Continuous-time Kalman

1 Statement of the problem

Here we study the one-dimensional case of Kalman filtering in continuous time. This is treated in Section 6.5 of the text. The filtering equations are given there in Theorem 6.5.4, which is stated for the multidimensional situation. To avoid messy expressions, the proof is only given in the text for the one-dimensional case. We will do the same here, but our proof will be formulated a little differently.

The statement of the theorem in the one-dimensional case is as follows.

Theorem 1 (One-dimensional Kalman) *Let a, c, h, g be real constants with $g > 0$. Let W and B be independent Wiener processes. Let X_0 be a Gaussian real-valued random variable. Assume that X_0, W, B are independent. Let X and Y be processes satisfying:*

$$X_t = X_0 + \int_0^t aX_s\,ds + \int_0^t c\,dW_s, \tag{1}$$

$$Y_t = \int_0^t h\,dX_s + \int_0^t g\,dB_s. \tag{2}$$

Let

$$\bar{X}_t = E\left(X_t|\sigma\left(Y_{[0,t]}\right)\right), \tag{3}$$

$$\bar{\Sigma}_t = E\left(X_t - \bar{X}_t\right)^2. \tag{4}$$

Then $\bar{X}$ satisfies

$$\bar{X}_t = EX_0 + \int_0^t \left(a - \frac{h^2}{g^2}\bar{\Sigma}_s\right)\bar{X}_s\,ds + \int_0^t \frac{h}{g^2}\bar{\Sigma}_s\,dY_s \tag{5}$$

and $\bar{\Sigma}$ satisfies

$$\bar{\Sigma}_t' = \left(2a - \frac{h^2}{g^2}\bar{\Sigma}_t\right)\bar{\Sigma}_t + c^2. \tag{6}$$

The purpose of these notes is to prove Theorem 1. It is easy to see that we can assume without loss of generality that

$$EX_0 = 0. \tag{7}$$

So from now on we will assume this.

2 Stochastic integrals over an interval

Lemma 1 (Integrating over intervals) *Let $a, b \in [0, \infty)$ with $a \leq b$. Let g be nonrandom and is continuously differentiable. Let V be any random variable.*

$$\int_0^b g(s)\ dW_s - \int_0^a g(s)\ dW_s = \\ g(b)\,(W_b - V) - g(a)\,(W_a - V) - \int_a^b g'(s)\,(W_s - V)\ ds. \quad (8)$$

In particular

$$\int_0^b g(s)\ dW_s - \int_0^a g(s)\ dW_s = g(b)\,(W_b - W_a) - \int_a^b g'(s)\,(W_s - W_a)\ ds. \quad (9)$$

Proof

$$\int_0^b g(s)\ dW_s - \int_0^a g(s)\ dW_s = g(b)W_b - g(a)W_a - \int_a^b g'(s)W_s\, ds.$$

Also

$$0 = V\,(g(b) - g(a)) - \int_a^b g'(s)V\, ds.$$

Subtracting the second equation from the first proves the lemma.

3 A bit of Itô calculus

In the approach in our text the stochastic integral

$$\int_0^t g(s)\ dW_s$$

is only defined rigorously when the integrand g is nonrandom and is continuously differentiable. This is all we need for our course.

As we noted in class, because the stochastic integral is additive with respect to intervals of integration, the definition in the text extends easily to the case in which the integrand g is nonrandom and piecewise continuously differentiable.

Of course this is still rather special. As remarked in class, one can actually define the stochastic integral for very general random integrands g as long as they are *nonanticipating*. To make sure that all the properties we are used will hold, one should also assume that $g(t)$ has finite second moment for each t.

We should be aware that Itô's stochastic integral is defined for general integrands, but for our work we won't use general integrals.

Let W be a Wiener process. We are often interested in a process X which can be written in the following way.

$$X_t = X_0 + \int_0^t f(s)\ ds + \int_0^t g(s)\ dW_s. \tag{10}$$

We often write (10) more concisely as

$$dX_t = f(t)dt + g(t)dW_t, \tag{11}$$

The ordinary integral part,

$$\int_0^t f(s)\ ds,$$

is defined even when f is a rather general random process, of course.

Itô showed that a product rule holds.

Lemma 2 (Itô's product rule for stochastic differentials) *Let (W_t) be a Wiener process and let $dX_t = f(t)\,dt + g(t)\,dW_t$, $dY_t = h(t)\ dt + k(t)\ dW_t$, where f, h, g, k can be random, and are all assumed to be nonanticipating. Then*

$$d(XY)_t = X_t\,dY_t + Y_t\,dX_t + gk\,dt. \tag{12}$$

We won't prove Lemma 2, so we won't use it.

However, we only really need the following consequence of Lemma 2 which would obtained by taking expectations throughout. To make it more convenient, we'll prove this lemma for the case of stochastic integrals against several Wiener processes (Lemma 2 extends to this situation also).

