The Korean Communications
in Statistics Vol. 4, No. 1, 1997
pp. 193-205

Bayesian Estimation Procedure in Multiprocess Discount Generalized Model[1)]

Joong Kweon Sohn[2)], Sang Gil Kang[3)] and Joo Yong Shim[3)]

Abstract

The multiprocess dynamic model provides a good framework for the modeling and analysis of the time series that contains outliers and is subject to abrupt changes in pattern. In this paper we consider the multiprocess discount generalized model with parameters having a dependent non-linear structure. This model has nice properties such as insensitivity to outliers and quick reaction to abrupt change of pattern in parameters.

1. Introduction

Dynamic systems have been used by communications and control engineers to the state of a system as it evolves through time since the works of Kalman(1960) developed an recursive estimation procedure for the state variables of a linear dynamic system. Ho and Lee(1964) studied the dynamic linear model within Bayesian framework. Duncan and Horn(1972) introduced the Kalman filter by relating the dynamic linear model to random β regression theory using the time varying random parameters as state variables. Harrison and Stevens(1976) summarized the foundations of Bayesian forecasting as the parametric or statespace model, the probabilistic information on model parameters, the sequential model definition which describes the dynamic behavior of model parameters and some uncertainty in choosing the underlying model from a number of discrete alternatives. Masreliez and Martin(1977) developed robust Bayesian estimates for a state space model where either the state noise is Gaussian and the observation noise is heavy-tailed, or vice versa. West(1981) developed an approximation to the sequential updating of the distribution of location parameters of a linear time series model. He examined the behavior of the resulting non-linear filter algorithm. Ameen and Harrison(1985) developed normal discount Bayesian models in order to overcome some practical disadvantages of dynamic linear models. West, Harrison, and

---

1) The Present Studies were Supported by the Basic Science Research Institute Program, Ministry of Education, 1994, Project No. BSRI-94-1403.
2) Professor, Department of Statistics, Kyungpook National University, Daegu 702-701, Korea.
3) Department of Statistics, Kyungpook National University, Daegu 702-701, Korea.

Migon(1985) developed the dynamic generalized linear model for applications in non-linear, non-normal time series and regression problems. Kitagawa(1987) developed a non-normal state space model for non-stationary time series, where the observation and system noise distributions are non-normal.

The multiprocess dynamic linear model was developed by Harrison and Stevens(1971,1976) for the time series that contain outliers and are subject to abrupt changes in pattern. Smith and West(1983) and Smith, Gordon, Knapp and Trimble(1983) described a related monitoring procedure for detecting various forms of kidney failure in renal transplant patients. West(1986) introduced a method of monitoring the predictive performance of a class of Bayesian models. West and Harrison(1986) studied the method of model monitoring and adapting to structural changes in the time series. Bolstad(1986) presented Harrison-Stevens forecasting algorithm and the multiprocess dynamic linear model. Bolstad(1988) developed the multiprocess dynamic generalized linear model. Whittaker and Fruhwirth-Schnatter(1994) used to a triangular multiprocess Kalman filter for detecting bacteriological growth in routine monitoring of feedstuff. Bolstad(1995) developed the multiprocess dynamic poisson model for estimating and forecasting a poisson random variable with a time-varying parameter. Sohn and Kang(1996) developed the multiprocess non-linear dynamic normal model.

In this paper, we develop multiprocess discount generalized models with non-linear structure. In Section 2, we develop the recursive estimation for the multiprocess discount generalized model with parameter non-linearities. Here the model is only partially specified in terms of their first and second moments. In Section 3, we study the proposed recursive estimations for the generalized exponential growth model by using Monte Carlo simulation study.

2. Recursive Estimation of Multiprocess Discount Generalized Model

2.1 Introduction

The multiprocess dynamic linear model by Harrison and Stevens(1971, 1976) is described as follows. The observation is normally distributed and the mean of the observation distribution is a linear function of parameters. The parameter vector is updated by a linear transform plus a perturbation which has an expected value equal to a zero vector. Perturbations at different times are independent of each other. However the perturbation distribution depends on the perturbation index variable at each time. The perturbation index variables at different times are independent and are outcomes of independent multinomial trials with known prior probabilities.

Ameen and Harrison(1985) introduced the normal discount Bayesian model to overcome some practical disadvantages associated with the dynamic linear models. The normal discount Bayesian model updates the variance-covariance matrix of the parameter vector by pre- and post-multiplication by a discount matrix, instead of updating by adding the perturbation variance matrix in the dynamic generalized linear model. This also has the same effect of increasing the variance and in many cases modelers and forecasters have more intuitive feel for the appropriate discount matrix than for a perturbation variance matrix.

