

Kalman Filter of Dynamic Hierarchical Model

Hanadi Saleh A. Al Thibah,

Department of Mathematics, College of Science and Arts, Najran University,
KSA

Sultan Ali M. Salem,

Department of Mathematics, College of Education, Hodeidah University,
Yemen

Doi:10.19044/esj.2020.v16n24p136 [URL:http://dx.doi.org/10.19044/esj.2020.v16n24p136](http://dx.doi.org/10.19044/esj.2020.v16n24p136)

Abstract

In a specific way, the dynamic hierarchical model has been presented alongside with the derivation of the final formula of the Kalman filter. The filtering coefficient used along with the equations necessary for the filtering process has also been determined. Most of the related works were studied which gave rise to the problem statement of filtering problems placed under the case of $s = t$. Most of the basic concepts of the dynamic hierarchical linear model were also displayed based on some previous works. A mathematical formula was also formulated and derived to calculate the dynamic hierarchical Kalman filter model, which results in a repetitive measure to estimate the model parameters. The proposed derived formula reduces the error associated with the model and achieves a successful optimal estimation of the parameters. This proves that the Kalman coefficient is the best filtering for any normal probability distribution and provides the least variance among the estimates. This study also provides an illustrative example of the model with the filtering process concerned. It was further illustrated that the findings can be used in practical applications, which reveals the fields that can be investigated in this area.

Keywords: Deterministic Dynamic Hierarchical Model, Kalman filter gain, Multivariate Gaussian Distribution, Bayesian inference

Introduction

With the applied problems of dynamic linear models, an interesting area of researchers' concerns was noted which makes an inference on the unobserved states θ_t or predicts future observations Y_t , based on a part of the observations. Usually, estimation and forecasting problems are solved by computing the conditional distributions of the interest quantities by considering the available

information. In order for the estimation and forecasting recursions to do this, the state vector θ_s needs to be estimated by computing the conditional densities $p(\theta_s \setminus y_1, y_2, \dots, y_t)$, where the densities $p(y_t \setminus \theta_t)$ and $p(\theta_t \setminus \theta_{t-1})$ have been specified.

On the other hand, computing $p(\theta_s \setminus y_1, y_2, \dots, y_t)$ is totally dependent on the values of s which is often called filtering when $S = t$, state prediction when $S > t$, and smoothing when $S < t$.

However, most of the related works were studied which gave rise to the problem statement of filtering problems placed under the case of $S = t$ (Petris, et al., 2009). Also, most of the basic concepts of the dynamic hierarchical linear model were displayed based on some previous works (Melsa, 1978; Gamerman & Migon, 1993; Migon et al., 2005; Da-Silva et al., 2016; Mohn et al., 2015; Aktekin et al., 2018; Taylor, 2009; Terui & Ban, 2014; Berry & West, 2019; Chen et al., 2019; McAlinn & West, 2019).

In this paper, a mathematical formula will be formulated and derived to calculate the Kalman filter for a dynamic hierarchical model, which will represent a repetitive measure to estimate the model parameters. The proposed derived formula reduces the error associated with the model and achieves a successful optimal estimation of the parameters. This proves that the Kalman coefficient is the best filtering for any normal probability distribution, which provides the least variance among the estimates.

Model Specification

The dynamic hierarchical linear model consists of three equations. The first is the observation equation, the second is the structural equation, and the third is the system equation. It is expressed as follows:

$$\begin{aligned} Y_t &= F_{1t} \theta_{1t} + v_{1t} & , & & v_{1t} &\sim N(0, V_{1t}) \\ \theta_{1t} &= F_{2t} \theta_{2t} + v_{2t} & , & & v_{2t} &\sim N(0, V_{2t}) \\ \theta_{1t} &= G_t \theta_{2,t-1} + w_t & , & & w_t &\sim N(0, W_t) \end{aligned}$$

All the disturbance terms v_{1t}, v_{2t} and w_t are independent with known variance matrices. F_{1t}, F_{2t} are known matrices, G_t is known as matrix, and θ_{1t}, θ_{2t} are unknown vectors (Mohn et al., 2015).

The values $\{F_{1t}, F_{2t}, G_t, V_{1t}, V_{2t}, W_t\}$ indicated may change or may not change over time but with the intention of generalization. In this study, all these values represent constant values at each time t , and the model is, then, referred to as the static-dynamic hierarchical model.

