Research article

Specification of initial Kalman recursions of symmetric nonlinear state-space model

M. Tasi'u a,*, H.G. Dikko a, O.I. Shittu b, I.A. Fulatan c

a Department of Statistics, Ahmadu Bello University, Zaria Nigeria
b Department of Statistics, University of Ibadan Nigeria
c Department of Mathematics, Ahmadu Bello University, Zaria Nigeria

ARTICLE INFO

Keywords:
Mathematics
Statistics
Time Series analysis
Nonlinear models
State-Space model
Kalman filter
Kalman gain
Kalman recursion
Vec operator
Kronecker products
System update

ABSTRACT

A new class of nonlinear Time Series model referred to as Symmetric Nonlinear State-Space Model (SNSSM) was successfully developed using Kalman filter methodology. Some vital properties of the SNSSM such as optimal Kalman gain and optimal filter state covariance were derived. We finally initialized the filter which enabled us obtained the initial Kalman recursions under stationarity and nonstationarity assumptions. Under the former, the mean and variance were obtained unconditionally using Kronecker products and vec operator. But under the later, the mean and variance/covariance of the system were conditionally obtained using a well-known marginal and conditional property of multivariate normal distribution. It is expected that the former will be better than the later if the system is stationary, otherwise the later will be better.

1. Introduction

A State-Space Model (SSM) consists of a transition/state equation and a measurement equation. The transition equation formulates the dynamics of the state variables and the measurement equation relates the observed variables to the unobserved transition vector. The state vector can contain trend, seasonal, cyclical and regression components together with an error term also known as innovation. However, the stochastic behavior of the state variable, its association to the data and the covariance structure of the errors depend on parameters that are almost always unknown; (Sascha, 2009). The word 'Kalman' was from the original author/inventor of the filter, Rudolf Kalman (1930–2016). So this is nothing but a name given to filters in this concept. Basically, all filters share a common attribute: to give opportunity for something to pass through, at the same time seize the opportunity for something to pass through. A Kalman filter also acts as a filter, but its operation is somewhat more complex and harder to understand! Rhudy et al. (2017).

Researchers from different fields across the world are contributing immensely to the development of the State-Space/Kalman filter models both theoretically and emphatically. Theoretically for example, Hamilton (1994) gave a State-Space representation of a linear dynamic system. The wisdom behind this representation is to capture all the dynamics of the unobserved measurement vector \( Y_t \) in terms of unobserved state vector, say \( X_t \). He imposed some restrictions on the parameters of the measurement vector that would ensure the stability of the process. He further proposed a general form of linear State-Space model with a constant parameter; and he derived an optimal forecast of the system via a well-established result for normal variables; all the needed components of the linear Kalman filter algorithm have been derived. The major limitation of the Hamilton's work on the State-Space modeling and Kalman filtering is the assumption of linearity, the frame work was designed to

* Corresponding author.
E-mail address: dagastatistician@gmail.com (M. Tasi'u).

https://doi.org/10.1016/j.heliyon.2020.e05152
Received 24 January 2020; Received in revised form 17 July 2020; Accepted 30 September 2020
2405-8440/© 2020 Published by Elsevier Ltd. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).
handle a linear or approximately linear system, as such, it cannot handle any nonlinear system.

Extended Kalman Filter (EKF), Particle Filter (PF), and Unscented Kalman Filter (UKF), are the most commonly used techniques to estimate the state of a nonlinear dynamic system (Gao et al., 2018) and (Hu et al., 2020). The EKF is an approximate technique that utilized the first-order Taylor series expansion to linearize the nonlinear dynamic system so that the classical Kalman filter can be applied. It was however observed that if the nonlinear dynamic model to linearize has a high degree of nonlinearity, the linearization usually results to a very large estimation error which make the filtering solution of the EKF unsuitable as well as unreliable (Del Moral, 1996) and (Hu et al., 2020). To overcome the limitations of the EKF, Particle Filter (PF) was proposed.

The PF (which is sometimes called Sequential Monte Carlo (SMC)) is a recursive Bayesian optimal filtering technique that approximates the posteriori distribution of a state of dynamic system (havven large amount of particles) based on Monte Carlo algorithm (Gao et al., 2018) and (Anons, 2019). However, computational complexity of the PF as a result of the large number of particles (samples) required in the process of resampling serves as its major drawback. To handle the limitations of both EKF and PF, the Unscented Kalman Filter (UKF) was proposed.

The UKF is a member of a set of filters called Linear Regression Filters also known as Sigma-Point Filters (Terejanu, 2019). This method forms a linear regression between $n$ sample (sigma) points drawn from prior distribution of a random variable to linearize a nonlinear function of the random variable. The UKF is regarded as a derivative-free alternative to the EKF because it uses a deterministic sampling approach to linearize the nonlinear dynamic system (Wan and van der Merwe, 2019).

Comparing to EKF, the estimation accuracy and convergence rate of the UKF is higher. Comparing to PF, the implementation of the UKF is simpler and its computation is easier (Gao et al., 2018). It was however reported that the implementation of the UKF requires that system models are exactly pre-defined; therefore if the said models involve uncertainties, the estimation of UKF will be deteriorated (Hu et al., 2020).

