A New Martingale Approach to Kalman Filtering

ARUNABHA BAGCHI

Department of Applied Mathematics, Twente University of Technology, Enschede, Postbus 217, The Netherlands

Communicated by Karl-Johan Astrom

ABSTRACT

A new derivation of continuous-time Kalman Filter equations is presented. The underlying idea has been previously used to derive the smoothing equations. A unified approach to filtering and smoothing problems has thus been achieved.

I. INTRODUCTION

Recently many rigorous derivations of continuous-time Kalman Filter equations have been obtained [1, 2, 3]. The most general nonlinear problem has been studied in [3] but its specialization to the linear case obscures the simplicity of the linear problem. An elegant proof in the linear case has been proposed in [1] which exploits a result on the estimation of one martingale from another. As pointed out in [4], the "state martingale" used in [1] does not yield the smoothing equations and a different martingale has, therefore, been proposed. This paper shows that the same martingale can be used to derive the filtering equations also, thus unifying the martingale technique initiated in [1] to derive both the filtering and smoothing equations.

2. PROBLEM FORMULATION

Let us consider the linear stochastic equations (continuous version)

\[ x(t; \omega) = \int_0^t A(\sigma) x(\sigma; \omega) \, d\sigma + \int_0^t B(\sigma) \, dW(\sigma; \omega) \]  
[2.1]

\[ y(t; \omega) = \int_0^t C(\sigma) x(\sigma; \omega) \, d\sigma + \int_0^t D(\sigma) \, dW(\sigma; \omega) \]  
[2.2]

for \( 0 \leq t \leq T \) where \( x(t; \omega) \) and \( y(t; \omega) \) take values in \( n \)- and \( m \)-dimensional Euclidean spaces \( \mathbb{R}^n \) and \( \mathbb{R}^m \), respectively, \( W(t; \omega) \) is a \( p \)-dimensional Wiener process, and \( A(t), B(t), C(t), \) and \( D(t) \) are appropriate dimensional matrix-valued functions. Assume that these coefficient functions are all continuous and \( D(t)D(t)^* > 0 \) on the interval \([0, T]\) of interest, where \( * \) stands for the transpose.

Let $\beta(s)$ be the smallest $\sigma$-algebra generated by the process $y(\sigma; \omega)$, $0 \leq \sigma \leq s$ completed with respect to sets of measure 0 and $\beta(s-)$ the smallest $\sigma$-algebra generated by the process $y(\sigma; \omega)$, $0 \leq \sigma < s$ completed with respect to sets of measure 0. Then since $y(t; \omega)$ is continuous in $t$ with probability one, $\beta(s) = \beta(s-)$. 

Let $\hat{x}(t | s) = E [x(t) | \beta(s) ]$. Then it is well known [5, p. 441] that $\hat{x}(t) = \hat{x}(t | t)$ is the best minimum variance estimate of $x(t)$ based on the observation $y(\sigma; \omega)$, $0 \leq \sigma \leq t$ and is called the filtered estimate of $x(t)$. Since $\hat{x}(t | s)$ is a martingale in $s$ for fixed $t$ we have from [6, p. 121]

$$\hat{x}(t | s) = \int_0^s \gamma_{12}(\tau) dZ_1(\tau; \omega)$$

(2.3)

where $Z_1(t; \omega)$, the so-called innovation process, is defined as

$$Z_1(t; \omega) = y(t; \omega) - \int_0^t C(\sigma) \hat{x}(\sigma; \omega) d\sigma$$

and

$$\gamma_{12}(\tau) = P_{12}(\tau) P_{22}(\tau)^{-1}$$

where

$$P_{12}(\tau) = \lim_{\Delta \to 0} \frac{1}{\Delta} E \left[ (\hat{x}(t | \tau+\Delta) - \hat{x}(t | \tau))(Z_1(t+\Delta) - Z_1(\tau))^* | \beta(\tau) \right] ,$$

$$P_{22}(\tau) = \lim_{\Delta \to 0} \frac{1}{\Delta} E \left[ (\int_\tau^{\tau+\Delta} dZ_1(\sigma; \omega))(\int_\tau^{\tau+\Delta} dZ_1(\sigma; \omega))^* | \beta(\tau) \right] .$$