Where,

Y_t is $(n \times 1)$ the observations vector of the model at the time t .

F_{1t} is a known $(n \times n)$ matrix at the time t .

F_{2t} is a known $(n \times r)$ matrix at the time t .

θ_{1t} is $(n \times 1)$ the parameters vector of the model at the time t , it is also called the state vector which is not exactly known.

θ_{2t} is $(r \times 1)$ the structure parameters vector of the model at the time t .

G_t is a known $(r \times r)$ system matrix at the time t .

v_{1t} is $(n \times 1)$ the observation error vector at the time t . It is normally distributed with zero mean and variance matrix V_{1t} is known at time t .

v_{2t} is $(r \times 1)$ the structure error vector at the time t . It is normally distributed with zero mean and variance matrix V_{2t} is known at time t .

w_t is $(n \times 1)$ the system error vector at time t . It is normally distributed with zero mean and variance matrix W_t is known at the time t .

The model is composed of three parts: the observation equation, the structural equations, and the system equation. The observation equation describes the distribution of observations, the structural equations describe the structure of parameter hierarchy, and the system equation describes the form of the evolution of the parameters through time.

Kalman Filter for Dynamic Hierarchical Models

In general, the computations of the conditional densities when solving filtering and forecasting problems may not be easy. The dynamic hierarchical models are of relevant cases where the general recursions simplify considerably. This is due to the fact that the dynamic hierarchical model includes a Gaussian distribution on the random error, and it has nice properties such as closure under marginalization and conditioning distribution. This implies that if the random vector $(y_1, y_2, \dots, y_t)'$ has a Gaussian distribution for any $t \geq 1$, it follows the marginal and conditional distributions of any sub-vector of $(y_1, y_2, \dots, y_t)'$ which are also Gaussian distribution. Since all the relevant distributions are Gaussian, it suffices to compute their means and covariances. The following theorem of Kalman shows the algorithm to find filter densities (Petris et al., 2009; Mohn et al., 2015).

Therefore, through the equations of the model stated above and the statistical description of some components of the model, as well as the information indicated around it, it is possible to find a candidate of Kalman filter by the following theorem. The idea of this theorem was taken from the

research of Gamerman and Migon (1993), but the formulation of this theorem and proof are based on the researcher's method.

Theorem 3.1: The initial information about the parameter (at the time $t = 0$) is :

$$(\theta_{20} \setminus D_0) \sim N(m_{20}, C_{20}) \tag{3.1}$$

Suppose that at a certain time $t - 1$, all the relevant information available up to that time is denoted as D_{t-1} , i.e., D_{t-1} is the set of all observations up to time $t - 1$. In the forecast value of some scalar quantity, it is denoted by Y_t with observed values of this quantity as y_t . It follows that $D_t = \{Y_t, D_{t-1}\}$. This means that $D_t = \{Y_t, D_{t-1}\}$ represents the available information until time t . Thus, the information at a certain time $t - 1$ can be expressed as follows:

- The probability distribution of the parameter $\theta_{2,t-1}$ given D_{t-1} is

$$(\theta_{2,t-1} \setminus D_{t-1}) \sim N(m_{2,t-1}, C_{2,t-1}) \tag{3.2}$$

Thus

- Prior distribution at time t is

For the parameter θ_{2t} is $(\theta_{2t} \setminus D_{t-1}) \sim N(a_{2t}, R_{2t})$ (3.3)

Where

$$a_{2t} = G_t m_{2,t-1} \quad \text{and} \quad R_{2t} = G_t C_{2,t-1} G_t' + W_{tt}$$

For the parameter θ_{1t} is $(\theta_{1t} \setminus D_{t-1}) \sim N(a_{1t}, R_{1t})$ (3.4)

Where

$$a_{1t} = F_{2t} a_{2t} \quad \text{and} \quad R_{1t} = F_{2t} R_{2t} F_{2t}' + V_{2t}$$

- Predictive Distributions (one-step ahead) at time t is

$$(Y_t \setminus D_{t-1}) \sim N(f_t, Q_t) \tag{3.5}$$

Where

$$f_t = F_{1t} a_{1t} \quad \text{and} \quad Q_t = F_{1t} R_{1t} F_{1t}' + V_{1t}$$

Then, the posterior distribution at time t is

1) For the parameter θ_{1t} is $(\theta_{1t} \setminus D_t) \sim N(m_{1t}, C_{1t})$ (3.6)

