mu_0$, $P_{0|0} = \Sigma_0$.

At time t ($t = 1, \ldots, T$), we predict ξ_t and Y_t conditional on the information set of time $t-1$.

$$\begin{pmatrix} \xi_t \\ Y_t \end{pmatrix} \left| Y_1^{t-1} \sim N \left[\begin{pmatrix} \widehat{\xi}_{t|t-1} \\ \widehat{Y}_{t|t-1} \end{pmatrix}, \begin{pmatrix} P_{t|t-1} & L_{t|t-1} \\ L'_{t|t-1} & D_{t|t-1} \end{pmatrix} \right] \right.,$$

where

$$\begin{aligned}
\widehat{\xi}_{t|t-1} &= f_t\left(Y_1^{t-1}\right) + F_t \widehat{\xi}_{t-1|t-1}, \\
\widehat{Y}_{t|t-1} &= g_t\left(Y_1^{t-1}\right) + H_t \widehat{\xi}_{t|t-1} + J_t \widehat{\xi}_{t-1|t-1}, \\
P_{t|t-1} &= F_t P_{t-1|t-1} F_t' + Q_t, \\
D_{t|t-1} &= H_t P_{t|t-1} H_t' + R_t + J_t P_{t-1|t-1} J_t' + H_t F_t P_{t-1|t-1} J_t' \\
&\quad + J_t P_{t-1|t-1} F_t' H_t' + H_t S_t + S_t' H_t', \\
L_{t|t-1} &= P_{t|t-1} H_t' + F_t P_{t-1|t-1} J_t' + S_t.
\end{aligned}$$

The update step follows from the fact that $\xi_t \left| Y_1^t \sim N\left(\widehat{\xi}_{t|t}, P_{t|t}\right)\right.$, where

$$\widehat{\xi}_{t|t} = \widehat{\xi}_{t|t-1} + L_{t|t-1}\left(D_{t|t-1}\right)^{-1}\left(Y_t - \widehat{Y}_{t|t-1}\right),$$

$$P_{t|t} = P_{t|t-1} - L_{t|t-1}\left(D_{t|t-1}\right)^{-1} L'_{t|t-1}.$$

This completes a recursion cycle and the filter proceeds to the next period.

The TVD feature is embodied in the time-varying size of the matrixes, while the recursion formulae do not change. It is possible that in some period we have no state or measurement vector, which can be interpreted as a zero-dimension column vector (i.e., a 0×1 vector). As long as a programming platform adopts the conformable matrix algebra for empty matrixes[1], the above formulae still apply, though they can be simplified as follows:

If ξ_t is empty, $\widehat{\xi}_{t|t-1}, P_{t|t-1}, L_{t|t-1}, \widehat{\xi}_{t|t}, P_{t|t}$ are empty while $\widehat{Y}_{t|t-1} = g_t\left(Y_1^{t-1}\right) + J_t \widehat{\xi}_{t-1|t-1}$ and $D_{t|t-1} = R_t + J_t P_{t-1|t-1} J'_t$. In other words, the prediction and update on ξ_t are skipped. Note that in the next period, states can reappear relying on past observations. That is, $\xi_{t+1} = f_{t+1}\left(Y_1^t\right) + \varepsilon_{t+1}$.

If Y_t is empty, $\widehat{Y}_{t|t-1}, D_{t|t-1}, L_{t|t-1}$ are empty while $\widehat{\xi}_{t|t} = \widehat{\xi}_{t|t-1}$ and $P_{t|t} = P_{t|t-1}$. In other words, we update the states by making a one-period-ahead prediction. In addition, empty Y_t does not contribute to the likelihood evaluation.