Bolstad(1988) proposed the multiprocess dynamic generalized linear model by applying the multiprocess idea on the dynamic generalized linear model.

The multiprocess dynamic generalized linear model can be extended to the multiprocess non-linear dynamic generalized model by introducing non-linear parameter evolution and predictor functions.

In this section, we introduce the multiprocess discount generalized model and suggest the recursive estimation procedure for it. Assumptions of the dynamic discount Bayesian model are same as those of the dynamic generalized linear model except that the prior distribution of β_t given Y_{t-1} will be replaced by a distribution with mean vector $G_t \hat{\beta}_{t-1}$ and variance-covariance matrix $B_t G_t V_{t-1} G_t' B_t$, that is, $(\beta_t \mid Y_{t-1}) \sim (G_t \hat{\beta}_{t-1}, B_t G_t' V_{t-1} G_t' B_t)$, when the posterior distribution at time $t-1$ is a distribution with mean vector $\hat{\beta}_{t-1}$ and variance-covariance matrix V_{t-1}, that is, $(\beta_{t-1} \mid Y_{t-1}) \sim (\hat{\beta}_{t-1}, V_{t-1})$. Here B_t is the discount matrix, a diagonal matrix of discount factors. The effect is similar to that of adding a perturbation in the dynamic linear model. One can see that the mean of the subsequent prior distribution is unchanged and the variance matrix has been inflated to allow for an increased uncertainty. However, the variance matrix inflation is multiplicative instead of additive and this produces some slight differences.

The multiprocess extension of this model allows the discount matrix to have one of k possible values $B_t^{(1)}, \cdots, B_t^{(k)}$, depending on the value of the discount index variable I_t. The discount index variables are an independent outcomes of multinomial random trials with known prior probabilities $P(I_t = j) = \pi_t^{(j)}$ which may change over time. Thus this model is formulated as follows.

Let I_t be perturbation index variable at time t.

$$P(I_t = j) = \pi_t^{(j)} \quad \text{for } j = 1, \cdots, k.$$

The observation model is

$$p(y_t \mid \theta_t) = \exp[c(y_t, \phi) + \phi(\theta_t y_t - b(\theta_t))],$$
$$h(\theta_t) = \eta_t = F_t(\beta_t), \tag{2.1}$$

where $F_t(\cdot)$ is a known non-linear regression function. When $I_t = j$, the evolution equation is

$$\beta_t = g_t(\beta_{t-1}) + r_t, \tag{2.2}$$

where $g_t(\cdot)$ is a known non-linear evolution function and r_t, the perturbation vector at time t.

2.2 Recursive Estimation

The initial conditions for the estimation at time $t-1$ require that the first and second moments for each of k posterior distributions are known. Each distribution is conditional on the perturbation index variable at time $t-1$ having been one of the k possible values. Thus the posterior distribution of β_t given $I_{t-1}=i$ and Y_{t-1} is known as a distribution with mean vector $\widehat{\beta_{t-1}^{(i)}}$ and variance-covariance matrix, $V_{t-1}^{(i)}$, i.e., $(\beta_{t-1} \mid I_{t-1}=i, Y_{t-1}) \sim (\widehat{\beta_{t-1}^{(i)}}, V_{t-1}^{(i)})$. The notation $Y_{t-1}=y_{t-1}, y_{t-2}, \cdots, y_1$ denotes all the observation up to and including y_{t-1}. Also required is that the posterior probabilities of perturbation index variable at time $t-1$, $q_{t-1}^{(i)}=P(I_{t-1}=i \mid Y_{t-1})$, is known.

For the structure of parameter non-linearities, we suggest the linearization technique. Various linearization techniques have been developed for dynamic non-linear models. The most straightforward, and easily interpreted approach is the one that is based on the use of first order Taylor series approximations to non-linear regression and evolution functions. This requires the assumption that both $F_t(\cdot)$ and $g_t(\cdot)$ be differentiable functions of their vector arguments. A Taylor series expansion of the evolution function and observation equation are given as follows.

