자연과학연구 제5호

最小 最大推定量과 베이즈 推定量으로서의 Kalman 필터에 關하여

金 知 坤

상명대학교 자연과학연구소
1998. 12

最小 最大推定量과 베이즈 推定量으로서의 Kalman 필터에 關하여

金 知 坤
祥明大學校 師範大學 數學敎育科

On the Kalman Filter as the Bayes Estimator and the Minimax Estimator

Kim Jee Gon

Department of Mathematical Education Sangmyung University
Seoul 110-743, KOREA

摘要

본 논문은 最小 最大理論과 베이즈 理論을 逐次的으로 使用하여서 最小 最大推定量과 베이즈 推定量으로서의 離散 Kalman 필터를 誘導하고, 推定量의 比較基準인 平均제곱誤差(MSE)에 의하여 이들 두 Kalman 필터의 效率을 比較檢討하였다.

Abstract

In this paper we try to derive the discrete Kalman Filters as the Minimax estimator and the Bayes estimator, by using recursive the Minimax theory and the Bayes theory, and then we compare and examine the efficiencies of these two Kalman Filters by means of the mean square error(MSE) which is the criterion for comparison of estimators.

* 본 연구는 1998학년도 상명대학교 자연과학연구소 연구지원비에 의하여 수행되었음.

1. Introduction

Since the introduction in the mid 1950s ,the filtering techniques developed by Kalman, and by Kalman and Bucy([6],[7]) have been widely known and widely used in all areas of applied sciences. Starting with applications in aerospace engineering, those include quality control,navigation and many others. Indeed, the Kalman Filter is based on a Bayesian estimation technique and many of the resulting methodologies have a Bayesian foundation, thus it is interesting that the theory and methodology of linear dynamic models is not very familiar to statisticians. Recently,however statisticians are begining to study the linear dynamic models and use it with some of the standard statistical problems([2],[8],[11]). The standard approach for estimating the parameters is to use the equations associated with the Kalman Filter. A good introduction and derivation of the formulas is found in Meinhold and Singpurwalla ([9]), and Klugman([8]).

The Kalman Filter that will be considered in this paper,is an inference procedure that consists of a linear model defined at discrete times t=1,2,3,... and a stochastic linear relation between the unknown parameters at time t+1 and at t. Thus we follow the following steps ; at first given initial prior mean and variance using the stochastic linear relation above the linear Bayes estimator(BE) and the Minimax-linear estimator(MILE) are obtained at t=1, at second from the stochastic linear relation above the updated dispersion is obtained, at third the linear BE , the MILE and the updated dispersion in second step serve as a new prior mean and dispersion, at fourth using the new prior mean and dispersion the linear BE and the MILE are obtained again, finally we will repeate these processes.

The aim of this paper is to derive the minimax and the Bayes versions of the Kalman Filter by using recursive the most general parametric of estimators from both the Bayes estimator(BE) and the minimax estimator(MILE), and then compare, and examines the efficiencies of these two Kalman Filters. This paper is divided into three parts. The first part(Section 2) summarizes the general results of the BE and the MILE which have include only materials rather directly relevant to our discussions in the sequel. The second part(Section 3) formulates the Kalman Filters as the BE and the MILE by using recussive the most general parametric form of the BE and the MILE. The third part(Section 4) compares and examines the efficiencies of these two Kalman Filters by means of MSE(the mean square error) which is the criterion for the comparison.

2. The general results of the MILE and the BE

In this section, we will summarize the well known results of the MILE and the BE which have included only materials rather directly relevant to our discussions in the sequel. Minimax estimation in linear models has recently received attention in statistical literature. If one has prior information on the unknown parameter vector β such that β may be assumed to lie in a concentration ellipsoid,the resulting unbiased the MILE has the same form of the BE. A fairly extensive discussion of the problems of minimax-estimation can be found in Rao([10]) and Toutenburg([12]).