Where

$$m_{1t} = a_{1t} + R_{1t} F_{1t}' Q_t^{-1} (Y_t - f_t) \tag{3.7}$$

$$C_{1t} = R_{1t} - R_{1t} F_{1t}' Q_t^{-1} F_{1t} R_{1t}' \tag{3.8}$$

Since

$$k_{1t} = R_{1t} F_{1t}' Q_t^{-1} \tag{3.9}$$

In Equation (3.9), k_{1t} is called the Kalman factor (or filter gain) of θ_{1t} and equation (3.7) is called the filtering equation θ_{1t} .

2) For the parameter θ_{2t} is $(\theta_{2t} \setminus D_t) \sim N(m_{2t}, C_{2t})$

Where

$$m_{2t} = a_{2t} + R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} (m_{1t} - a_{1t}) \tag{3.10}$$

$$C_{2t} = R_{2t} - R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} (R_{1t} - C_{1t}) R_{1t}^{-1} F_{1t} F_{2t}' R_{2t}' \tag{3.11}$$

$$\text{Since } k_{2t} = R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} \tag{3.12}$$

In equation (3.12), k_{2t} is called the Kalman factor (or filter gain) of θ_{2t} and equation (3.10) is called the filtering equation θ_{2t} .

Proof: Validity posterior distribution is proven. The posterior distribution of the parameter θ_{1t} at the time t given D_t can be obtained by using Bayes theorem as follows:

$$p(\theta_{1t} \setminus D_t) \propto p(Y_t \setminus \theta_{1t}) p(\theta_{1t} \setminus D_{t-1}) \tag{3.13}$$

Since $(Y_t \setminus \theta_{1t}) \sim N(F_{1t} \theta_{1t}, V_{1t}^{-1})$, then

$$p(Y_t \setminus \theta_{1t}) \propto e^{-\frac{1}{2}(Y_t - F_{1t} \theta_{1t})' V_{1t}^{-1} (Y_t - F_{1t} \theta_{1t})} \tag{3.14}$$

And since $(\theta_{1t} \setminus D_{t-1}) \sim N(a_{1t}, R_{1t}^{-1})$, then

$$p(\theta_{1t} \setminus D_{t-1}) \propto e^{-\frac{1}{2}(\theta_{1t} - a_{1t})' R_{1t}^{-1} (\theta_{1t} - a_{1t})} \tag{3.15}$$

Put equation (3.14) and equation (3.15) in equation (3.13), the result obtained is

$$p(\theta_{1t} \setminus D_t) \propto e^{-\frac{1}{2}[(Y_t - F_{1t} \theta_{1t})' V_{1t}^{-1} (Y_t - F_{1t} \theta_{1t}) + (\theta_{1t} - a_{1t})' R_{1t}^{-1} (\theta_{1t} - a_{1t})]} \tag{3.16}$$

Taking the natural logarithm and multiplying it with -2 for both sides of the equation (3.16) gives:

$$\begin{aligned} & -2Lnp(\theta_{1t} \setminus D_t) \propto (Y_t - F_{1t} \theta_{1t})' V_{1t}^{-1} (Y_t - F_{1t} \theta_{1t}) + (\theta_{1t} - a_{1t})' R_{1t}^{-1} (\theta_{1t} - a_{1t}) \\ & \quad \propto (Y_t' - \theta_{1t}' F_{1t}') V_{1t}^{-1} (Y_t - F_{1t} \theta_{1t}) + (\theta_{1t}' - a_{1t}') R_{1t}^{-1} (\theta_{1t} - a_{1t}) \\ & \alpha Y_t' V_{1t}^{-1} Y_t - Y_t' V_{1t}^{-1} F_{1t} \theta_{1t} - \theta_{1t}' F_{1t}' V_{1t}^{-1} Y_t + \theta_{1t}' F_{1t}' V_{1t}^{-1} F_{1t} \theta_{1t} \\ & \quad + \theta_{1t}' R_{1t}^{-1} \theta_{1t} - \theta_{1t}' R_{1t}^{-1} a_{1t} - a_{1t}' R_{1t}^{-1} \theta_{1t} + a_{1t}' R_{1t}^{-1} a_{1t} \\ & = h + \theta_{1t}' (R_{1t}^{-1} + F_{1t}' V_{1t}^{-1} F_{1t}) \theta_{1t} - 2\theta_{1t}' (R_{1t}^{-1} a_{1t} + F_{1t}' V_{1t}^{-1} Y_t) \end{aligned} \tag{3.17}$$