A Taylor series expansion of the evolution functions about the estimate $\widehat{\beta_{t-1}^{(i)}}$ of β_{t-1}, gives

$$g(\beta_{t-1}) = g_t(\widehat{\beta_{t-1}^{(i)}}) + G_t(\beta_{t-1} - \widehat{\beta_{t-1}^{(i)}}) + R_1(\beta_{t-1} - \widehat{\beta_{t-1}^{(i)}}),$$

where $R_1(\beta_{t-1} - \widehat{\beta_{t-1}^{(i)}})$ is a remainder term which is a function of quadratic and higher order terms of $(\beta_{t-1} - \widehat{\beta_{t-1}^{(i)}})$ and G_t is the known $n \times n$ matrix derivative of the evolution matrix evaluated at the estimate $\widehat{\beta_{t-1}^{(i)}}$,

$$G_t = \left[\frac{\partial g_t(\beta_{t-1})}{\partial \beta_{t-1}'} \right]_{\beta_{t-1} = \widehat{\beta_{t-1}^{(i)}}}.$$

Assuming that terms other than the linear term are negligible, the linearized expression of the evolution equation becomes

$$\begin{aligned} \beta_t &\approx g_t(\widehat{\beta_{t-1}^{(i)}}) + G_t(\beta_{t-1} - \widehat{\beta_{t-1}^{(i)}}) + r_t \\ &= h_t + G_t\beta_{t-1} + r_t, \end{aligned} \tag{2.3}$$

where $h_t = g_t(\widehat{\beta_{t-1}^{(i)}}) - G_t \widehat{\beta_{t-1}^{(i)}}$ is also known. Similarly the non-linear regression function is also linearized about the expected value $a_t = h_t + G_t \widehat{\beta_{t-1}^{(i)}}$ for β_t,

$$F_t(\beta_t) = F_t(a_t) + H_t'(\beta_t - a_t) + R_2(\beta_t - a_t),$$

where $R_2(\beta_t - a_t)$ is a remainder term which is a function of quadratic and higher order terms of $(\beta_t - a_t)$ and H_t is the known n-vector derivative of F_t evaluated at the prior mean a_t,

$$H_t = \left[\frac{\partial F_t(\beta_t)}{\partial \beta_t} \right]_{\beta_t = a_t}.$$

Assuming the linear term dominates the expansion, the non-linear regression function is linearized as

$$\begin{aligned} \eta_t &= F_t(\beta_t) \\ &\approx f_t + H_t'(\beta_t - a_t), \end{aligned} \tag{2.4}$$

where $f_t = F_t(a_t)$.

(1) Evolution Step

In this step, evolving to time t, we find the prior distributions about β_t and η_t depend on the combination of possible models applying at both $t-1$ and t. Each of these k distributions is calculated to time t conditional on $I_t = j$. The joint prior distribution of β_t and η_t and conditional distribution of β_t given η_t are given as follows.

By using the discount matrix and evolution equation (2.3), the mean vector and variance-covariance matrix of the prior distribution of β_t given $I_{t-1} = i, I_t = j$ and Y_t are $\widehat{\beta_t^{(i,j)}} = h_t + G_t \widehat{\beta_{t-1}^{(i)}}$ and $C_t^{(i,j)} = B_t^{(j)} G_t V_{t-1}^{(i)} G_t' B_t^{(j)}$, respectively. That is,

$$(\beta_t \mid I_{t-1} = i, I_t = j, Y_{t-1}) \sim (\widehat{\beta_t^{(i,j)}}, C_t^{(i,j)}). \tag{2.5}$$

By using the prior distribution (2.5) and the observation equation (2.4), we obtain the joint distribution of β_t and η_t.

$$E[\eta_t \mid I_{t-1} = i, I_t = j, Y_{t-1}] = f_t + H_t'(\widehat{\beta_t^{(i,j)}} - a_t)$$

and

$$Var[\eta_t \mid I_{t-1} = i, I_t = j, Y_{t-1}] = H_t' C_t^{(i,j)} H_t,$$

respectively. The covariance of β_t and η_t given $I_{t-1} = i, I_t = j$ and Y_{t-1} is

$$Cov[\beta_t, \eta_t \mid I_{t-1}=i, I_t=j, Y_{t-1}] = C_t^{(i,j)} H_t.$$

Therefore the mean vector and variance-covariance matrix of the joint distribution of β_t and η_t given $I_{t-1}=i, I_t=j$ and Y_{t-1} are mean vector $(\widehat{\beta_t^{(i,j)}}, f_t + H_t'(\widehat{\beta_t^{(i,j)}} - a_t))'$ and variance-covariance matrix

$$\begin{pmatrix} C_t^{(i,j)} & C_t^{(i,j)} H_t \\ H'_t C_t^{(i,j)} & H'_t C_t^{(i,j)} \end{pmatrix},$$