Similarly, Antti (2001), proposed a State-Space model that is nonlinear in nature; the model is called Switching Nonlinear State-State model (Switching Nonlinear SSM) which is a combination of the SSM and Hidden Markov Model (HMM). As we all know, the state vector is continuous in the SSM where as it is discrete in the HMM. In this case, the author designed the SSM part to model the short-term changes/dynamics of the data, while the HMM describes the long-term dynamics of the data and controls the SSM. He further described the development of a switching Nonlinear SSM and a Bayesian learning algorithm for its parameters. It uses multilayer perception networks to model the nonlinear function of the SSM. In the same vain; the author stated that “computational complexity of the Nonlinear SSM algorithm sets serious limitation for the Bayesian switching SSM for most practical uses and its learning algorithm is very slow”. In addition to the above stated limitations by the author, the model frame work is more general! It did not take into consideration the nature of series/data (such as symmetry or asymmetry); and avoiding this fact of great importance would lead to drawing a very wrong inference in the system under study.

A class of robust Kalman like filters was also developed by many researchers such as Zorzi (2017), Levy and Nikoukhah (2013), Levy and Zorzi (2016), Zorzi and Levy (2015), among other. These filters were designed to avoid large errors usually encountered in the implementation of the classical Kalman filters. These robust Kalman like filters are performing according to risk sensitive approach. The sensitivity to large errors was tuned by what is known as risk sensitivity parameter. In addition to the risk sensitive approach, a divergence-based minimax approaches were also been proposed to perform the robust filtering, where the robust filter is obtained by minimizing the mean square error according to least favorable model in a bounded continuum (ball) on the Tau-divergence family. The gain matrix of these kind of filters is updated according to the risk sensitivity.

In a quest for Kalman filter like algorithm that will model and predict Time Series characterized by a Regime-Switching behavior, Raphael (2016), proposed another nonlinear State-Space model known as Modified State-Space Model (MSSM). The MSSM allows for the introduction of nonlinear function: Logistic Smooth Transition Autoregressive (LSTAR) model in the state equation of the CSSM and this transformed the CSSM from linear to nonlinear model. The basic limitation of the Raphael’s work is the asymmetric behavior of its state equation (which is inappropriate for modelling many financial series such as exchange rates). We therefore addressed this challenge by developing a nonlinear Time Series model with symmetric state equation known as Symmetric Nonlinear State-Space Model (SNSSM). Hence, the aim of this paper is to specify the initial equations of the developed system which would enable us obtain the Kalman recursions of the SNSSM under stationary and nonstationary situations.

2. Methodology

In this section, we would sequentially demonstrate the procedure upon which we followed to develop the SNSSM. To enable the reader(s) grasp the idea effectively, we introduced the existing and classical methods given in section 2.1. The developed methodology: SNSSM is presented in section 2.2; finally, the initial equations of the SNSSM were proposed/specify in section 3.

2.1. Classical state-space models (CSSM)

As stated in section 1 above, the State-Space model is a system of two equations as given in (1)(1) and (2)(2):

$$Y_{t+1} = HX_{t+1} + V_{t+1}$$  
(1)

$$X_{t+1} = \psi X_t + \omega_{t+1}$$  
(2)

Eq. (1) called measurement (observation) equation, describes the relation between the observed Time Series, $Y_{t+1}$ and the (possibly unobserved) state $X_{t+1}$. Eq. (2) called the State (transition) equation, describes the evolution of the state variables as being driven by the stochastic process of innovations $\omega_{t+1}$.

The terms $V_{t+1}$ and $\omega_{t+1}$ are the measurement and the process noise respectively. Usually one assumes normal innovations, such that $V_{t+1} \sim N(0, \sigma_v^2)$ and $\omega_{t+1} \sim N(0, \sigma_\omega^2)$. Similarly, these error terms $V_{t+1}$ and $\omega_{t+1}$ are assumed to be serially independent and independent of each other at all time periods as well as uncorrelated with the initial state. The role of $V_{t+1}$ in (1) is to account for any uncertainty in the measurement of the output (i. e. it tells us how much or little we can trust the equation). The parameter $H$ is an unknown that links the unobservable variables and regression effects of the state equation with the observation equation, $\psi$ is an unknown parameter that determines how the observation and state equations evolve (change) in time (Yu, 2015).

Moreover, one can decides to look at (1)(1) and (2)(2) in multivariate perception which can subsequently be written as

$$Y_{t+1} = HX_{t+1} + V_{t+1}$$  
(3)

$$X_{t+1} = \psi X_t + \Omega_{t+1}$$  
(4)

In this case, $E(V_{t+1}Y_{t+1}) = L$ and $E(\Omega_{t+1}\Omega_{t+1}^T) = Z$

However, the matrices $V_{t+1}$ and $\Omega_{t+1}$ are not really implemented/ included in evaluations of (3)(3) and (4)(4) because they are assumed to be random innovations with zero mean, but instead are always used in determination of any information about the observation and state error covariance matrices $L$ and $Z$. The system matrices $H$ and $\psi$ are in general vary with time, but would not change with respect to states/transition.
In most cases, they are regarded as constants. \( \psi \) Contains the coefficients of the transition terms in (4) and \( H \) performs similar task in (3) (Rhudy et al., 2017).