Where $Y_t V_t^{-1} F_t \theta_t = \theta_t' F_t' V_t^{-1} Y_t$, $\theta_t' R_t^{-1} a_t = a_t' R_t^{-1} \theta_t$, and h is constant, all fixed quantities are included which do not contain θ_t .

$$i.e., \quad h = Y_t' V_t^{-1} Y_t + a_t' R_t^{-1} a_t$$

Now, an equivalent formula can be derived for the first amount in equation (3.17) as follows:

$$\begin{aligned} R_t^{-1} + F_t' V_t^{-1} F_t &= (R_t^{-1} + F_t' V_t^{-1} F_t) (R_t - R_t F_t' Q_t^{-1} F_t R_t') (R_t - R_t F_t' Q_t^{-1} F_t R_t')^{-1} \\ &= (I - F_t' Q_t^{-1} F_t R_t + F_t' V_t^{-1} F_t R_t - F_t' V_t^{-1} F_t R_t F_t' Q_t^{-1} F_t R_t') (R_t - R_t F_t' Q_t^{-1} F_t R_t')^{-1} \\ &= (I + F_t' V_t^{-1} F_t R_t - (V_t + F_t R_t F_t') V_t^{-1} F_t' Q_t^{-1} F_t R_t') (R_t - R_t F_t' Q_t^{-1} F_t R_t')^{-1} \\ &= (I + F_t' V_t^{-1} F_t R_t - Q_t V_t^{-1} F_t' Q_t^{-1} F_t R_t') (R_t - R_t F_t' Q_t^{-1} F_t R_t')^{-1} \\ &= (I + F_t' V_t^{-1} F_t R_t - F_t' V_t^{-1} F_t R_t') (R_t - R_t F_t' Q_t^{-1} F_t R_t')^{-1} \\ &= (R_t - R_t F_t' Q_t^{-1} F_t R_t')^{-1} \\ &= C_t^{-1} \end{aligned} \tag{3.18}$$

Also, an equivalent formula can be derived for the second amount in equation (3.17) as follows:

$$\begin{aligned} R_t^{-1} a_t + F_t' V_t^{-1} Y_t &= R_t^{-1} a_t + F_t' V_t^{-1} Y_t + F_t' V_t^{-1} F_t a_t - F_t' V_t^{-1} F_t a_t \\ &= (R_t^{-1} + F_t' V_t^{-1} F_t) a_t + F_t' V_t^{-1} (Y_t - F_t a_t) \\ &= (R_t^{-1} + F_t' V_t^{-1} F_t) (a_t + (R_t^{-1} + F_t' V_t^{-1} F_t)^{-1} F_t' V_t^{-1} (Y_t - f_t)) \\ &= (R_t^{-1} + F_t' V_t^{-1} F_t) (a_t + (R_t^{-1} + F_t' V_t^{-1} F_t)^{-1} F_t' V_t^{-1} Q_t Q_t^{-1} (Y_t - f_t)) \\ &= (R_t^{-1} + F_t' V_t^{-1} F_t) (a_t + (R_t^{-1} + F_t' V_t^{-1} F_t)^{-1} F_t' V_t^{-1} (V_t + F_t R_t F_t') Q_t^{-1} (Y_t - f_t)) \\ &= (R_t^{-1} + F_t' V_t^{-1} F_t) (a_t + (R_t^{-1} + F_t' V_t^{-1} F_t)^{-1} (R_t^{-1} + F_t' V_t^{-1} F_t) R_t F_t' Q_t^{-1} (Y_t - f_t)) \\ &= (R_t^{-1} + F_t' V_t^{-1} F_t) (a_t + R_t F_t' Q_t^{-1} (Y_t - f_t)) \\ &= C_t^{-1} m_t \end{aligned} \tag{3.19}$$

This is because from equation (3.18),

$$R_t^{-1} + F_t' V_t^{-1} F_t = C_t^{-1}$$

Then substituting equation (3.18) and equation (3.19) in equation (3.17), the result obtained is given as:

$$\begin{aligned} -2Lnp(\theta_t \setminus D_t) &= h + \theta_t' C_t^{-1} \theta_t - 2\theta_t' C_t^{-1} m_t \\ &= h + (\theta_t - m_t)' C_t^{-1} (\theta_t - m_t) \end{aligned}$$