respectively. That is,

$$\begin{bmatrix} \beta_t \\ \eta_t \end{bmatrix} \mid I_{t-1}=i, I_t=j, Y_{t-1} \sim \left(\begin{bmatrix} \widehat{\beta_t^{(i,j)}} \\ f_t + H_t'(\widehat{\beta_t^{(i,j)}} - a_t) \end{bmatrix}, \begin{bmatrix} C_t^{(i,j)} & C_t^{(i,j)} H_t \\ H_t' C_t^{(i,j)} & H_t' C_t^{(i,j)} H_t \end{bmatrix} \right). \quad (2.6)$$

By using the method of linear Bayes estimation, the moments of the conditional distribution of β_t given η_t are directly obtained. Therefore the mean vector and variance-covariance matrix the conditional distribution of the parameter vector given η_t are

$$(\widehat{\beta_t^{(i,j)}} + C_t^{(i,j)} H_t (H_t' C_t^{(i,j)} H_t)^{-1} (\eta_t - (f_t + H_t'(\widehat{\beta_t^{(i,j)}} - a_t)))$$

and

$$(C_t^{(i,j)} - C_t^{(i,j)} H_t (H_t' C_t^{(i,j)} H_t)^{-1} H_t' C_t^{(i,j)}),$$

respectively. That is,

$$(\beta_t \mid I_{t-1}=i, I_t=j, \eta_t, Y_{t-1}) \sim [\ (\widehat{\beta_t^{(i,j)}} + C_t^{(i,j)} H_t (H_t' C_t^{(i,j)} H_t)^{-1} (\eta_t - (f_t + H_t'(\widehat{\beta_t^{(i,j)}} - a_t))), \\ (C_t^{(i,j)} - C_t^{(i,j)} H_t (H_t' C_t^{(i,j)} H_t)^{-1} H_t' C_t^{(i,j)})\ ].$$

(2) Updating Step

In this step we update the prior distribution of the parameter given the observation y_t. Assume that the prior distribution $(\theta_t \mid I_{t-1}=i, I_t=j, Y_{t-1})$ has the conjugate prior distribution $CP(\gamma_t^{(i,j)}, \delta_t^{(i,j)})$. The parameters $\gamma_t^{(i,j)}$ and $\delta_t^{(i,j)}$ are chosen to be consistent with the moments for η_t in joint distribution(2.6). The relationship between the moments of the sampling parameter prior distribution and the moments of η_t is called the guide relationship by West, Harrison, and Migon(1985).

Clearly the joint distribution of y_t and θ_t is

$$f(y_t, \theta_t \mid I_{t-1}=i, I_t=j, Y_{t-1})$$
$$=f(y_t \mid \theta_t) f(\theta_t \mid I_{t-1}=i, I_t=j, Y_{t-1})$$
$$=\exp[c(y_t, \phi)+c(\gamma_t^{(i,j)}, \delta_t^{(i,j)})+\theta_t(\gamma_t^{(i,j)}+y_t\phi)-(\delta_t^{(i,j)}+\phi)b(\theta_t)]$$

and the marginal distribution of y_t is

$$f(y_t \mid I_{t-1}=i, I_t=j, Y_{t-1})$$
$$=\int f(y_t, \theta_t \mid I_{t-1}=i, I_t=j, Y_{t-1})d\theta_t$$
$$=\exp[c(y_t, \phi)+c(\gamma_t^{(i,j)}, \delta_t^{(i,j)})-c(\gamma_t^{(i,j)}+\phi y_t, \delta_t^{(i,j)}+\phi)]$$

Thus posterior distribution of θ_t given $I_{t-1}=i, I_t=j$ and Y_t is

$$f(\theta_t \mid I_{t-1}=i, I_t=j, Y_t)$$
$$=f(y_t, \theta_t \mid I_{t-1}=i, I_t=j, Y_{t-1})/f(y_t \mid I_{t-1}=i, I_t=j, Y_{t-1})$$
$$=\exp[c(\gamma_t^{(i,j)}+\phi y_t, \delta_t^{(i,j)}+\phi)+\theta_t(\gamma_t^{(i,j)}+\phi y_t)-(\delta_t^{(i,j)}+\phi)b(\theta_t)].$$