IV. The state space form of ARMA

Let $\{Z_t\}$ be a univariate $ARMA(p,q)$ process

$$Z_t = c + \sum_{i=1}^{p} \phi_i Z_{t-i} + \varepsilon_t + \sum_{i=1}^{q} \theta_i \varepsilon_{t-i},$$

where $\varepsilon_t \sim iidN\left(0, \sigma^2\right)$. There are various ways to write an ARMA model into its state space form. In Akaike (1973, 1974) and Jones (1980), the state vector is chosen as the projection of $Z_t, \ldots, Z_{t+r-1}$ on the information set of time t, where $r \equiv \max(p, q+1)$. Hamilton (1994) explores the fact that the lagged sum of an AR process is an ARMA process. The state vector keeps track of r recent values of a latent $AR(p)$ process. de Jong and Penzer (2004) extend the idea of Pearlman (1980) and propose a canonical form in which the length of the state vector is reduced to $\max(p,q)$.

Our state space representation of the ARMA model distinguishes from the well-known SSMs in three aspects. First, it is more parsimonious. The length of the state vector is q except for the initial p periods when states have dynamic dimensions. Second, it is more general. The well-known SSMs are most suitable for a stationary ARMA process with the initial states coming from the stationary distribution. Our representation can conveniently handle other types of initial distributions and time-varying parameters. Third, it is more intuitive. States simply consist of the disturbance terms and unobserved presample values in the ARMA model.

Let $W_t = \left(Z_t, ..., Z_{t-p+1}, \varepsilon_t, ..., \varepsilon_{t-q+1}\right)'$, $t = 0, ..., T$. Assume $W_0 \sim N(\mu, \Sigma)$. The ARMA literature distinguishes the exact likelihood and the conditional likelihood. The exact likelihood approach assumes W_0 is conformable with the stationary distribution of ARMA process. The conditional likelihood method treats either W_0 or W_p as deterministic. The well-known SSMs are all suitable for exact likelihood evaluation, but apparently have difficulty handling the conditional likelihood since the states are not expressed in terms of Z_t or ε_t. The flexible SSM accommodates both exact and conditional likelihood by properly specifying the initial states.

The state vector in our flexible SSM is

$$\xi_t = \left(Z_0, Z_{-1} ..., Z_{t-p+1}, \varepsilon_t, ..., \varepsilon_{t-q+1}\right)'.$$

By definition, $\xi_0 = W_0$. Note that the state vector shrinks in the initial p periods. Starting from time p, the state vector only contains disturbances $\left(\varepsilon_t, ..., \varepsilon_{t-q+1}\right)'$.

Let $\Phi = (\phi_1, ..., \phi_p)$, $\Theta = \left(\theta_1, ..., \theta_q\right)$, $E_i = \begin{pmatrix} I_i & 0_{i,1} \end{pmatrix}$, where I_i and $0_{i,1}$ are identity and zero matrixes respectively, whose subscripts indicate matrix dimension. Then the transition equation in time $t = 1, ..., p$ is given by

$$\xi_t = \begin{pmatrix} E_{p-t} & 0_{p-t,q} \\ 0_{1,p-t+1} & 0_{1,q} \\ 0_{q-1,p-t+1} & E_{q-1} \end{pmatrix} \xi_{t-1} + \begin{pmatrix} 0_{p-t,1} \\ \varepsilon_t \\ 0_{q-1,1} \end{pmatrix},$$

and the measurement equation is

$$Z_t = c + \sum_{i=1}^{t-1} \phi_i Z_{t-i} + \begin{pmatrix} 0_{1,p-t} & 1 & 0_{1,q-1} \end{pmatrix} \xi_t + (\phi_t, ..., \phi_p, \Theta)\, \xi_{t-1}.$$

Note that at time $t = p$, $E_{p-t}, 0_{p-t,q}, 0_{p-t,1}, 0_{1,p-t}$ are empty, but the formulae still apply.