By adding and subtracting the fixed quantity $m_{1t}C_{1t}^{-1}m_{1t}'$, and conducting the analysis as well as the integration of the fixed quantity with the constant h , it gives:

$$p(\theta_{1t} \setminus D_t) \propto e^{-\frac{1}{2}(\theta_{1t} - m_{1t})' C_{1t}^{-1} (\theta_{1t} - m_{1t})}$$

$$i.e., (\theta_{1t} \setminus D_t) \sim N(m_{1t}, C_{1t})$$

This represents the posterior probability distribution of the parameter θ_{1t} given D_t .

Also, from equation (3.19) and equation (3.18), the following is obtained:

$$m_{1t} = a_{1t} + R_{1t} F_{1t}' Q_t^{-1} (Y_t - f_t) \tag{3.20}$$

$$C_{1t} = R_{1t} - R_{1t} F_{1t}' Q_t^{-1} F_{1t} R_{1t}' \tag{3.21}$$

Equation (3.20) represents the Kalman filter which is called the filtering equation and the Kalman factor θ_{1t} is $k_{1t} = R_{1t} F_{1t}' Q_t^{-1}$.

Now, the posterior probability distribution of the parameter θ_{2t} is as follows:

The covariance between θ_{1t} and θ_{2t} given D_{t-1} is

$$\begin{aligned} Cov(\theta_{1t}, \theta_{2t} \setminus D_{t-1}) &= Cov((F_{1t} F_{2t} \theta_{2t} + F_{1t} V_{2t} + V_{1t}), \theta_{2t} \setminus D_{t-1}) \\ &= F_{1t} F_{2t} Cov(\theta_{2t}, \theta_{2t} \setminus D_{t-1}) \\ &= F_{1t} F_{2t} Var(\theta_{2t} \setminus D_{t-1}) \\ &= F_{1t} F_{2t} R_{2t} \\ &= R_{2t}' F_{2t}' F_{1t}' \end{aligned}$$

Therefore, it can be written as:

$$\begin{pmatrix} \theta_{2t} \\ \theta_{1t} \end{pmatrix} \setminus D_{t-1} \sim N \left\{ \begin{pmatrix} a_{2t} \\ a_{1t} \end{pmatrix}, \begin{pmatrix} R_{2t} & R_{2t} F_{2t}' F_{1t}' \\ F_{1t} F_{2t} R_{2t}' & R_{1t} \end{pmatrix} \right\} \tag{Theorem}$$

From this, the following is obtained:

$$\begin{aligned} E(\theta_{2t} \setminus \theta_{1t}, D_{t-1}) &= a_{2t} + R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} (\theta_{1t} - a_{1t}) \\ Var(\theta_{2t} \setminus \theta_{1t}, D_{t-1}) &= R_{2t} - R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} F_{1t} F_{2t} R_{2t}' \end{aligned}$$

Then, the posterior probability distribution of the parameter θ_{2t} given D_t is as follows

$$\begin{aligned} E(\theta_{2t} \setminus D_t) &= E(E(\theta_{2t} \setminus \theta_{1t}, D_{t-1}) \setminus D_t) \\ &= m_{2t} = a_{2t} + R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} (m_{1t} - a_{1t}) = E(a_{2t} + R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} (\theta_{1t} - a_{1t})) \end{aligned} \tag{3.22}$$

$$\begin{aligned}
 \text{Var}(\theta_{2t} \setminus D_t) &= \text{Var}(E(\theta_{2t} \setminus \theta_{1t}, D_{t-1}) \setminus D_t) + E(\text{Var}(\theta_{2t} \setminus \theta_{1t}, D_{t-1}) \setminus D_t) \\
 &= \text{Var}(a_{2t} + R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} (\theta_{1t} - a_{1t})) + E(R_{2t} - R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} F_{1t} F_{2t} R_{2t}') \\
 &= (R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1}) C_{1t} (R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1})' + R_{2t} - R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} F_{1t} F_{2t} R_{2t}' \\
 &= R_{2t} - R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} (F_{1t} F_{2t} R_{2t}' - C_{1t} R_{1t}^{-1} F_{1t} F_{2t} R_{2t}') \\
 &= R_{2t} - R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} (R_{1t} - C_{1t}) R_{1t}^{-1} F_{1t} F_{2t} R_{2t}' \\
 &= C_{2t} \tag{3.23}
 \end{aligned}$$

From equations (3.22) and (3.23), the following can be written as $(\theta_{2t} \setminus D_t) \sim N(m_{2t}, C_{2t})$.