Now consider linear model (an observation equation),

$$Y = X\beta + \varepsilon, \tag{2.1}$$

where the matrix X is a known $n \times m$ matrix of rank $s \le m$, the unknown parameter β is an m dimensional vector, and Y and ε are n dimensional random variables. Assume that β is a random variable with known prior mean and known dispersion given by

$$E(\beta) = \theta \quad \text{and} \quad D(\beta) = F_m, \tag{2.2}$$

and assume that

$$E(\varepsilon \mid \beta) = 0 \quad \text{and} \quad D(\varepsilon \mid \beta) = \sigma^2 I . \tag{2.3}$$

where I denotes an appropriate identity matrix. Let the parameter space,

$$\Theta = \{\beta : (\beta - \theta)' G (\beta - \theta) \le 1 \}, \tag{2.4}$$

where G is a positive definite(PD) matrix. Consider a linear estimator of the parametric form

$$P'\hat{\beta} = P'\theta + L'(Y - X\theta), \tag{2.5}$$

where the symbole $(\cdot)'$ means the transposition of $(\cdot)$. The risk or MSE of $P'\hat{\beta}$ is

$$\begin{aligned} MSE(P'\hat{\beta}) &= P'E(\hat{\beta} - \beta)(\hat{\beta} - \beta)'P \\ &= E[P'(\theta - \beta) + L'(Y - X\theta)][(\theta - \beta)'P + (Y' - \theta' X')L] \\ &= (L'X - P')(\beta - \theta)(\beta - \theta)'(X'L - P) + \sigma^2 L'L, \end{aligned} \tag{2.6}$$

$$(\therefore MSE(\beta) = Variance(\beta) + (Bias\ in\ \beta)^2).$$

On the ellipsoid (2.4) or equivalently

$$(\beta - \theta)(\beta - \theta)' \le G^{-1} = F_m, \tag{2.7}$$

the maximum of (2.6) is

$$MSE(P'\hat{\beta}) = (L'X - P')F_m(X'L - P) + \sigma^2 L'L . \tag{2.8}$$

Differentiating with respect to L, let the result be equal to zero,

$$\frac{\partial}{\partial L} MSE(P'\hat{\beta}) = 2XF_mX'L - 2P'F_mX' + 2L\sigma^2 = 0 \tag{2.9}$$

Thus, since XF_mX' is a non-negative definite(NND) matrix, the expression in (2.9) is minimized when

$$L' = P'F_mX'(XF_mX' + \sigma^2 I)^{-1} . \tag{2.10}$$

Thus, substituting (2.10) into (2.5) the MILE is given by

$$P'\tilde{\beta}_m = P'\theta + P'F_mX'(XF_mX' + \sigma^2 I)^{-1}(Y - X\theta) . \tag{2.11}$$

Next the BE under the assumptions of (2.1), (2.3) and $E(\beta) = \theta$, $D(\beta) = F_b$ is given by the lemma 3.1 in J.G.Kim([4]). We have this by the following lemma.

<u>Lemma 2.1</u> The general parametric form of the BE relative to the assumptions in (2.1), (2.3) and $E(\beta) = \theta$, $D(\beta) = F_b$ is given by

$$P'\tilde{\beta}_b = P'\theta + P'F_bX'(XF_bX' + \sigma^2 I)^{-1}(Y - X\theta) . \tag{2.12}$$

We observe that the MILE (2.11) and the BE (2.12) were each obtained by solving a different optimization problem. However the forms of the MILE and the BE are exactly the same if $D(\beta) = F_m = F_b$.

The criterian for comparison of estimators will usually be the mean square error (MSE). With respect to a matrix loss function the MSE is the matrix

$$MSE(P'\hat{\beta}) = P'E[(\hat{\beta} - \beta)(\hat{\beta} - \beta)']P , \tag{2.13}$$

where $\hat{\beta}$ is an estimator of unknown parameter β . It was shown that more precise prior information led to an estimator with smaller average MSE in the theorem 4.2 in J.G.Kim([5]). We take this result as the following theorem with rewritting form.

Theorem 2.2 Suppose that $F_m - F_b$ is non-negative definite. Let $P'\tilde{\beta}_b$ be the BE associated with prior mean θ and dispersion F_b . Let $P'\tilde{\beta}_m$ be the MILE associated with prior mean θ and dispersion F_m . Then

$$MSE(P'\tilde{\beta}_b) \leq MSE(P'\tilde{\beta}_m) . \tag{2.14}$$

Hence we see that if $F_m \leq F_b$, then

$$\begin{aligned} MSE(P'\tilde{\beta}_b) &= P'F_bP - P'F_bX'(XF_bX' + \sigma^2 I)^{-1}XF_b \\ &\leq P'F_mP - P'F_mX'(XF_mX' + \sigma^2 I)^{-1}XF_m = MSE(P'\tilde{\beta}_m) . \end{aligned} \tag{2.15}$$