For time $t = p+1,...,T$, the state space form has time-invariant parameters and dimensions:

$$\xi_t = \begin{pmatrix} 0_{1,q} \\ E_{q-1} \end{pmatrix} \xi_{t-1} + \begin{pmatrix} \varepsilon_t \\ 0_{q-1,1} \end{pmatrix},$$
$$Z_t = c + \sum_{i=1}^{p} \phi_i Z_{t-i} + \begin{pmatrix} 1 & 0_{1,q-1} \end{pmatrix} \xi_t + \Theta \xi_{t-1}.$$

Suppose the distribution of W_0 is explicitly specified (as in the case of the conditional likelihood), we can immediately apply the flexible SSM. However, we often do not articulate the initial distribution but require W_0 generated from the stationary distribution (as in the case of the exact likelihood). Note that the transition matrix at time 1 is not square, so we cannot invert it to generate a stationary distribution. However, we can make it square by temporarily expanding $\xi_1 = W_1$ so that $\xi_1 = c_1 + F_1 \xi_0 + \widetilde{\varepsilon}_1$, where

$$c_1 = \begin{pmatrix} c \\ 0_{p+q-1,1} \end{pmatrix}, F_1 = \begin{pmatrix} \Phi & \Theta \\ E_{p-1} & 0_{p-1,q} \\ 0_{1,p} & 0_{1,q} \\ 0_{q-1,p} & E_{q-1} \end{pmatrix}, \widetilde{\varepsilon}_1 = \begin{pmatrix} \varepsilon_t \\ 0_{p-t,1} \\ \varepsilon_t \\ 0_{q-1,1} \end{pmatrix}.$$

Then the stationary distribution can be generated by

$$E(\xi_0) = \left(I_{(p+q)} - F_1\right)^{-1} c_1, \tag{3}$$
$$vec\left[Var(\xi_0)\right] = \left(I_{(p+q)^2} - F_1 \otimes F_1\right)^{-1} vec(Q_1).$$

where Q_1 is the covariance matrix of $\widetilde{\varepsilon}_1$, that is, a $(p+q)\times(p+q)$ matrix of zeros except for $(1,1)$, $(1,p+1)$, $(p+1,1)$,$(p+1,p+1)$ elements being σ^2.

In summary, our flexible SSM employs TVD states to handle the initial distribution, but from time $p+1$ to T, the fixed-length state vector only tracks the MA part of the series. The AR part is predetermined and thus put in the measurement equation.

V. MIXED FREQUENCY REGRESSION

One feature of the flexible SSM is that lagged observations in the transition equation, bringing in richer dynamics between the states and observables. We illustrate its usage by a mixed frequency vector autoregression (VAR) model. Temporal aggregation in the state space framework is explored by Zadrozny (1988), Mittnik and Zadrozny (2004), Mariano and Murasawa (2003, 2010), Hyung and Granger (2008).

Consider a bivariate $VAR(1)$ model

$$\begin{pmatrix} Z_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} + \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \begin{pmatrix} Z_{1,t-1} \\ Z_{2,t-1} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix}.$$

We can observe $\{Z_{2,t}\}$, but $\{Z_{1,t}\}$ is aggregated every other period, namely $\overline{Z}_{1,t} = Z_{1,t-1} + Z_{1,t}$, $t = 2,4,6,...,T$.

In the standard SSM, the state vector tracks variables in recent two periods, namely $\xi_t = (Z_{1,t}, Z_{2,t}, Z_{1,t-1}, Z_{2,t-1})'$. The transition and measurement equations are explained in detail in Mariano and Murasawa (2003, 2010). The flexible SSM only admits unobserved variables in the state vector. Let $\xi_0 = (Z_{1,0}, Z_{2,0})$ and $\xi_t = Z_{1,t}$ for all $t = 1,...,T$.