As stated earlier, the MMSM of Raphael (2016) was developed with the aid of LSTAR model: a family of Smooth Transition Autoregressive (STAR) model. The STAR model is a nonlinear Time Series model that allows for regime-switching behavior through a continuous transition function. Alenka et al. (2014), Terasvirla (1994) and Zhou (2010). In this context, the data-generating process to be modeled is viewed as a linear process that switches between numbers of regimes according to some rules. It has been assumed that there is a continuum of switches, that is, there is a smooth transition from one extreme regime to the other. It consists of three stages: specification, estimation and evaluation; Iqubeal (2016), and Usman et al. (2018). The STAR model of order \( p \) is given as,

\[
x_{t+1} = \psi x_t [1 - G(x_{t-\delta}; \gamma, C)] + \psi x_t G(x_{t-\delta}; \gamma, C) + \omega_{t+1}
\]

where \( x_t = (1, x_{t-1}, x_{t-2}, ..., x_{t-p})' \), \( \psi = (\psi_0, \psi_1, ..., \psi_p)' \) and \( J = \{1, 2, \ldots, n\} \), \( G(x_{t-\delta}; \gamma, C) \) is known as transition function which allows the model to switch between different regimes smoothly. It is bounded between zero and one, i.e., \( 0 \leq G(x_{t-\delta}; \gamma, C) \leq 1 \). This property also makes Eq. (5) capable to explain the two extreme states as well as a continuum of states lie in-between these extreme states/ regimes. The variable \( x_{t-\delta} \) is known as transition variable and \( \delta \) is delay parameter. \( \gamma > 0 \), is a smoothness parameter for the transition function \( G(x_{t-\delta}; \gamma, C) \). \( C \) is a location or threshold parameter and it represents the point of transition between the two extreme regimes. Transition function, \( G(x_{t-\delta}; \gamma, C) \) causes the nonlinear dynamics in the model, and can have different functional choices. For each choice of transition function switching behavior; Yaya and Shittu (2016). The most common choices are logistic and exponential forms as given in Eqs. (6) and (7) respectively.

\[
G(x_{t-\delta}; \gamma, C) = \frac{1}{1 + \exp [\gamma (x_{t-\delta} - C)]}
\]

(6)

\[
G(x_{t-\delta}; \gamma, C) = 1 - \exp [\gamma (x_{t-\delta} - C)^2]
\]

(7)

Note: If (6) is considered as \( G(x_{t-\delta}; \gamma, C) \) in (5), then (5) is called Logistic Smooth Transition Autoregressive (LSTAR) model. Similarly, if (7) is considered as \( G(x_{t-\delta}; \gamma, C) \) in (5), then (5) is called Exponential Smooth Transition Autoregressive (ESTAR) model.

Comparing between the two transition functions: (6(6) and (7)(7), the logistic is changing monotonically with \( X_{t-\delta} \), while the exponential is changing symmetrically at \( C \) with \( X_{t-\delta} \). To visualize the asymmetric and symmetric features of the two transition functions: logistic and exponential, see Figure 1 below.

2.2. Symmetric nonlinear state-space model (SNSSM)

In order to handle the shortcomings of the MSSM highlighted earlier, we modified our system of equations given in (1(1) and (2(2) by replacing the state equation: (2) of the system with the STAR model given in (5) and considering (7) as a transition function assuming that \( p = 2 \) (for crystal clear). Hence, the new state equation is given as:

\[
X_{t+1} = \psi_1 X_t (1 - e^{-\gamma (X_t - C)}) + \psi_2 X_t (1 - e^{-\gamma (X_t - C)^2}) + \omega_{t+1}
\]

(8)

Note that when \( \gamma \to \infty \), then Eq. (8) becomes Eq. (2) which is linear; the same situation happens when \( \gamma \to 0 \).

Differentiating Eq. (8) partially with respect to the previous state \( X_t \), we have

\[
\frac{dX_{t+1}}{dX_t} = \psi_1 X_t (-2 \gamma (X_t - C)) e^{-\gamma (X_t - C)} + \psi_2 e^{-\gamma (X_t - C)^2}
\]

(9)

\[
\frac{dX_{t+1}}{dX_t} = \psi_1 (2 \gamma X_t (X_t - C)) e^{-\gamma (X_t - C)^2} + \psi_2 (1 - e^{-\gamma (X_t - C)^2})
\]

which gives

\[
= -2 \psi_1 \gamma X_t (X_t - C) e^{-\gamma (X_t - C)^2} + \psi_2 (1 - e^{-\gamma (X_t - C)^2})
\]

(10)

expanding and equation to zero, we have

\[
(\psi_1 - \psi_2)(1 - G(X_t)) + 2 \gamma X_t^2 + 2 \gamma CX_t + 1 + \psi_2 = 0
\]

(11)

simplifying further we have

\[
-2 \gamma (X_t - C)^2 = \frac{-(\psi_1 - \psi_2)(1 - G(X_t))}{(\psi_1 - \psi_2)}
\]

(12)

Figure 1. Logistic and exponential transition functions with varying values of gamma (\( \gamma \)).
taking the L. C. M. of the R. H. S. of (12) and simplifying further gives

$$\tilde{X}_t = -c + \sqrt{\frac{2c + 2c^2}{\gamma^2}} \frac{X_{t-1} - \gamma X_{t-2}}{\gamma(1-\gamma)}$$

(13)

Note that (13) is an estimate of the previous state with the following regularity conditions:

$$C > 0, 0 < \gamma < \infty, \psi_1 > 0, \psi_2 > 0, \psi_1 > \psi_2, 0 < G(X_t) < 1$$ and $$\infty < X_{t-2} < \infty.$$ Recall that $$G(X_t) = G(X_{t-1}; \gamma, c) = 1 - \epsilon^{-\gamma(t)X_{t-1}}$$ as in (7).