Therefore the posterior distribution of θ_t given $I_{t-1}=i, I_t=j$ and Y_t is the conjugate posterior $CP(\gamma_t^{(i,j)}+\phi y_t, \delta_t^{(i,j)}+\phi)$ and the probability density function

$$f(\theta_t \mid I_{t-1}=i, I_t=j, Y_t)$$
$$=\exp[c(\gamma_t^{(i,j)}+\phi y_t, \delta_t^{(i,j)}+\phi)+\theta_t(\gamma_t^{(i,j)}+\phi y_t)-(\delta_t^{(i,j)}+\phi)b(\theta_t)]. \tag{2.7}$$

The guide relationship is used to relate the posterior distribution of the sampling parameter back to the distribution for η_t. From posterior distribution(2.7), the posterior mean and variance of $\eta_t=h(\theta_t)$ are calculated and denoted by $\widehat{\eta_t^{(i,j)}}=E[h(\theta_t) \mid I_{t-1}=i, I_t=j, Y_t]$ and $U_t^{(i,j)}=Var[h(\theta_t) \mid I_{t-1}=i, I_t=j, Y_t]$, respectively. This completes the determination of the first two moments of the conditional distributions of η_t at time t posterior to the observations Y_t.

Now we need to relate this back to find the moments of the posterior distribution of the parameter.

By using the Bayes theorem, one can get

$$f(\beta_t, \eta_t \mid I_{t-1}=i, I_t=j, Y_t)$$
$$=f(\beta_t \mid I_{t-1}=i, I_t=j, \eta_t, Y_{t-1}) \cdot f(\eta_t \mid I_{t-1}=i, I_t=j, Y_t).$$

and hence

$$f(\beta_t \mid I_{t-1}=i, I_t=j, Y_t) = \int f(\beta_t, \eta_t \mid I_{t-1}=i, I_t=j, Y_t) d\eta_t.$$

By taking the expectation and variance for this distribution, we obtain the expressions as follows.

$$E[\beta_t \mid I_{t-1}=i, I_t=j, Y_t] = E[E(\beta_t \mid I_{t-1}=i, I_t=j, \eta_t, Y_{t-1}) \mid y_t]$$

and

$$\begin{aligned} Var[\beta_t \mid I_{t-1}=i, I_t=j, Y_t] &= E[Var(\beta_t \mid I_{t-1}=i, I_t=j, \eta_t, Y_{t-1}) \mid y_t] \\ &\quad + Var[E(\beta_t \mid I_{t-1}=i, I_t=j, \eta_t, Y_{t-1}) \mid y_t]. \end{aligned}$$

Thus we obtain the mean vector and variance-covariance matrix as

$$\begin{aligned} \widehat{\widehat{\beta_t^{(i,j)}}} &= E[\beta_t \mid I_{t-1}=i, I_t=j, Y_t] \\ &= \widehat{\beta_t^{(i,j)}} + C_t^{(i,j)} H_t (H_t' C_t^{(i,j)} H_t)^{-1} (\widehat{\eta_t^{(i,j)}} - (f_t + H_t' (\widehat{\beta_t^{(i,j)}} - a_t))) \end{aligned}$$

and

$$\begin{aligned} V_t^{(i,j)} &= Var[\beta_t \mid I_{t-1}=i, I_t=j, Y_t] \\ &= C_t^{(i,j)} - C_t^{(i,j)} H_t (H_t' C_t^{(i,j)} H_t)^{-1} H_t' C_t^{(i,j)} \\ &\quad + C_t^{(i,j)} H_t (H_t' C_t^{(i,j)} H_t)^{-1} U_t^{(i,j)} (H_t' C_t^{(i,j)} H_t)^{-1} H_t' C_t^{(i,j)}, \end{aligned}$$

respectively. Therefore

$$(\beta_t \mid I_{t-1}=i, I_t=j, Y_t) \sim (\widehat{\widehat{\beta_t^{(i,j)}}}, V_t^{(i,j)}). \tag{2.8}$$

This completes the determination of the posterior distributions of the parameters. To complete the development of the recursive estimation, we need to determine the posterior probabilities of the perturbation indices given the present observation. These can be used to detect a change for situation in which the change of pattern is of more interest than the forecast itself. Using the Bayes theorem, we have

$$\begin{aligned} P_t^{(i,j)} &= P(I_{t-1}=i, I_t=j \mid Y_t) \\ &= \frac{q_{t-1}^{(i)} \pi_t^{(j)} \exp[c(y_t, \phi) + c(\gamma_t^{(i,j)}, \delta_t^{(i,j)}) - c(\gamma_t^{(i,j)} + \phi y_t, \delta_t^{(i,j)} + \phi)]}{P(y_t \mid Y_{t-1})}, \end{aligned}$$

for $i=1,\cdots,k$ and $j=1,\cdots,k$. The quantity $P(y_t \mid Y_{t-1})$ is a normalizing constant. Hence the $P_t^{(i,j)}$ are all completely determined.