Where

$$\begin{aligned}
 m_{2t} &= a_{2t} + R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} (m_{1t} - a_{1t}) \\
 C_{2t} &= R_{2t} - R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1} (R_{1t} - C_{1t}) R_{1t}^{-1} F_{1t} F_{2t} R_{2t}'
 \end{aligned}$$

Here, $k_{2t} = R_{2t} F_{2t}' F_{1t}' R_{1t}^{-1}$ is called the Kalman factor of θ_{2t} .

Thus, the proof is over.

This is an explanation of the relationship between the sequence stages. Based on this, it is worthy to note that Bayesian inference provides results. Thus, this gives the opportunity to conduct the inference and recognition process of parameter θ_{1t} and parameter θ_{2t} through a probability distribution.

Example: Cross-section of Random Samples of Linear Growing Exchangeable Means.

This example is a generalization of Cross-section of Random Samples with Steady Exchangeable Means with a non-trivial system equation given by

$$\begin{aligned}
 \text{Observation equation} \quad Y_t \setminus \beta_t &\sim N(\beta_t, \sigma^2 I_n), & t = 1, 2, \dots, n \\
 \text{Structure equation} \quad (\beta_t \setminus \mu_t) &\sim N(\mu_t, \tau^2 I_n) \\
 \text{System equation} \quad \begin{aligned} \mu_t &= \mu_{t-1} + \delta_{t-1} + w_{1t} \\ \delta_t &= \delta_{t-1} + w_{2t} \end{aligned} &, \quad \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix} \sim N(0, W_t)
 \end{aligned}$$

This model represents a dynamic hierarchical model, which can be written in the general format as follows:

$$\begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ & & \dots & \\ 0 & 0 & \dots & 1 \end{pmatrix} \begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} + \begin{pmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1n} \end{pmatrix}, \quad \begin{pmatrix} v_{11} \\ v_{12} \\ \vdots \\ v_{1n} \end{pmatrix} \sim N(0, \sigma^2 I_n)$$

$$\begin{pmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_n \end{pmatrix} = \begin{pmatrix} 10 \\ 10 \\ \vdots \\ 10 \end{pmatrix} \begin{pmatrix} \mu_t \\ \delta_t \end{pmatrix} + \begin{pmatrix} v_{21} \\ v_{22} \\ \vdots \\ v_{2n} \end{pmatrix}, \quad \begin{pmatrix} v_{21} \\ v_{22} \\ \vdots \\ v_{2n} \end{pmatrix} \sim N(0, \tau^2 I_n)$$

$$\begin{pmatrix} \mu_t \\ \delta_t \end{pmatrix} = \begin{pmatrix} 11 \\ 01 \end{pmatrix} \begin{pmatrix} \mu_{t-1} \\ \delta_{t-1} \end{pmatrix} + \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix}, \quad \begin{pmatrix} w_{1t} \\ w_{2t} \end{pmatrix} \sim N(0, W_t)$$

In this model, it is noted that

$$r = 2, \quad \theta_{1t} = (\beta_1, \beta_2, \dots, \beta_n)', \quad F_{2t} = F_2 = (1_n, 0_n), \quad \theta_{2t} = (\mu_t, \delta_t)', \quad F_{1t} = F_1 = I_n,$$

$$V_{1t} = \sigma^2 I_n, \quad V_{2t} = \tau^2 I_n, \quad G_t = G = \begin{pmatrix} 11 \\ 01 \end{pmatrix} \text{ and } W_t = W = \begin{pmatrix} \sigma_\mu^2 & 0 \\ 0 & \sigma_\delta^2 \end{pmatrix}$$

From the above, this model is described by the amounts

$$\left\{ I_n, (1_n, 0_n), \begin{pmatrix} 11 \\ 01 \end{pmatrix}, \sigma^2 I_n, \tau^2 I_n, \begin{pmatrix} \sigma_\mu^2 & 0 \\ 0 & \sigma_\delta^2 \end{pmatrix} \right\}, \quad \text{For each time } t.$$

Where 0_p is a p -dimensional vector of 0_s . The parameter δ_t represents the unit growth experienced by the means of the observation levels.