One would ask whether (13) inherited the symmetrical feature of the exponential transition function given in (7) or not? To answer this, we need to visualize (13) to see if it is really symmetric; even though it will give two graphs because of the presence of [±] signs, the two graphs are given in Figure 2 below:

Figure 2 clearly show that the symmetrical properties of Eq. (7) were inherited by (13); hence our system of equation: Eqs. (1) and (2) is now nonlinear as well as symmetric. The system is therefore capable of handling/modeling any symmetric nonlinear series such as exchange rate.

Hence, the Kalman filter algorithms of the proposed system (SNSSM) was developed using an established marginal and conditional property of multivariate normal distribution. This procedure was also adopted by Hamilton (1994), Zivot (2006) and Tsay (2010). Thus, the algorithms (also known as Kalman recursions) are given in the predictions and updating equations below:

- Predictions:

$$X_{t+1|t} = \psi X_t + Z$$

$$F_{t+1} = \psi F_t + Z$$

(14)

where $$X_t$$ as given in (13), $$F_t$$ is the error covariance matrix of the previous state and $$Z$$ is as previously defined.

- Updating:

$$X_{t+1|t+1} = X_{t+1|t} + HF_{t+1}Q_{t+1}^{-1}(Y_{t+1} - Y_{t+1|t})$$

$$F_{t+1} = F_{t+1|t} - HF_{t+1|t}Q_{t+1|t}H^T F_{t+1|t}$$

(15)

where $$Y_{t+1|t} = HX_{t+1|t}, Q_{t+1}^{-1} = HF_{t+1|t}H^T + L$$ and $$L$$ is as previously defined.

Note: the initialization process of the above algorithms is discussed in section 3; and that is the overall objective of this paper.

3. Specifying the initial Kalman recursions of the developed SNSSM

Initialization of the Kalman filter usually involves two possible situations: It may be that the system under study is stationary and it may also be that the system under study is nonstationary. So, in Kalman filter theory, if the series under study is stationary (which is very rare in real life), the mean and variance are usually obtained unconditionally. Similarly, if the said system is nonstationary, the mean and variance/covariance are usually obtained conditionally. Details of these two cases were given below:

3.1. Case I: the system is assumed to be stationary

Consider our system of Eqs. (3) and (4); if the system is covariance stationary, then the state vector $$X_t$$ is also covariance stationary. Hence, the mean and variance/covariance will be determined unconditionally using Kronecker products and vec operator. The vec operator is a technique of transforming a matrix to a column vector by stacking the columns of the original matrix (say D) on top of the other sequentially1. For details on this see example Knut et al. (2005), Helmut (2005), Caswell and Daalen (2016), Casarin et al. (2017), David (2008) among others.

Now, the initialization is started by considering the state Eq. (4) and assuming that the initial value of the state vector $$X_t$$ is chosen from normal distribution (note: it is customary at the initial stage to let $$t = 0$$ with mean $$\mu_0 = \bar{X}_0$$ and covariance $$F_0$$; and we can recall that for this process to be stationary, there is need for all the eigenvalues of $$\psi$$ to lie within the unit circle.

Now, we can write the unconditional mean of (4) as

$$E(X_0) = \bar{X}_0 = 0$$

(16)

And the unconditional variance is given as

$$F_0 = E(X_0X_0')$$

(17)

As an illustration, the unconditional variance of $$X_t$$ given in (18) can be calculated from (4). To achieve this; we postmultiplied the state equation by its transpose, taking the expectation and equating all cross products to zero, that is

$$X_{t+1|t} = \psi X_t + Z$$

$$F_{t+1|t} = \psi F_t + Z.$$

(14)

1 If the vec operator is applied to an $$n \times r$$ matrix $$D$$, it will return an $$m \times 1$$ column vector defined as vec$$(D)$$; this is achieved by stacking each column of $$D$$ on top of the next column ordered from left to right. Consider for example if the product $$KLM$$ is defined, then vec$$(KLM) = (M \otimes K) \cdot$$ vec$$(L)$$. Here, vec is refer to a stacking operator of the column and $$\otimes$$ is the Kronecker product.
\[ E(X_i, X_{i+1}) = E((\psi X_i + \Omega_{i1})|\psi X_i + \Omega_{i1})' \]

\[ = E((\psi X_i + \Omega_{i1})|\psi X_i + \Omega_{i1}) \]