3. FILTERING EQUATIONS

Let us consider (2.3) for $s < t$. From [6, p. 127], for any $\tau > 0$

$$P_{22}(\tau) = D(\tau)D(\tau)^*$$

while for $\tau < t$

$$P_{12}(\tau) = \lim_{\Delta \to 0} \frac{1}{\Delta} E \left[ (\hat{x}(t | \tau+\Delta) - \hat{x}(t | \tau))(Z_1(t+\Delta) - Z_1(\tau))^* | \beta(\tau) \right]$$

where $\tau$ and $\tau + \Delta$ are both less than $t$. Now

$$x(t) = \Phi(t, \tau + \Delta) x(\tau + \Delta) + \int_\tau^{\tau+\Delta} \Phi(t, \sigma) dW(\sigma)$$

where $\Phi(t, \tau)$ is the fundamental matrix of dimension $n \times n$ satisfying
\[ \frac{d \Phi(t, \tau)}{dt} = A(t) \Phi(t, \tau) \quad \Phi(\tau, \tau) = I. \]

Hence

\[ E[x(t) | \beta(\tau + \Delta)] = \Phi(t, \tau + \Delta) \hat{x}(\tau + \Delta) \]

and

\[ E[x(t) | \beta(\tau)] = \Phi(t, \tau) \hat{x}(\tau). \]

With this, we have

\[ P_{12}(\tau) = \lim_{\Delta \to 0} \frac{1}{\Delta} E\left[ (\Phi(t, \tau + \Delta) \hat{x}(\tau + \Delta) - \Phi(t, \tau) \hat{x}(\tau)(Z_o(\tau + \Delta) - Z_o(\tau))^* | \beta(\tau) \right] \]

\[ = \Phi(t, \tau) \lim_{\Delta \to 0} \frac{1}{\Delta} E\left[ (\Phi(t, \tau + \Delta) \hat{x}(\tau + \Delta) - \hat{x}(\tau)(Z_o(\tau + \Delta) - Z_o(\tau))^* | \beta(\tau) \right] \]

\[ = \Phi(t, \tau) \lim_{\Delta \to 0} \frac{1}{\Delta} E\left\{ (\Phi(t, \tau + \Delta) - I) \hat{x}(\tau + \Delta) + \hat{x}(\tau + \Delta) - \hat{x}(\tau) \right\} \]

\[ \{Z_o(\tau + \Delta) - Z_o(\tau) \}^* | \beta(\tau) \}

\[ = \Phi(t, \tau) \lim_{\Delta \to 0} \frac{1}{\Delta} \Phi(t, \tau + \Delta) - I E\left[ \hat{x}(\tau + \Delta)(Z_o(\tau + \Delta) - Z_o(\tau))^* | \beta(\tau) \right] \]

\[ + \Phi(t, \tau) \lim_{\Delta \to 0} \frac{1}{\Delta} E\left[ (\hat{x}(\tau + \Delta) - \hat{x}(\tau))(Z_o(\tau + \Delta) - Z_o(\tau))^* | \beta(\tau) \right] \]

\[ = \Phi(t, \tau) \lim_{\Delta \to 0} \frac{1}{\Delta} E\left[ \hat{x}(\tau + \Delta)(Z_o(\tau + \Delta) - Z_o(\tau))^* | \beta(\tau) \right] \]

since

\[ \lim_{\Delta \to 0} \frac{\Phi(t, \tau + \Delta) - I}{\Delta} = -A(t) \]

exists and

\[ \lim_{\Delta \to 0} E\left[ \hat{x}(\tau + \Delta)(Z_o(\tau + \Delta) - Z_o(\tau))^* | \beta(\tau) \right] = 0 \]

in $L_1$ sense.
Let us define the error \( e(t) = x(t) - \hat{x}(t) \). Then

\[
\hat{x}(\tau + \Delta) - x(\tau + \Delta) - e(\tau + \Delta) = \int_0^\tau A(\sigma) x(\sigma) \, d\sigma + \int_0^\tau B(\sigma) \, dW(\sigma) - e(\tau + \Delta)
\]

and so

\[
\hat{x}(\tau + \Delta) - \hat{x}(\tau) = e(\tau) - e(\tau + \Delta) + \int_\tau^{\tau + \Delta} A(\sigma) x(\sigma) \, d\sigma + \int_\tau^{\tau + \Delta} B(\sigma) \, dW(\sigma).
\]