For $t = 1$, the transition and measurement equations are given by

$$\begin{aligned} \xi_1 &= c_1 + \begin{pmatrix} \phi_{11} & \phi_{12} \end{pmatrix} \xi_0 + \varepsilon_{1,1}, \\ Z_{2,1} &= c_2 + \begin{pmatrix} \phi_{21} & \phi_{22} \end{pmatrix} \xi_0 + \varepsilon_{2,1}. \end{aligned}$$

In the following periods, the transition equation is

$$\xi_t = c_1 + \phi_{12} Z_{2,t-1} + \phi_{11} \xi_{t-1} + \varepsilon_{1,t},$$

and the measurement equation in the odd-numbered periods is given by

$$Z_{2,t} = c_2 + \phi_{22} Z_{2,t-1} + \phi_{21} \xi_{t-1} + \varepsilon_{2,t}.$$

and in the even-numbered periods is

$$\begin{pmatrix} \overline{Z}_{1,t} \\ Z_{2,t} \end{pmatrix} = \begin{pmatrix} 0 \\ c_2 + \phi_{22} Z_{2,t-1} \end{pmatrix} + \begin{pmatrix} 1 \\ 0 \end{pmatrix} \xi_t + \begin{pmatrix} 1 \\ \phi_{21} \end{pmatrix} \xi_{t-1} + \begin{pmatrix} 0 \\ \varepsilon_{2,t} \end{pmatrix}.$$

In the standard SSM, the state vector has four dimensions, while the flexible SSM only keeps track of the scalar $Z_{1,t}$ as the state vector except in the first period. The state and measurement equations simply replicate the original $VAR(1)$ process and the aggregation constraint.

VI. Dynamic factor model with missing data

Factor models have wide applications in macroeconomic forecasting (e.g., Stock and Watson, 2002; Forni et al., 2003; Schumacher, 2007), monetary policy analysis (Bernanke et al., 2005; Stock and Watson, 2005) and business cycle transmission study (Eickmeier, 2007). We consider a factor model with randomly missing data similar to Jungbacker et al. (2011), but propose a more parsimonious state space representation.

Let Y_t be a $n \times 1$ vector of observations, determined by a $m \times 1$ vector of common factors f_t and a $n \times 1$ vector of idiosyncratic terms v_t such that

$$Y_t = \Lambda f_t + v_t. \tag{4}$$

Both common factors and idiosyncratic components follow AR(1) processes

$$f_t = F f_{t-1} + \varepsilon_t,$$
$$v_t = \Phi v_{t-1} + u_t,$$

where $\varepsilon_t \sim iidN(0, Q)$ and $u_t \sim iidN(0, R)$.

We follow the notations of Jungbacker et al. (2011) in handling missing data in Y_t. Consider an $n \times 1$ vector Z_t. The vector $Z_t(o_s)$ contains all elements of Z_t that correspond to observed entries in Y_s $(t, s = 1, \ldots, T)$. In other words, o_s is a logical index indicating the observed entries in Y_s and we use o_s to select corresponding elements in Z_t. Similarly, $Z_t(m_s)$ contains all elements of Z_t that correspond to missing entries in Y_s. We can also use logical indexes to extract corresponding rows and/or columns of a $n \times n$ matrix A. For example, $A(o_s, :)$ denotes row selections, $A(:, o_s)$ denotes column selections, and $A(m_s, o_s)$ denotes both row and column selections.

In principle, we can track both f_t and v_t as states and straightforwardly write the model into the state space form. However, v_t is of length n, which is typically much larger than m. It is inconvenient to work on a high-dimension SSM. Jungbacker et al. (2011) solve this problem by partially differencing Eq. (4),

$$Y_t = \Phi Y_{t-1} + \Lambda f_t - \Phi \Lambda f_{t-1} + u_t, \tag{5}$$

If data are available in both Y_t and Y_{t-1}, Eq. (5) serves as the measurement equation with f_t as states. Otherwise, the measurement equation is switched to Eq. (4) with both f_t and v_t in the state vector.

Our flexible SSM solely relies on Eq. (5) as the measurement equation and v_t never enters the state vector. First rewrite Eq. (5) as

(6) $$Y_t = \Phi Y_{t-1} + G f_{t-1} + w_t,$$

where $G = \Lambda F - \Phi\Lambda$, $w_t = \Lambda\varepsilon_t + u_t$. $\begin{pmatrix} \varepsilon_t \\ w_t \end{pmatrix} \sim N\left[0, \begin{pmatrix} Q & Q\Lambda' \\ \Lambda Q & \Lambda Q\Lambda' + R \end{pmatrix}\right]$.