(3) Collapsing Step

To proceed to time $t+1$, we need to remove the dependence of the joint posterior $P(\beta_t \mid Y_t)$ on $k \times k$ possible combination of $I_{t-1}=i$ and $I_t=j$ for $i=1,\cdots,k$ and $j=1,\cdots,k$. The principle that the effect of different models at time $t-1$ are negligible for time $t+1$ is applied for approximating such mixtures. After collapsing the posterior distribution, mean vector and variance-covariance matrix of β_t are obtained as follows.

By using the posterior index probabilities at time t, the posterior distribution of β_t given $I_t=j$ and Y_t is represented as a k component mixtures of β_t given $I_{t-1}=i, I_t=j$ and Y_t. Thus the posterior distribution is

$$\begin{aligned} f(\beta_t \mid I_t=j, Y_t) &= \sum_{i=1}^{k} f(\beta_t \mid I_{t-1}=i, I_t=j, Y_t) \cdot \frac{P(I_{t-1}=i, I_t=j \mid Y_t)}{P(I_t=j \mid Y_t)} \\ &= \sum_{i=1}^{k} (q_t^{(j)})^{-1} P_t^{(i,j)} f(\beta_t \mid I_{t-1}=i, I_t=j, Y_t), \end{aligned}$$

where $q_t^{(j)} = \sum_{i=1}^{k} P_t^{(i,j)}$.

Also by using the technique of approximation of mixture, the mean vector and variance-covariance matrix are

$$\widehat{\beta_t^{(j)}} = E[\beta_t \mid I_t=j, Y_t] = \sum_{i=1}^{k} (q_t^{(j)})^{-1} P_t^{(i,j)} \widehat{\beta_t^{(i,j)}}$$

and

$$\begin{aligned} V_t^{(j)} &= Var[\beta_t \mid I_t=j, Y_t] \\ &= \sum_{i=1}^{k} (q_t^{(j)})^{-1} P_t^{(i,j)} [V_t^{(i,j)} + (\widehat{\beta_t^{(i,j)}} - \widehat{\beta_t^{(j)}})(\widehat{\beta_t^{(i,j)}} - \widehat{\beta_t^{(j)}})'], \end{aligned}$$

respectively. Therefore the posterior distribution of β_t given $I_t=j$ and Y_t is

$$(\beta_t \mid I_t=j, Y_t) \sim (\widehat{\beta_t^{(j)}}, V_t^{(j)}). \tag{2.9}$$

We are now in the same position as when started the recursive estimation procedure, so we are ready to repeat the process when the next observation becomes available.

2.3 Forecast Distributions

At time t, the distributions required by the forecaster are the distributions of $(\beta_{t+1} \mid I_t=i, I_{t+1}=j, Y_t)$, $(\theta_{t+1} \mid I_{t+1}=j, Y_t)$ and $(y_{t+1} \mid I_t=i, I_{t+1}=j, Y_t)$. These forecast distributions are as follows.

The forecast distribution of y_{t+1} given $I_t = i, I_{t+1} = j$ and Y_t is

$$f(y_{t+1} \mid I_t = i, I_{t+1} = j, Y_t)$$
$$= \sum_{i=1}^{k} \sum_{j=1}^{k} \pi_{t+1}^{(j)} q_t^{(i)} \exp[c(y_{t+1}, \phi) + c(\gamma_{t+1}^{(i,j)}, \delta_{t+1}^{(i,j)}) - c(\gamma_{t+1}^{(i,j)} + \phi y_{t+1}, \delta_{t+1}^{(i,j)} + \phi)].$$

The mean vector and variance-covariance matrix of the forecast distribution for β_{t+1} given $I_t = i, I_{t+1} = j$ and Y_t are $\widehat{\beta_{t+1}^{(i,j)}} = h_{t+1} + G_{t+1} \widehat{\beta_t^{(i)}}$ and $C_{t+1}^{(i,j)} = G_{t+1} V_t^{(i)} G'_{t+1} + R_{t+1}^{(j)}$, respectively. Thus

$$(\beta_{t+1} \mid I_t = i, I_{t+1} = j, Y_t) \sim (\widehat{\beta_{t+1}^{(i,j)}}, C_{t+1}^{(i,j)}).$$