Let \( P(t;\tau) = E[e(t;\omega) e(\tau;\omega)^*] \). Then

\[
P_{12}(\tau) = \lim_{\Delta \to 0} \frac{1}{\Delta} E[(e(\tau) - e(\tau + \Delta) + \int_\tau^{\tau + \Delta} A(\sigma) \, d\sigma + \int_\tau^{\tau + \Delta} B(\sigma) \, dW(\sigma)) \times (Z_0(\tau + \Delta) - Z_0(\tau))^* \beta(\tau)] .
\]

Now for any \( \Delta > 0 \), \( e(\tau + \Delta;\omega) \) is uncorrelated with \( y(\sigma;\omega), \sigma \leq \tau + \Delta \) and hence with \( Z_0(\sigma;\omega), \sigma \leq \tau + \Delta \). It is also uncorrelated with (and hence independent of) the random variables generating \( \beta(\tau) \). Hence

\[
E(e((\tau + \Delta);\omega)(\int_\tau^{\tau + \Delta} dZ_0(\sigma;\omega))^* \beta(\tau)) = 0.
\]

Furthermore, we have the following [6, p. 129]:

\[
E(e(\tau)(\int_\tau^{\tau + \Delta} dZ_0(\sigma;\omega))^* \beta(\tau)) = \int_\tau^{\tau + \Delta} P(\tau,\sigma) C(\sigma)^* \, d\sigma,
\]

\[
|E(\int_\tau^{\tau + \Delta} A(\sigma) x(\sigma) \, d\sigma)(\int_\tau^{\tau + \Delta} dZ_0(\sigma;\omega))^* \beta(\tau))| = O(|\Delta|^{3/2}),
\]

\[
E(\int_\tau^{\tau + \Delta} B(\sigma) \, dW(\sigma;\omega))(\int_\tau^{\tau + \Delta} dZ_0(\sigma;\omega))^* \beta(\tau)) = \int_\tau^{\tau + \Delta} B(\sigma) D(\sigma)^* \, d\sigma + O(\Delta^{3/2}).
\]

So

\[
P_{12}(\tau) = P(\tau) C(\tau)^* + B(\tau) D(\tau)^*
\]

where
Thus we finally get, for $s < t$,

$$\hat{x}(t \mid s) = \int_s^t \Phi(t, \tau) \left[ P(\tau) C(\tau)^* + B(\tau) D(\tau)^* \right] (D(\tau) D(\tau)^*)^{-1} dZ_\omega(\tau). \quad (3.1)$$

Now $\hat{x}(t \mid s)$ being a Martingale in $s$ for fixed $t$, we have from Doob [7, Theorem 4.3, p. 355]

$$\lim_{s \to t^-} x(t \mid s) = E \left[ x(t) \mid \beta(t^-) \right] = E \left[ x(t) \mid \beta(t) \right] = \hat{x}(t).$$

Hence taking limit in (3.1) as $s \to t^-$, we get

$$\hat{x}(t) = \int_t^T \Phi(t, \tau) \left[ P(\tau) C(\tau)^* + B(\tau) D(\tau)^* \right] (D(\tau) D(\tau)^*)^{-1} dZ_\omega(\tau),$$

or, writing

$$K(t) = \left[ P(t) C(t)^* + B(t) D(t)^* \right] (D(t) D(t)^*)^{-1},$$

$\hat{x}(t)$ is the solution of the stochastic integral equation

$$\hat{x}(t) = \int_t^T A \hat{x}(s) \, ds + \int_t^T K(s) \left[ dY(s) - C(s) \hat{x}(s) \, ds \right] \quad (3 \cdot 2)$$

and $P(t)$, the error covariance matrix that appears in $K(t)$, satisfies the well-known matrix Riccati equation [6, Corollary 2, p. 137]

$$\frac{d}{dt} P(t) = A(t) P(t) + P(t) A(t)^* + B(t) B(t)^* - \left[ P(t) C(t)^* + B(t) D(t)^* \right] (D(t) D(t)^*)^{-1} \left[ C(t) P(t) + D(t) B(t)^* \right] \quad (3 \cdot 3)$$

with $P(0) = O$.

4. CONCLUSION

A new derivation of linear recursive filtering equations is presented. This, with an earlier paper [4], enables us to give a unified rigorous approach to linear filtering and smoothing problems in continuous-time dynamical systems.
REFERENCES


*Received December, 1974*