Remember that if the state vector \( X_i \) is stationary, then \( E(X_i, X_{i+1}) = E(X_i, X_t) = E_i \); and for \( t = 0 \), Eq. (18) becomes

\[ F_0 = \psi F_0 \psi' + Z \]

Applying the vec operator to (20) (19), we have

\[ \text{vec}(F_0) = ((\psi') \otimes \psi) \cdot \text{vec}(F_0) + \text{vec}(Z) \]

which becomes

\[ \text{vec}(F_0) - (\psi \otimes \psi) \cdot \text{vec}(F_0) = \text{vec}(Z) \]

which is

\[ \text{vec}(F_0) = (I - \psi \otimes \psi) \cdot \text{vec}(Z) \]

where

\[ \text{vec}(Z) \]

is an \( n \times 1 \) column vector, \( I \) is an \( n \times n \) identity matrix, \( (\psi \otimes \psi) \) is an \( n^2 \times n^2 \) matrix and \( \text{vec}(Z) \) is an \( n^2 \times 1 \) column vector.

In the case of two variables (i.e., \( n = 2 \)), the unconditional variance equation for the initial Kalman recursions of the SNSSM representation given in Eq. (20) can be written out in full as

\[ \text{vec}(F_0) = (I - \psi \otimes \psi) \cdot \text{vec}(Z) \]

(21)

Simplifying further, we have

\[ \begin{bmatrix} f_{11}^0 \\ f_{21}^0 \\ f_{22}^0 \end{bmatrix} = \begin{bmatrix} 1 - \varphi_{11}^2 & -\varphi_{11} \varphi_{12} & -\varphi_{11} \varphi_{12} \\ -\varphi_{11} \varphi_{21} & 1 - \varphi_{12}^2 & -\varphi_{12} \varphi_{22} \\ -\varphi_{11} \varphi_{21} & -\varphi_{12} \varphi_{21} & 1 - \varphi_{11}^2 \end{bmatrix} \begin{bmatrix} z_{11} \\ z_{21} \\ z_{22} \end{bmatrix} \]

(22)

where

\[ f_{11}^0 \]

is the unconditional variance of the state \( x_{11}^0 \); \( f_{22}^0 \) is the unconditional variance of the state \( x_{22}^0 \) and \( f_{21}^0 \) is the unconditional covariance between the states \( x_{12}^0 \) and \( x_{22}^0 \). Hence, in order to handle the inversion problem of the \( 4 \times 4 \) matrix associated with Eq. (22), we need to partition the matrix and apply the principle of inverse of a partitioned matrix (matrix inversion lemma) accordingly.

Matrix inversion lemma is a tool that is found very useful in Time Series Analysis, Estimation Theory, Signal Processing and Control Engineering. Applying the lemma, the inverse of any partitioned matrix can be found provided that the elements (submatrices) on the main diagonal of the partitioned matrix are invertible square matrices, and the other elements (submatrices) may or may not be square (Dan, 2006).

Now, from (22) let the \( 4 \times 4 \) matrix be

\[ \kappa = \begin{bmatrix} \kappa_{11} & \kappa_{12} \\ \kappa_{21} & \kappa_{22} \end{bmatrix} \]

(23)

where

\[ \kappa_{11} = \begin{bmatrix} 1 - \varphi_{11}^2 & -\varphi_{11} \varphi_{12} \\ -\varphi_{11} \varphi_{21} & 1 - \varphi_{12}^2 \end{bmatrix}, \quad \kappa_{12} = \begin{bmatrix} -\varphi_{11} \varphi_{12} \\ -\varphi_{12} \varphi_{21} \end{bmatrix}, \quad \kappa_{21} = \begin{bmatrix} -\varphi_{11} \varphi_{21} \\ -\varphi_{12} \varphi_{22} \end{bmatrix}, \quad \kappa_{22} = \begin{bmatrix} 1 - \varphi_{12}^2 & -\varphi_{11} \varphi_{12} \\ -\varphi_{12} \varphi_{22} & 1 - \varphi_{11}^2 \end{bmatrix} \]

Some unique properties of \( \kappa \) are

1. \( (\kappa_{11} - \kappa_{22})^{-1} = \kappa_{11}^{-1} + \kappa_{22}^{-1} \)
2. \( |\kappa| = |\kappa_{11}| |\kappa_{22} - \kappa_{21} \kappa_{12}^{-1} \kappa_{11}| \)
3. \( \kappa^{-1} \) can be written as

\[ \kappa^{-1} = \begin{bmatrix} \kappa_{11} & \kappa_{12} \\ \kappa_{21} & \kappa_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \kappa_{11}^{-1} + \kappa_{22}^{-1} & -\kappa_{22}^{-1} \kappa_{12}^{-1} \kappa_{11}^{-1} \\ -\kappa_{22}^{-1} \kappa_{12}^{-1} & \kappa_{22}^{-1} \end{bmatrix} \]

(24)

assuming that \( \kappa_{11}^{-1} \) exists. And \( \theta = \kappa_{22} - \kappa_{21} \kappa_{12} \) alternatively

\[ \kappa^{-1} = \begin{bmatrix} \kappa_{11} & \kappa_{12} \\ \kappa_{21} & \kappa_{22} \end{bmatrix}^{-1} = \begin{bmatrix} \theta^{-1} & -\theta^{-1} \kappa_{12} \\ -\kappa_{21} \theta^{-1} & \kappa_{22} \theta^{-1} \end{bmatrix} \]

(25)

assuming that \( \kappa_{11}^{-1} \) exists. And \( \theta = \kappa_{11} - \kappa_{12} \kappa_{12}^{-1} \kappa_{11} \)

Substituting either Eq. (24) or (25) in Eq. (22), the unconditional variance equation for the initial Kalman recursions of the SNSSM representation given in Eq. (22) can easily be found. For the details of the above lemma, see (Knut et al., 2005) (Tsai, 2014), (Dan, 2006) and (Knut et al., 2008).