Note that Y_{t-1} can be decomposed into $Y_{t-1}(o_{t-1})$ and $Y_{t-1}(m_{t-1})$. Eq. (6) implies that Y_t is determined by $Y_{t-1}(o_{t-1})$, $Y_{t-1}(m_{t-1})$ and f_{t-1}. The first one is predetermined, while the last two are exactly the state vector of time $t-1$. It follows that the measurement equation is given by

$$\begin{aligned} Y_t(o_t) &= \Phi(o_t, o_{t-1}) Y_{t-1}(o_{t-1}) \\ &\quad + \begin{bmatrix} G(o_t,:) & \Phi(o_t, m_{t-1}) \end{bmatrix} \begin{bmatrix} f_{t-1} \\ Y_{t-1}(m_{t-1}) \end{bmatrix} + w_t(o_t), \end{aligned}$$

and the transition equation is

$$\begin{aligned} \begin{pmatrix} f_t \\ Y_t(m_t) \end{pmatrix} &= \begin{pmatrix} 0 \\ \Phi(m_t, o_{t-1}) Y_{t-1}(o_{t-1}) \end{pmatrix} \\ &\quad + \begin{bmatrix} F_t & 0 \\ G(m_t,:) & \Phi(m_t, m_{t-1}) \end{bmatrix} \begin{bmatrix} f_{t-1} \\ Y_{t-1}(m_{t-1}) \end{bmatrix} + \begin{bmatrix} \varepsilon_t \\ w_t(m_t) \end{bmatrix}. \end{aligned}$$

Compared with the state space representation of Jungbacker et al. (2011), our flexible SSM represents the same process but has some advantages. First, the state vector is shorter. Second, no restrictions are put on Φ. The transition equation presented in Jungbacker et al. (2011) is based on a diagonal Φ so that $v_t(m_{t-1})$ only depends on $v_{t-1}(m_{t-1})$. Non-diagonal extension is possible but will be cumbersome. Third, states need not to be reshuffled. In Jungbacker et al. (2011), a selection matrix is employed to re-order the states to facilitate transition. Fourth, the representation is intuitive. The transition and measurement equation are symmetric and they largely resemble Eq. (4) and Eq. (6). The elements in Y_t, no matter as the states or observations, always get access to Y_{t-1} partially from the past observations and partially from the previous states.

VII. ASYMMETRIC STOCHASTIC VOLATILITY

Conditional heteroskedasticity models are widely used in financial volatility forecasting. Empirical evidence from Kim et al. (1998), Danielsson (1998), Yu (2002) suggests that stochastic volatility (SV) models often outperform GARCH models in characterizing the stylized fact of volatility clustering. Another stylized feature that both models attempt to capture is the asymmetric volatility due to the leverage effect. EGARCH by Nelson (1991) and GJR-GARCH by Glosten et al. (1993) allow (a non-linear function of) signed returns affect the conditional variance. However, SV models typically accommodate asymmetry by negatively correlated disturbances. See, among others, Harvey and Shephard (1996), Jacquier et al. (2004), Kirby (2006). By the Cholesky decomposition, two correlated variables can be represented by linear combinations of two uncorrelated variables. Therefore, SV models with correlated innovations allow a specific form that signed returns affect volatility, usually in a linear manner. See Eq. (7) in Harvey and Shephard (1996) and Eq. (2.3) in Yu (2005).

The flexible SSM supports any function of past observations in the transition equation. Therefore, our version of the SV model can characterize asymmetric volatility as flexibly as in the GARCH model. Suppose the observed returns $\{r_t\}$ are generated by

$$r_t = \sigma_t \cdot z_t, \tag{7}$$

$$\ln \sigma_t^2 = f(r_{t-1}) + \gamma \ln \sigma_{t-1}^2 + u_t, \tag{8}$$

where $z_t \sim iidN(0,1)$, $u_t \sim iidN\left(0,\lambda^2\right)$ and z_t, u_t are mutually independent. $f(r_{t-1})$ can be any non-linear function of past returns. For example, similar to a GJR-GARCH model we can set $f(r_{t-1}) = \alpha r_{t-1}^2 + \beta r_{t-1}^2 I(r_{t-1} < 0)$, where $I(\cdot)$ is an indicator function.