The forecast distribution for θ_{t+1} given $I_t = i, I_{t+1} = j$ and Y_t is the conjugate distribution $CP(\gamma_{t+1}^{(i,j)}, \delta_{t+1}^{(i,j)})$, that is,

$$(\theta_{t+1} \mid I_t = i, I_{t+1} = j, Y_t) \sim CP(\gamma_{t+1}^{(i,j)}, \delta_{t+1}^{(i,j)}).$$

3. Monte Carlo Simulation Study

In this section, we study the performances of the Bayesian estimation proposed in Section 2 via Monte Carlo simulation for the multiprocess discount generalized model.

We consider the generalized exponential growth models by Migon and Gamerman(1993). Let y_t, $t = 1, 2, \cdots, n$, be a time series of interest. The model is normally distributed with mean θ_t and variance $V(\theta_t)/\phi_t$, that is, $(y_t \mid \theta_t, \phi_t) \sim N(\theta_t, V(\theta_t)/\phi_t)$ where $V(\theta_t)$ describes a particular known variance law and ϕ_t is a known scale factor and θ_t is the level of the process related the parameter β_t through the non-linear regression $h(\theta_t) = F_t(\beta_t)$. Since the canonical link function in the normal case is the identity, thus to complete the model specification the functions F and g are defined as

$$\eta_t = F_t(\beta_t) = \beta_{1t}$$

and

$$g_t(\beta_{t-1}) = \begin{pmatrix} \beta_{1t-1} + \beta_{2t-1} \\ \beta_{2t-1}\beta_{3t-1} \\ \beta_{3t-1} \end{pmatrix}.$$

β_{1t-1} is the level, β_{2t-1} is the growth in the level and β_{3t-1} is the damping factor for the model at time $t-1$. The non-linearity in the model is due to multiplicative effect of β_3. The simulation study was be carried out with the following example on an artificially generated time series. The time series consists of 80 normally distributed random variables and are the following change pattern. The time series data start with no change with an outlier of at the 12th observation. At the 21st observation, the growth change starts and continues up to the 30th observation. From the 31th observation to the 50th observation there is no change. From the 51th observation, the damping factor change starts and continues up to the 60th observation. At the 61th observation, level change starts and continues up to the 80th observation with an outlier at the 72th observation.

The forecast and the actual observations are shown in Figure 3.1 and the forecast errors are shown in Figure 3.2. From these figures, it can be summarized as follows.

(i) The developed models give good estimates by using past data as well as present data when the time series is in a stable pattern.

(ii) The developed models are not quite sensitive to an outlier.

(iii) The developed models react quickly when a change occurs. But when a change occurs, the forecast error is slightly increasing.