Thus, the filtering process can be done by building a computer program according to the mathematical steps in the above theory.

Conclusion

In this study, a problem statement was examined that stated the filtering problems placed under the case of $S = t$ (Petris et al., 2009; Mohn et al., 2015). Furthermore, most of the basic concepts of the dynamic hierarchical linear model were displayed based on some previous works. A mathematical formula was also formulated and derived to calculate the dynamic hierarchical Kalman filter model. This resulted in a repetitive measure to estimate the model parameters that tried to proffer the solution. The dynamic hierarchical model has been presented alongside with the derivation of the final formula of the Kalman filter. The filtering coefficient used along with the equations necessary for the filtering process was also determined.

The results derived so far were applied when the variance matrices were assumed as known. In fact, it is rare to happen. Therefore, appropriate measures in applications must be taken to allow variance estimation. The proposed study made a simple and effective knowledge assumption of the scalar factor variance and presented the model structure. Here, many applications allowed the same assumption $V_{1t} = \sigma^2 I_n$, where n is the

dimension of Y_t . Also, a conjugate analysis was possible if all variances V_{it} , w_t and C_{k_0} were scaled by σ^2 and the unknown factor with an inverted gamma prior distribution. Here, the result of this theorem remains valid except for the multiplication of all variances σ^2 .

The observed variance is studied due to its estimation, where the proposed method of application could be used for one set of data, *i.e.*, $i = 1$ and could be generalized to more than one set of data. The proposed derived formula reduced the error associated with the model and achieved a successful optimal estimation of the parameters. This proved that Kalman coefficient is the best filtering for any normal probability distribution and provides the least variance among the estimates. Furthermore, this study provided an illustrative example of the model with the filtering process concerned. Finally, it was illustrated that the findings could be used in practical applications, which reveals the fields that can be investigated in this area.

Future work

It is recommended that researchers should carry out studies in three areas: Variance estimation, Non-Linear models, and Non-normal observations.

References:

1. Aktekin, T., Polson, N. G., & Soyer, R. (2018). "Sequential Bayesian analysis of multivariate count data". *Bayesian Analysis*, 13, 385–409.
2. Berry, L.R. & West, M. (2019). "Bayesian forecasting of many count-valued time series". *Journal of Business and Economic Statistics*, doi:10.1080/07350015.2019.1604372.
3. Chen, X., Banks, D., & West, M. (2019). "Bayesian dynamic modeling and monitoring of network flows". *Network Science*, doi: forthcoming.
4. Da-Silva, C.Q. & Migon, H. (2016). "Hierarchical dynamic beta model". *Revstat Statistical Journal* 14.1: 49-73.
5. Gamerman, D. & Migon, H. (1993). "Dynamic Hierarchical Models". *Journal of the Royal Statistical Society. Series B (Methodological)*, vol.55, No.3., pp.629-642.
6. McAlinn, K. & West, M. (2019). "Dynamic Bayesian predictive synthesis in time series forecasting". *Journal of Econometrics* 210,155-169.
7. Melsa, J. L. (1978). "Decision and estimation theory". (No. 04; QA279. 4, M4.).New York-Hill.
8. Migon, H., Gamerman, D., Lopes, H. F., & Ferreira, M. A. (2005). "Dynamic models". *Handbook of statistics*, 25, 553-588.

9. Mohn, M.S., Naik, N., Gemson, R. M. O., & Ananthasayanam, M. R. (2015). "Introduction to the Kalman Filter and Tuning its Statistics for Near Optimal Estimates and Cramer Rao Bound". arXiv preprint arXiv: 1503.04313 .
10. Petris, G., Petrone, S., & Campagnoli, P. (2009). "Dynamic Linear Models With R". Springer, London.
11. Taylor, G. (2009). "A Hierarchical Kalman filter". Taylor fry consulting actuaries. Level 8, 30 Clarence Street. Sydney NSW 2000. Centre for Actuarial Studies, Faculty of Economics and Commerce, University of Melbourne, Parkville VIC 3052. Australia.
12. Terui, N. & Ban M. (2014). "Multivariate time series model with hierarchical structure for over dispersed discrete outcomes". Journal of Forecasting 33,379-390.