3. The Kalman Filters as the MILE and the BE

In this section, in the light of the iterative procedure in the previous section 1, we will now derive the mathematical presentation of Kalman Filters as the MILE and the BE by using the results of the section 2. For discrete time points t=0,1,2,3..., consider a linear model(an observation equation) ([1]),

$$Y(t) = X(t)\beta(t) + \varepsilon(t) , \tag{3.1}$$

where $Y(t)$ is an n dimensional vector of observation, $X(t)$ a fixed nonrandom $n \times m$ matrix(or design matrix) of rank $s(t) \leq m$, $\beta(t)$ an m dimensional random vector of unknown parameter(or the state of the system at time t), $\varepsilon(t)$ an n dimensional observation error vector. The error vector $\varepsilon(t)$ satisfies

$$E(\varepsilon(t) \mid \beta(t)) = 0 \quad \text{and} \ , D(\varepsilon(t) \mid \beta(t)) = \sigma_\varepsilon^2(t)I , \tag{3.2}$$

where the symbols E and D denote respectively the mean and the dispersion of $(\cdot)$, and I denotes the appropriate identity matrix. The $\beta(t)$ are random variables and thus the dynamic feature, that is, the stochastic linear relationship between $\beta(t)$ and $\beta(t-1)$ is given by the system equation,

$$\beta(t) = M(t)\beta(t-1) + V(t) \qquad (t = 1, 2, 3, \ldots) , \tag{3.3}$$

where $M(t)$ is an $m \times m$ nonrandom system transition matrix and $V(t)$ is an m dimensional system error vector satisfying ,

$$E(V(t) \mid \beta(t)) = 0 \quad \text{and} \quad D(V(t) \mid \beta(t)) = \sigma_w^2(t)I . \tag{3.4}$$

The vector $V(t)$ and vector $\sigma^2(t)I$ are vector white noise ;

$$E(V(t), V'(\tau)) = \begin{cases} \sigma_w^2(t) & \text{for } t = \tau \\ 0 & otherwise \end{cases} \quad \text{and} \quad E(\varepsilon(t), \varepsilon'(\tau)) = \begin{cases} \sigma_\varepsilon^2(t) & \text{for } t = \tau \\ 0 & otherwise \end{cases} . \tag{3.5}$$

The disturbances $V(t)$ and $\sigma^2(t)I$ are assumed to be uncorrelated at all logs ;

$$E(V(t), \varepsilon'(t)) = 0 \qquad \text{for all } t \text{ and } \tau , \tag{3.6}$$

and also we assume that $\beta(t)$ and $Y(t)$ are uncorrelated with any realization of $V(t)$ or $\sigma^2(t)I$,

$$E(V(t), \beta'(t)) = 0\ , \quad E(\sigma(t)I, \beta'(t)) = 0 \quad \text{and} \quad E(V(t), Y'(t)) = 0\ , \tag{3.7}$$

for $t = 1, 2, 3 \ldots$. For $t = 0$ let the prior assumption be

$$E(\beta(0)) = \mu \quad \text{and} \quad D(\beta(0)) = F_m(0)\ , \tag{3.8}$$

where μ and $F_m(0)$ are the initial values of an m dimensional vector and $m \times m$ PD matrix. From (3.3) notice that

$$E(\beta(1) \mid Y(0)) = M(1)\mu \ \text{and}\ D(\beta(1) \mid Y(0)) = M(1)F_m(0)M'(1) + \sigma_w^2(1)\ , \tag{3.9}$$

where the symbole "'" means the transposition . Let

$$\mu(1) = M(1)\mu \ \text{ and } \ F_m(1) = M(1)F_m(0)M'(1) + \sigma_w^2(1)\ . \tag{3.10}$$

This is the new prior mean and dispersion . Now let the parameter space at time $t = 1$ be

$$\Omega(1) = \{\beta(1):\ (\beta(1) - M(1)\mu)(\beta(1) - M(1)\mu)' \le G^+(1) = F_m(1)\}\ , \tag{3.11}$$

where $G^+(1)$ is PD . Following the method of section 2, a linear estimator of the parameteric form

$$P'\hat{\beta}(1) = P'\mu + L'(1)(Y(1) - X\mu) \tag{3.12}$$

is considered. At the first step ,the maximum MSE of (3.12) on ellipsoid (3.11) is obtained in terms of $L(1)$, and at the second step, the $L(1)$ that minimize the value of the expression obtained in the first step is found, and at final step the optimum estimator is obtained. Now , for $\mu(2) = M(2)\hat{\beta}(1)$ and