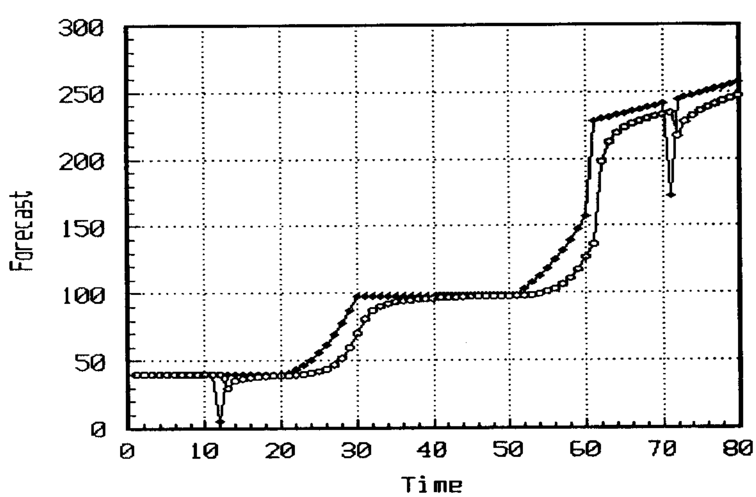

Figure 3.1 Observed(•) and Forecast(◦) Value

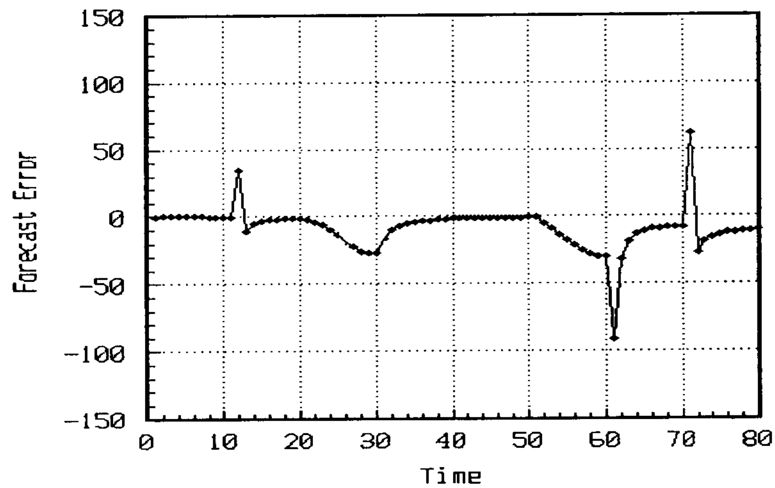

Figure 3.2 Forecast Error

References

[1] Ameen, J. R. M. and Harrison, P. J. (1985), Normal Discount Bayesian Models, *Bayesian Statistics 2*, eds. J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, Amsterdam, North Holland.

[2] Bolstad, W. M. (1986), Harrison-Stevens Forecasting and the Multiprocess Dynamic Model, *The American Statistician*, 40, 129-135.

[3] Bolstad, W. M. (1988), Estimation in the Multiprocess Dynamic Generalized Linear Model, *Communications in Statistics: Theory and Method*, 17, 4179-4204.

[4] Bolstad, W. M. (1995), The Multiprocess Dynamic Poisson Model, *Journal of the American Statistical Association*, 90, 227-232.

[5] Duncan, B. and Horn, S. D. (1972), Linear Dynamic Recursive Estimation from the Viewpoint of Regression Analysis, *Journal of the American Statistical Association*, 67, 815-821.

[6] Harrison, P. J. and Stevens, C. F. (1971), A Bayesian Approach to Short-term Forecasting, *Operations Research Quarterly*, 22, 341-362.

[7] Harrison, P. J. and Stevens, C. F. (1976), Bayesian Forecasting, *Journal of the Royal Statistical Society B*, 38, 205-247.

[8] Ho, Y. C. and Lee, R. C. K. (1964), A Bayesian Approach to Problems in Stochastic Estimation and Control, *IEEE Transactions on Automatic Control*, 33-39.

[9] Kalman, R. E. (1960), A New Approach to Linear Filtering and Prediction Problems, *Journal of Basic Engineering*, 95-108.

[10] Kitagawa, W. (1987), Non-Gaussian State-Space Modeling of Nonstationary Time Series, *Journal of the American Statistical Association*, 82, 1041-1061.

[11] Masreliez, C. J. and Martin, R. D. (1977), Robust Bayesian Estimation for the Linear Model and Robustifying the Kalman Filter, *IEEE Transactions on Automatic Control*, 22, 361-371.

[12] Migon, H. S. and Gamerman, D. (1993), Generalized Exponential Growth Models - A Bayesian Approach, *Journal of Forecasting*, 12, 573-584.

[13] Smith, A. F. M. and West, M. (1983), An Application of Multiprocess Kalman Filter to Renal Transplant Data, *Biometrics*, 39, 867-878.

[14] Smith, A. F. M., West, M., Gordon, K., Knapp, M. and Trimble, I. C.(1983), Monitoring Kidney Transplant Patients, *The Statistician*, 32, 46-54.

[15] Sohn and Kang (1996), Bayesian Estimation Procedure in Multiprocess Non-Linear Dynamic Normal Model, *The Korean Communication in Statistics*, 3, 155-168.

[16] West, M. (1981), Robust Sequential Approximate Bayesian Estimation, *Journal of the Royal Statistical Society B*, 43, 157-166.

[17] West, M. (1986), Bayesian Model Monitoring, *Journal of the Royal Statistical Society B*, 48, 70-78.

[18] West, M. and Harrison, P. J. (1986), Monitoring and Adaptation in Bayesian Forecasting Models, *Journal of the American Statistical Association*, 81, 741-750.

[19] West, M., Harrison, P. J. and Migon, H. S. (1985), Dynamic Generalized Linear Models and Bayesian Forecasting, *Journal of the American Statistical Association*, 80, 73-97.

[20] Whittaker, J. and Fruhwirth-Schnatter, S. (1994), A Dynamic Changepoint Model for Detecting the Onset of Growth in Bacteriological Infections, *Applied Statistics*, 43, 625-640.