The essence of Harvey and Shephard (1996)'s state space method is the inference conditional on the signs of the observations. We also explore this idea. Imagine that z_t is generated by two steps, first a Bernoulli draw to determine its sign, and then an independent draw from the half normal to determine its magnitude. Similarly, an observed r_t can also be interpreted as two observations, its sign s_t and its monotone transformed magnitude $y_t \equiv \ln r_t^2$. The former is generated by an independent Bernoulli draw, and the latter is generated according to Eq. (7)

$$y_t = \ln \sigma_t^2 + \ln z_t^2.$$

That is, y_t is a draw from the log of $\chi^2(1)$, shifted by $\ln \sigma_t^2$. The normality approximation of this system is proposed by Harvey et al. (1994),

$$(9) \qquad y_t = -1.27 + \ln \sigma_t^2 + v_t,$$

where $v_t \sim iidN\left(0, \frac{\pi^2}{2}\right)$.

For the purpose of parameter estimation by quasi-maximum likelihood, the likelihood of $\{y_t, s_t\}$ is proportional to that of $\{y_t\}$ conditional on realized $\{s_t\}$. In other words, $\{s_t\}$ can be treated as if an exogenous series. If we let $\ln \sigma_t^2$ be the state variable, Eq. (9) serves as the measurement equation. The transition equation can be obtained from Eq. (8),

$$\ln \sigma_t^2 = f\left(s_{t-1}\sqrt{e^{y_{t-1}}}\right) + \gamma \ln \sigma_{t-1}^2 + u_t.$$

Compared with the existing state space approaches to the asymmetric SV model, our formulation is both general and straightforward. Since asymmetry is not embodied in the correlation between disturbances, there is no need to compute the disturbance moments conditional on the sign s_t, which is usually cumbersome as in Harvey and Shephard (1996).

VIII. Conclusion

In the standard SSM, the states are detached from observations due to its own autoregression. The observations are noise-ridden representation of the states. In this paper, the SSM is examined from a new angle. Our SSM is flexible due to the symmetry and two-way dynamics between the states and observations. This feature merits concise translation from a structure model to its state space form. In additional, the translation is straightforward. Relevant but unobserved components are placed in the state vector while all observables are in the measurement vector. The number of unobserved (observed) variables in the model may vary over time, so the length of the state (measurement) vector is also dynamic.

Despite the differences in interpreting the system dynamics, the same Kalman filter can be applied to both the standard and flexible SSM. The latter is more concise and thus the Kalman filter is expected to run faster.

REFERENCES

Akaike, H., 1973. Maximum likelihood identification of gaussian autoregressive moving average models. Biometrika 60 (2), 255–265.

Akaike, H., 1974. Markovian representation of stochastic processes and its application to the analysis of autoregressive moving average processes. Annals of the Institute of Statistical Mathematics 26, 363–387.

Basdevant, O., 2003. On applications of state-space modelling in macroeconomics. Reserve Bank of New Zealand Discussion Paper Series.

Bernanke, B., Boivin, J., Eliasz, P. S., 2005. Measuring the effects of monetary policy: A factor-augmented vector autoregressive (favar) approach. The Quarterly Journal of Economics 120 (1), 387–422.

Chan, J., Koop, G., Leon-Gonzalez, R., Strachan, R., 2011. Time varying dimension models. Working papers, University of Strathclyde Business School, Department of Economics.

Danielsson, J., 1998. Multivariate stochastic volatility models: Estimation and a comparison with vgarch models. Journal of Empirical Finance 5 (2), 155–173.

de Jong, P., Penzer, J., 2004. The arma model in state space form. Statistics and Probability Letters 70 (1), 119–125.