$$\begin{aligned}
D(\beta(2) \mid Y(1)) &= E[(\beta(2) - E(\beta(2))][\beta(2) - E(\beta(2))]' \\
&= E[M(2)\beta(1) + V(2) - M(2)\hat{\beta}(1)][M(2)\beta(1) + V(2) - M(2)\hat{\beta}(1)]' \\
&= M(2)E[(\beta(1) - \hat{\beta}(1))(\beta(1) - \hat{\beta}(1))']M'(2) + E(V(2)V'(2)) \\
&= M(2)P(1 \mid 1)M'(2) + \sigma_w^2(2) = F_m(2)\ ,
\end{aligned}$$

where , on $\Omega(1)$

$$\begin{aligned}
P(1 \mid 1) &= \max E_\beta[(\beta(1) - \hat{\beta}(1))(\beta(1) - \hat{\beta}(1))'] \\
&= F_m(1) - F_m(1)X'(1)(X(1)F_m(1)X'(1) + \sigma_\varepsilon^2(1)I)^{-1}X(1)F_m(1)\ ,
\end{aligned}$$

the optimum estimator is obtained again . At the t'th step let

$$\Omega(t) = \{\beta(t):\ (\beta(t) - M(t)\hat{\beta}(t-1))(\beta(t) - M(t)\hat{\beta}(t-1))' \le F_m(t)\}, \tag{3.13}$$

where

$$F(t) = M(t)P(t-1 \mid t-1)M'(t) + \sigma_w^2(t) \tag{3.14}$$

and ,on $\Omega(t-1)$ ([3]),

$$\begin{aligned}
P(t-1 \mid t-1) &= \max\ E_\beta[(\beta(t-1) - \hat{\beta}(t-1))(\beta(t-1) - \hat{\beta}(t-1))'] \\
&= F_m(t-1) - F_m(t-1)X'(t-1)(X(t-1)F_m(t-1)X'(t-1) + \sigma_\varepsilon^2(t-1)I)^{-1}X(t-1)F_m(t-1)\ .
\end{aligned} \tag{3.15}$$

From (2.6) the linear estimator of the parameteric form

$$P'\hat{\beta}(t) = P'M(t)\hat{\beta}(t-1) + L'(t)(Y(t) - X(t)M(t)\hat{\beta}(t-1)) \tag{3.16}$$

has MSE of $P'\hat{\beta}(t)$ with

$$\begin{aligned}
MSE(P'\hat{\beta}(t)) &= P'E_\beta[(\beta(t) - \hat{\beta}(t))(\beta(t) - \hat{\beta}(t))']P \\
&= (L'(t)X(t) - P')(\beta(t) - M(t)\hat{\beta}(t-1))(\beta(t) - M(t)\hat{\beta}(t-1))'(X'(t)L(t) - P) + \sigma^2(t)L'(t)L(t)\ .
\end{aligned} \tag{3.17}$$

Its maximum on $\Omega(t)$ is

$$MSE(P'\hat{\beta}(t)) = (L'(t)X(t) - P')F_m(t)(X'(t)L(t) - P) + \sigma_\varepsilon^2(t)L'(t)L(t)\ . \tag{3.18}$$

Differentiating with respect to $L(t)$, let the result be equal to zero ;

$$\frac{\partial}{\partial L(t)} MSE(P'\hat{\beta}(t)) = 2X(t)F_m(t)X'(t)L(t) - 2P'F_m(t)X'(t) + 2L(t)\sigma_\varepsilon^2(t) = 0 \ . \tag{3.19}$$

Thus , since $X(t)F_m(t)X'(t)$ is NND , the expression in (3.19) is minimized if

$$L'(t) = P'F_m(t)X'(t)(X(t)F_m(t)X'(t) + \sigma_\varepsilon^2(t)I)^{-1} \ . \tag{3.20}$$

Thus , from (3.16) and (3.20) the resulting minimax version of the Kalman Filter is

$$P'\tilde{\beta}_m(t) = P'M(t)\hat{\beta}(t-1) + P'F_m(t)X'(t)(X(t)F_m(t)X'(t) + \sigma_\varepsilon^2(t)I)^{-1}(Y(t) - X(t)M(t)\hat{\beta}(t-1)) \ . \tag{3.21}$$

Similarly , for $t=0$ let the prior assumptions be

$$E(\beta(0)) = \mu \qquad \text{and} \qquad D(\beta(0)) = F_b(0) \tag{3.22}$$

where μ and $F(0)$ be the initial values of $m \times m$ PD matrix . From (3.3) notice that

$$E(\beta(1) \mid Y(0)) = M(1)\mu \quad \text{and} \quad D(\beta(1) \mid Y(0)) = M(1)F_b(0)M'(1) + \sigma_w^2(1) \ . \tag{3.23}$$