3.2 Case II: the system is assumed to be nonstationary

Recalling our system of equations: (3)(3) and (4)(4); and assuming that the system is nonstationary. Under this situation, the mean and variance/covariance of the initial state \( X_0 \) were usually obtained conditionally. To achieve this, we used a well-known marginal and conditional property of the multivariate normal distribution; [see, Rencher (2002), for more details even though we can achieve same from the prediction and updating equations: (14)(14) and (15)(15) by letting \( t = 1 \), but for better understanding, we choose to follow this path.

From (3)(3) and (4)(4), let \( X_{i-1} \) and \( Y_{i-1} \) denote \( n \times 1 \) and \( m \times 1 \) subvectors respectively whose joint normal distribution is given as

\[ J = \begin{bmatrix} X_{i-1} \\ Y_{i-1} \end{bmatrix} \sim N_2(\mu, \Sigma) \]

(26)

Then the conditional distribution of \( X_{i-1} \) given \( Y_{i-1} \) is multivariate normal with mean \( \mu \) and variance-covariance matrix \( \Sigma \); i.e., \( X_{i-1}|Y_{i-1} \sim N(\mu, \Sigma) \); where

\[ \mu = E(X_{i-1}|Y_{i-1}) = \mu_{i|Y} = \mu + \tau_{YX} \tau_{XV}^{-1}(Y_{i-1} - \mu) \]

(27)

\[ \tau = \text{var}(X_{i-1}|Y_{i-1}) = \tau_{XX} - \tau_{XY} \tau_{VY}^{-1} \tau_{YX} \]

(28)

In the theory of the State-Space model with Gaussianity assumption, the system innovations \( \Omega_{i-1} \) and \( \chi_{i-1} \) are normally distributed and the initial state vector \( X_0 \) is also normally distributed. We can therefore write the state equation at \( t = 0 \) as

\[ X_i = \psi X_0 + \Omega_i \]

(29)

Since \( X_0 \sim N(\mu_0, F_0) \), \( \Omega_i \sim N(0, \Theta) \) also \( X_i \) and \( \Omega_i \) are independent, so

\[ E(X_i) = \mu_{i|Y} = \psi \hat{X}_0 \]

(30)

Similarly,

\[ \tau_{XX} = \text{var}(X_i) = F_{i0} = E[(X_i - E(X_i))(X_i - E(X_i))^\top] \]
\[ F_{1|0} = E\left[ (X_t - \mu) (X_0 - \mu) \right] \]

using the facts from (29) and (30), we have

\[ F_{1|0} = E\left[ (\psi X_0 + \Omega - \psi \tilde{X}_0) (\psi X_0 + \Omega - \psi \tilde{X}_0) \right] \]

that is

\[ F_{1|0} = \psi E\left( X_0 - \tilde{X}_0 \right) \psi' + E(\Omega, \Omega') \]

therefore

\[ \tau_{XX} = F_{1|0} = \psi F_0 \psi' + Z \]

It follows that \( X_t \sim N(\mu_{1|0}, F_{1|0}) \).

---

**Figure 3.** Flow Chart for the Kalman Recursions of Symmetric Nonlinear State-Space Model.
Note that (30)(30) and (31)(31) are the prediction equation and prediction error of the state \( X_t \) at \( t = 0 \) respectively.

And for the measurement equation at \( t = 0 \), we have

\[
Y_1 = HX_1 + V_1
\]  

(32)

Since \( X_t \sim N(\mu_{10}, F_{10}) \), \( V_t \sim N(0, L) \), \( X_t \) and \( V_t \) are independent, so \( Y_1 \) is also distributed normal with mean and variance given below

\[
E(Y_1) = \mu_{10} = H\mu_1;
\]

\[
\text{Var}(Y_1) = HF_{10}H' + L
\]  

(33)

and

\[
\tau_{Y1} = \text{var}(X_1) = E((Y_1 - E(Y_1))(Y_1 - E(Y_1)))
\]

\[
= E((X_1 - \mu_{10})(Y_1 - \mu_{10})')
\]

\[
= E((X_1 - \mu_{10})(HX_1 + V_1 - H\mu_{10}))
\]

\[
= HE[((X_1 - \mu_{10})(X_1 - \mu_{10}')) + E(V_1V_1')]
\]

therefore

\[
\tau_{Y1} = HF_{10}H' + L
\]  

(34)

We can notice that (33)(33) and (34)(34) are the prediction equation and error dispersion for the measurement equation \( Y_t \) at \( t = 0 \) respectively.