Durbin, J., Koopman, S. J., 2012. Time Series Analysis by State Space Methods: Second Edition. Oxford University Press.

Eickmeier, S., 2007. Business cycle transmission from the us to germany–a structural factor approach. European Economic Review 51 (3), 521–551.

Forni, M., Hallin, M., Lippi, M., Reichlin, L., 2003. Do financial variables help forecasting inflation and real activity in the euro area? Journal of Monetary Economics 50 (6), 1243–

1255.

Glosten, L. R., Jagannathan, R., Runkle, D. E., 1993. On the relation between the expected value and the volatility of the nominal excess return on stocks. Journal of Finance 48 (5), 1779–1801.

Hamilton, J. D., 1994. Time Series Analysis. Princeton University Press: Princeton.

Harvey, A. C., 1991. Forecasting, Structural Time Series Models and the Kalman Filter. Cambridge University Press.

Harvey, A. C., Pierse, R. G., 1984. Estimating missing observations in economic time series. Journal of the American Statistical Association 79 (385), 125–131.

Harvey, A. C., Shephard, N., 1996. Estimation of an asymmetric stochastic volatility model for asset returns. Journal of Business & Economic Statistics 14 (4), 429–34.

Hyung, N., Granger, C. W., 2008. Linking series generated at different frequencies. Journal of Forecasting 27 (2), 95–108.

Jacquier, E., Polson, N. G., Rossi, P. E., 2004. Bayesian analysis of stochastic volatility models with fat-tails and correlated errors. Journal of Econometrics 122 (1), 185–212.

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

Jungbacker, B., Koopman, S., van der Wel, M., 2011. Maximum likelihood estimation for dynamic factor models with missing data. Journal of Economic Dynamics and Control 35 (8), 1358–1368.

Kalman, R. E., 1960. A New Approach to Linear Filtering and Prediction Problems. Transactions of the ASME,Journal of Basic Engineering, 35–45.

Kim, S., Shephard, N., Chib, S., 1998. Stochastic volatility: Likelihood inference and comparison with arch models. Review of Economic Studies 65 (3), 361–93.

Kirby, C., 2006. Linear filtering for asymmetric stochastic volatility models. Economics Letters 92 (2), 284–292.

Mariano, R. S., Murasawa, Y., 2003. A new coincident index of business cycles based on monthly and quarterly series. Journal of Applied Econometrics 18 (4), 427–443.

Mariano, R. S., Murasawa, Y., 2010. A coincident index, common factors, and monthly real gdp. Oxford Bulletin of Economics and Statistics 72 (1), 27–46.

Mergner, S., 2009. Applications of State Space Models in Finance. Universitatsverlang Gottingen.

Mittnik, S., Zadrozny, P. A., 2004. Forecasting quarterly german gdp at monthly intervals using monthly ifo business conditions data. CESifo Working Paper Series.

Nelson, D. B., 1991. Conditional heteroskedasticity in asset returns: A new approach. Econometrica 59 (2), 347–70.

Pearlman, J. G., 1980. An algorithm for the exact likelihood of a high-order autoregressive-moving average process. Biometrika 67 (1), 232–233.

Schumacher, C., 2007. Forecasting german gdp using alternative factor models based on large datasets. Journal of Forecasting 26 (4), 271–302.

Stock, J. H., Watson, M. W., 2002. Macroeconomic forecasting using diffusion indexes. Journal of Business & Economic Statistics 20 (2), 147–62.

Stock, J. H., Watson, M. W., 2005. Implications of dynamic factor models for var analysis. NBER Working Papers 11467, National Bureau of Economic Research, Inc.

Yu, J., 2002. Forecasting volatility in the new zealand stock market. Applied Financial Economics 12 (3), 193–202.

Yu, J., 2005. On leverage in a stochastic volatility model. Journal of Econometrics 127 (2), 165–178.

Zadrozny, P., 1988. Gaussian likelihood of continuous-time armax models when data are stocks and flows at different frequencies. Econometric Theory 4 (01), 108–124.