Thus , from (2.12) we have

$$P'\hat{\beta}(1) = P'M(1)\mu + P'F_b(1)X'(1)(X(1)F_b(1)X'(1) + \sigma_\varepsilon^2(1)I)^{-1}(Y(1) - X(1)M(1)\mu) \ . \tag{3.24}$$

The $P'\hat{\beta}(2)$ is obtained for the prior assumptions

$$\mu(2) = M(2)\hat{\beta}(1) \quad \text{and} \quad F(2) = M(2)P(1 \mid 1)M'(2) + \sigma_w^2(2) \ , \tag{3.25}$$

where ([3])

$$\begin{aligned} P(1 \mid 1) &= E[(\beta(1) - \hat{\beta}(1))(\beta(1) - \hat{\beta}(1))'] \\ &= F_b(1) - F_b(1)X'(1)(X(1)F_b(1)X'(1) + \sigma_\varepsilon^2(1)I)^{-1}X(1)F_b(1) \ . \end{aligned} \tag{3.26}$$

Once $P'\hat{\beta}(t-1)$ is obtained , $P'\hat{\beta}(t)$ is the BE for the prior assumptions with

$$E(\beta(t) \mid Y(t-1)) = M(t)\hat{\beta}(t-1) \text{ and } D(\beta(t) \mid Y(t-1)) = M(t)P(t-1 \mid t-1)M'(t) + \sigma_w^2(t) \ , \tag{3.27}$$

where

$$\begin{aligned} P(t-1 \mid t-1) = {} & F_b(t-1) - F_b(t-1)X'(t-1)(X(t-1)F_b(t-1)X'(t-1) \\ & + \sigma_\varepsilon^2(t-1)I)^{-1}X(t-1)F_b(t-1) \ , \end{aligned} \tag{3.28}$$

Thus the resulting Bayes version of the Kalman Filter is given by

$$P'\tilde{\beta}_b(t) = P'M(t)\hat{\beta}(t-1) + P'F_b(t)X'(t)(X(t)F_b(t)X'(t) + \sigma_\varepsilon^2(t)I)^{-1}(Y(t) - X(t)M(t)\hat{\beta}(t-1)) \ . \tag{3.29}$$

The above derivations show how the Kalman Filters consist of an iterative MILE or BE where each iteration provides the prior information for the next step . The coefficient matrix in (3.21) or in (3.29) is known as the gain matrix and is denoted respectively ,

$$K_m(t) = F_m(t)X'(t)(X(t)F_m(t)X'(t) + \sigma_\varepsilon^2(t)I)^{-1} \ , \ K_b(t) = F_b(t)X'(t)(X(t)F_b(t)X'(t) + \sigma_\varepsilon^2(t)I)^{-1} . \tag{3.30}$$

Equations (3.21) and (3.29) along with the definitions of $K_m(t)$ and $K_b(t)$ in (3.30) will produce respectively

$$P'\tilde{\beta}_m(t) = P'M(t)\hat{\beta}(t-1) + P'K_m(t)(Y(t) - X(t)M(t)\hat{\beta}(t-1) \tag{3.31}$$

and

$$P'\tilde{\beta}_b(t) = P'M(t)\hat{\beta}(t-1) + P'K_b(t)(Y(t) - X(t)M(t)\hat{\beta}(t-1)) \ . \tag{3.32}$$

4. The Comparison of Efficiencies of two Kalman Filters

In section 2 theorem 2.2 shows that more precise prior information led to an estimator with smaller average MSE . In this section we will now show that for the Kalman Filter more precise initial prior information gives an estimator with a smaller MSE for each value of time t . Thus consider two Kalman Filters where the initial prior information is of the form

$$E(\theta(0)) = \theta \quad \text{and} \quad D(\theta(0)) = F_m(0) \tag{4.1}$$

$$E(\theta(0)) = \theta \quad \text{and} \quad D(\theta(0)) = F_b(0) \ . \tag{4.2}$$

Let $P'\tilde{\theta}_m(t)$ be the MILE associated with the initial prior information (4.1) , derived at that t'th stage. Let $P'\tilde{\theta}_b(t)$ be the BE associated with the initial prior information (4.2) , derived at that t'th stage. Then

$$F_b(0) \leq F_m(0) \ \Rightarrow \ MSE(P'\tilde{\theta}_b(t)) \leq MSE(P'\tilde{\theta}_m(t)) \ . \tag{4.3}$$