Now, the next step is to find the covariance of \( X_t \) and \( Y_t \) to enable us update \( J \) (26) as well as (27)(27) and (28)(28) i.e. the updating equations at \( t = 0 \). Recall that

\[
\text{cov}(X_1, Y_1) = \tau_{XY}
\]

\[
= E((X_1 - E(X_1))(Y_1 - E(Y_1)))
\]

\[
= E((X_1 - \mu_{10})(Y_1 - \mu_{10}'))
\]

\[
= E((X_1 - \mu_{10})(HX_1 + V_1 - H\mu_{10}))
\]

\[
= E((X_1 - \mu_{10})(X_1 - \mu_{10}'))H' + E(V_1V_1')
\]

therefore

\[
\tau_{XY} = F_{10}H
\]  

(35)

Hence, the update of \( J \) at \( t = 0 \) is

\[
J = \begin{pmatrix}
X_1 \\
Y_1
\end{pmatrix} \sim N_{n+m} \begin{pmatrix}
\mu_{10} \\
F_{10}H
\end{pmatrix} \begin{pmatrix}
I & F_{10}H
\end{pmatrix}^{-1} \begin{pmatrix}
H_{10} & H_{10}H' + L
\end{pmatrix}
\]

(36)

And the conditional distribution of \( X_t \) given \( Y_t \) is multivariate normal: \( X_t|Y_t \sim N(\mu_1, F_1) \), where \( \mu_1 \) and \( F_1 \) are the update of (27)(27) and (28)(28) at \( t = 0 \) and are given below

\[
\mu_1 = E(X_t|Y_t) = \mu_{10} + F_{10}H'(H_{10}H' + L)^{-1}(Y_1 - H\mu_{10})
\]

\[
= \mu_{10} + K_1(Y_1 - H\mu_{10})
\]  

(37)

and

\[
F_1 = \text{var}(X_t|Y_t)
\]

\[
= F_{10} - F_{10}H'(H_{10}H' + L)^{-1}H_{10}
\]

\[
= F_{10} - K_1H_{10}
\]

\[
= (I - K_1H)F_{10}
\]  

(38)

Note that (37)(37) and (38)(38) are the updating equations of the Kalman filter at \( t = 0 \). Similarly, the prediction equations: (30)(30) and (31)(31) together with the updating equations: (37)(37) and (38)(38) gives one complete Kalman filter's iteration also known as the Kalman recursion. Repeating the same process at \( t = 1, 2, ..., T \) (where \( T \) is the number of observations) yields the Kalman recursions in Kalman filter literature.

To illustrate how the algorithms provided above will work sequentially, it is necessary to provide a flow chart for the algorithms of the system (SNSSM). The flow chart shows/demonstrate the relationship between case I and II of section 3, and how they (the cases) collectively relate with section 2.2. Similarly, the flow chart provided the criteria that will enable the Kalman recursions (iterations) commence. The flow chart is given in Figure 3.

In summary, it is customary to initializes/starts the filter with some arbitrary values (a prior information) say \( \mu_{00} \) and \( F_{00} \). Use it to predict \( Y_{10} \) and \( Q_{10} \). Whenever the observation \( Y_1 \) becomes available, it will be used in the updating equations and compute \( \mu_{11} \) and \( F_{11} \), which at the same time considered as a prior for the subsequent observation. This process completes one Kalman recursion. It is very important to note that the effect of initial prior \( \mu_{00} \) and \( F_{00} \) is decreasing with the increase of time \( t \) (Yu, 2015), and (Tsay, 2010).

Declarations

**Author contribution statement**

M. Tasi’u: Conceived and designed the experiments; Wrote the paper.

H. G. Dikko: Performed the experiments.

O. I. Shittu: Analyzed and interpreted the data.

I. A. Fulatan: Contributed reagents, materials, analysis tools or data.

**Funding statement**

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

**Competing interest statement**

The authors declare no conflict of interest.

**Additional information**

No additional information is available for this paper.

**References**

Alenka, K., Peter, M., Bernhard, B., Darja, B., 2014. Nonlinear Econometric Models: the Smooth Transition Regression Approach, a Research Supported by CERGE-EI Foundation under Global Development Network (GDN).

Arnold, T., Bertus, M.J., Godfrey, J., 2008. A Simplified Approach to Understanding the Kalman Filter Technique, 8, Finance Faculty Publication, ROBINS School of Business, University of Richmond, pp. 1–23.

Antti, H., 2001. Nonlinear Switching State-Space Models. Master’s Thesis of the Department of Engineering Physics and Mathematics, Helsinki University of Technology.

Anon, 2019. Particle filter. Accessed at https://en.wikipedia.org/wiki/Particle_filter on Monday, July 6th, 2020 by 2:31:23 pm.

Casarin, R., Petinischi, Y., 2017. Matrix Algebra and Vector Spaces for Econometrics. PhD Lecture Note. School of Advanced Studies in Economics, University of Venice, Italy.

Caswell, H., Daalen, S.F., 2016. A note on the vec operation applied to unbalanced block-structured matrices. J. Appl. Math. 2016, 1–3.