In order to prove (4.3) we use mathematical induction and theorem 2.2 . By theorem 2.2 and (2.15) , the result is true for t = 1 ;

$$\begin{aligned} MSE(P'\tilde{\theta}_b(1)) &= P'F_b(1) - P'F_b(1)(X(1)F_b(1)X'(1) + \sigma^2_\varepsilon(1)I)^{-1}X(1)F_b(1) \\ &\leq P'F_m(1) - P'F_m(1)(X(1)F_m(1)X'(1) + \sigma^2_\varepsilon(1)I)^{-1}X(1)F_m(1) \\ &= MSE(P'\tilde{\theta}_m(1)) \ . \end{aligned} \tag{4.4}$$

Assume that it holds true for t = t ;

$$MSE(P'\tilde{\theta}_b(t)) \leq MSE(P'\tilde{\theta}_m(t)) \ . \tag{4.5}$$

Now

$$\begin{aligned} F_b(t+1) &= M(t+1)MSE[P'\tilde{\theta}_b(t)]M'(t+1) + \sigma^2_w(t+1) \\ &\leq M(t+1)MSE[P'\tilde{\theta}_m(t)]M'(t+1) + \sigma^2_w(t+1) = F_m(t+1) \ , \end{aligned} \tag{4.6}$$

and

$$\begin{aligned} F_b(t+2) &= M(t+2)MSE[P'\tilde{\theta}_b(t+1)]M'(t+2) + \sigma^2_w(t+2) \\ &\leq M(t+2)MSE[P'\tilde{\theta}_m(t+1)]M'(t+2) + \sigma^2_w(t+2) = F_m(t+2) \ , \end{aligned} \tag{4.7}$$

at stage $t+1$, from theorem 2.2 ,

$$MSE(P'\tilde{\theta}_b(t+1)) \leq MSE(P'\tilde{\theta}_m(t+1)) \ . \tag{4.8}$$

This completes the proof of (4.3) . From (4.3) we arrive at the following conclusions ;

$$\begin{aligned} F_b(0) < F_m(0) \ &\Rightarrow \ MSE(P'\tilde{\theta}_b(t)) < MSE(P'\tilde{\theta}_m(t)) \\ F_b(0) = F_m(0) \ &\Rightarrow \ MSE(P'\tilde{\theta}_b(t)) = MSE(P'\tilde{\theta}_m(t)) \\ F_b(0) > F_m(0) \ &\Rightarrow \ MSE(P'\tilde{\theta}_b(t)) > MSE(P'\tilde{\theta}_m(t)) \ . \end{aligned} \tag{4.9}$$

REFERENCES

1. Broemeling, L.D.(1985), "Bayesian analysis of linear models", Marcel Dekker, Inc.
2. Diderich, G.T.(1985), "The Kalman Filter from the perspective of Godberger-Theil estimators" ,The American statistician , 39(13), pp. 193-198.
3. Hamilton, J.D.(1994), "Time series analysis", Princeton University Press,Princeton, New Jersey.
4. Jee Gon, K.(1996), "A comparative Study on the estimators of Bayes and the generalized ridge", The Institute of Natural Science Sangmyung Uni. , Vol.3, pp. 159-169.
5. Jee Gon, K.(1995), "A comparative Study on the Bayes and Empirical Bayes inference", The Institute of Basic Science Sangmyung Uni. Vol.8, pp. 87-96.
6. Kalman, R.E.(1960), "A new approach to linear filtering and prediction problems", Journal of Basic Engineering, 82, pp. 34-45.
7. Kalman, R.E. and Bucy, R.S.(1961), "New results in linear filtering and Prediction", Journal of Basic Engineering, 83, pp. 95-108.
8. Klugman, S.A.(1992), "Bayesian statistic in Actuarial science with Emphasis on credibility", Kluwer Academic publishers.
9. Meinhold, R.J. and Singpurwalla, N.D.(1983), "Understanding the Kalman Filter", The American statistician, 37(2), pp. 123-127.
10. Rao, C.R.(1976), "The 1975 wald memorial lectures. Estimation of parameters in a linear model", Ann. Statist. 4, pp 1023-1037.
11. Sorenson, H.W.(1970), "Least squares estimation : from Gauss to Kalman", Kalman Filtering Theory and Application, Edited by H.W. Sorenson. IEEE Press, New york.
12. Toutenburg, H.(1982), "Prior information in linear models", John Wiley and Sons.