Dan, S., 2006. Optimal State Estimation: Kalman, H∞, and Nonlinear Approaches. A John Wiley and Sons, Inc, Publication.

David, A.H., 2008. Matrix Algebra from a Statistician’s Perspective. Springer-verlag Berlin Heidelberg, New York.

Del Moral, P., 1996. Non linear filtering: interacting particle solution. Markov Process.Related Fields 2 (4), 555–580.

Gao, B., Hu, G., Gao, S., Zhong, Y., Gu, C., 2018. Multi-sensor optimal data fusion for INS/ GNSS/INS integration based on unscented Kalman filter. Int. J. Contr. Autom. Syst. 16, 129–146.

Hamilton, J.D., 1994. State-space models. Chapter 50. In: Engle, R.F., McFadden, D.L. (Eds.), Handbook of Econometrics, 4. Elsevier Science B.V, Amsterdam, the Netherlands, pp. 3039–3080.
Helmut, L., 2005. New Introduction to Multiple Time Series Analysis. Springer-verlag Berlin Heidelberg, New York.

Hu, G., Ni, L., Gao, B., Zhu, X., Wang, W., Zhong, Y., 2020. Model predictive based unscented Kalman filter for hypersonic vehicle navigation with INS/GNSS integration. IEEE Access 8, 4814–4823.

Iqbal, M.A., 2016. On Smooth Transition Autoregressive (STAR) Models and Their Applications: an Overview. Library Avenue, New Delhi-110012. PhD (Agricultural Statistics) Thesis Roll No. 9068 I. A. S. R. I.

Knut, S., Arne, S., Peter, B., 2005. Economists’ Mathematical Manual, fourth ed. Springer-verlag Berlin Heidelberg, New York.

Knut, S., Hammond, P., Atle, S., Arne, S., 2008. Further Mathematics for Economic Analysis, second ed. Ashford Colour Press Ltd, Gosport, UK.

Levy, B.C., Nikoukhah, R., 2013. Robust state space filtering under incremental model perturbations subject to a relative entropy tolerance. IEEE Trans. Automat. Contr. 58, 682–695.

Levy, B.C., Zorzi, M., 2016. A contraction analysis of the convergence of risk-sensitive filters. SIAM J. Contr. Optim. 54 (4), 2154–2173.

Raphael, A.Y., 2016. Modified State-Space Model with Nonlinear State Equation. A PhD Thesis. Department of Statistics, University of Ibadan, Nigeria.

Rencher, A.C., 2002. Methods of Multivariate Analysis, second ed. John Wiley and Sons, Inc.

Rhudy, M.B., Rogner, A.S., Keaton, H., 2017. A Kalman filtering tutorial for undergraduate students. Int. J. Comput. Sci. Eng. Syst. 8 (1), 1–18.

Robert, P., Matti, R., Juha, A., Simo, A., 2016. Partitioned update Kalman filter. J. Adv. Informat. Fusion 11 (1), 3–14.

Sascha, M., 2009. Application of State-Space Models in Finance: an Empirical Analysis of the Time-Varying Relationship between Macroeconomics, Fundamentals and Pan-European Industry Portfolios. Universitätsverlag Göttingen.

Teravirala, T., 1994. Specification, estimation, and evaluation of smooth transition autoregressive models. J. Am. Stat. Assoc. 89, 208–218.

Terejanu, G.A., 2019. Unscented Kalman Filter Tutorial. Department of Computer Science and Engineering, University at Buffalo, Buffalo, NY, USA.

Tsai, R.S., 2010. Analysis of Financial Time Series, third ed. John Wiley and Sons, Inc.

Tsai, R.S., 2014. Multivariate Time Series Analysis. John Wiley and Sons Inc, Hoboken, New Jersey, USA.

Usman, M.U., Tuba, S., Guven, S., 2018. Forecasting performance of smooth transition autoregressive (STAR) model on travel and leisure stock index. J. Fin. Data Sci. 4 (2), 90–100.

Wat, E.A., van der Merwe, R., 2019. The Unscented Kalman Filter for Nonlinear Estimation. Sponsored by the NSF under grant: grant IRI-9712346. Department of Electrical and Computer Engineering, Oregon Graduate Institute of Science & Technology, Beaverton, Oregon, USA.

Yaya, O.S., Shittu, O.I., 2016. Symmetric variants of logistic smooth transition autoregressive models: Monte Carlo evidences. J. Mod. Appl. Stat. Methods 15 (1), 711–737.

Yu, J., 2015. Factor models: Kalman filter. Lecture Note in Advanced Economics. School of Economics, Singapore Management University, Singapore.

Zhou, J., 2010. Smooth Transition Autoregressive Models a Study of the Industrial Production Index of Sweden. A Master Thesis of Uppsala Universitet, Sweden.

Zivot, E., 2006. State-Space Models and Kalman Filter. Lecture Note of the Department of Economics, University of Washington, USA.

Zorzi, M., 2017. Robust Kalman filtering under model perturbations. IEEE Trans. Automat. Contr. 62 (6), 2902–2907.

Zorzi, M., Levy, B.C., 2015. On the convergence of a risk sensitive like filter. In: IEEE 54th Annual Conference on Decision and Control (CDC), December 15-18, 2015. Osaka, Japan.