Lemma 3 (Expected product rule for stochastic differentials)
Let $W^1, \ldots, W^m$ be independent Wiener processes and let

$$\begin{aligned} dX_t &= f(t)\,dt + g^1(t)\,dW_t^1 + \ldots + g^m(t)\,dW_t^m, \\ dY_t &= h(t)\ dt + k^1(t)\,dW_t^1 + \ldots + k^m(t)\,dW_t^m, \end{aligned} \tag{13}$$

where f, h are random and continuous and $g_1, \ldots, g_m, k_1, \ldots, k_m$ are nonrandom and continuously differentiable. We assume that $X_0, W^1, \ldots, W^m$ are independent and $Y_0, W^1, \ldots, W^m$ are independent. We assume that $f(t), h(t) \in L_2$ for each t. We also assume that f and k are nonanticipating, in the following simple sense:

$$Ef(s)\left(W_t^i - W_s^i\right) = 0,\ \ Eg(s)\left(W_t^i - W_s^i\right) = 0 \tag{14}$$

for all times s, t with $s \le t$ and all $i = 1, \ldots, m$. Then

$$E(X_tY_t) - E(X_0Y_0) = \\ E\int_0^t X_s h(s)\,ds + E\int_0^t Y_s f(s)\,ds + \sum_{i=1}^m E\int_0^t g^i(s)k^i(s)\,ds. \tag{15}$$

Proof This lemma is easy to prove by cases. Using bilinearity we can break up both sides of (15) into pieces which are simpler.

Case a Suppose that $dX = 0$. Then $X_t = X_0$ for all t. We must show that

$$EX_0\left(Y_t - Y_0\right) = E\int_0^t X_0 h(s)\,ds,$$

or in other words that

$$E\int_0^t X_0 h(s)\,ds = EX_0\int_0^t h(s)\,ds + \sum_{i=1}^m EX_0\int_0^t g^i(s)\,dW_s^i.$$

This follows since $X_0, \int_0^t g^i(s)\,dW_s^i$ are independent and $E\int_0^t g^i(s)\,dW_s^i = 0$ by equation (3) in Theorem 6.3.2 in the text. Hence the result holds.

Case b Suppose that $dY = 0$. The result holds just as in Case a.

Case c Suppose that $X_0 \equiv 0$ and $Y_0 \equiv 0$. If we can prove the result holds in this case we are done, using bilinearity to break up the general case.

Case c.1

$$dX_t = f(t)\,dt,\ dY_t = h(t)\,dt.$$

In this case the result follows from the ordinary product rule of calculus.

Case c.2

$$dX_t = g^i(t)\,dW_t^i,\ dY_t = k^i(t)\,dW_t^i.$$

In this case the result follows from equation (4) in Theorem 6.3.2 in the text.

Case c.3

$$dX_t = g^i(t)\,dW_t^i,\ dY_t = k^j(t)\,dW_t^j,\ i \neq j.$$

In this case the result follows from the independence of the processes W^i, W^j and equation (3) in Theorem 6.3.2 in the text.

Case c.4

$$dX_t = g^i(t)\,dW_t^i,\ dY_t = h(t)\,dt.$$

We must show that

$$E\left(X_t Y_t\right) = E\int_0^t X_s h(s)\,ds,$$

where

$$X_t = \int_0^t g^i(s)\,dW_s^i \equiv g^i(t)W_t^i - \int_0^t \left(g^i(s)\right)' W_s^i\,ds$$

and

$$Y_t = \int_0^t h(s)\,ds.$$

(Remember we are assuming now that $X_0 = Y_0 = 0$.)

Using equation (9) of Lemma 1, and the nonanticipating property, we see that

$$X_t - X_s \perp h(s)$$

for all $s \le t$. Thus

$$E\left(X_t Y_t\right) = E X_t \int_0^t h(s)\,ds = \int_0^t E X_t h(s)\,ds = \int_0^t E X_s h(s)\,ds = E \int_0^t X_s h(s)\,ds$$

as claimed.

Case c.5

$$dY_t = k^i(t)\,dW_t^i,\ dX_t = f(t)\,dt.$$

This is the same as Case c.4.

Using bilinearity, the lemma follows.

Handy notations

In order to work with Lemma 3 efficiently, we'll invent our own notation. This notation will never be popular, because it is not really needed as soon as one learns the real Itô formula. But it will save some writing.

Let Z be any random process such that EZ_t exists as a finite number for each t. Let u be continuous (possibly random), such that $E\left|u(s)\right| < \infty$ for each s. We write

$$d_E Z_t = u(t)\,dt \tag{16}$$

to mean that

$$EZ_t - EZ_0 = E\int_0^t u(s)\,ds. \tag{17}$$

When (13) holds we let

$$[dX_t] = f(t)\,dt,\ [dY_t] = h(t)\,dt.$$

We refer to $[dX_t]$ as the *finite total variation part* of the differential dX_t.

Finally, when (13) holds we let

$$\begin{aligned}\langle dX_t, dY_t\rangle &= \Big\langle f(t)\,dt + \sum_{i=1}^m g^i(t)\,dW_t^i, h\left(t\right)\ dt + \sum_{j=1}^m k^j(t)\,dW_t^j \Big\rangle \\ &= \sum_{i=1}^m g^i(t) k^i(t)\,dt.\end{aligned} \tag{18}$$

We refer to $\langle dX_t, dY_t \rangle$ as the *joint variation* of the differentials dX_t and dY_t.

From (18) we have the well-known Itô multiplication table for stochastic differentials:

$$\begin{aligned} \langle dt, dW_t^i \rangle &= 0, \\ \langle dW_t^i, dW_t^i \rangle &= dt, \\ \langle dW_t^i, dW_t^j \rangle &= 0, \ i \neq j. \end{aligned} \tag{19}$$

Using our notations we can now rewrite the conclusion of Lemma 3 in a somewhat more elegant form.

$$d_E\,(X_t Y_t) = X_t\ [dY_t] + Y_t\ [dX_t] + \langle dX_t, dY_t \rangle. \tag{20}$$

A skeptical person might argue that equation (20) only looks elegant because we have *hidden* the details of (15) inside our notation, but actually it is easy to calculate using (20) if we remember the two facts mentioned above: (i) the fact that $[dX_t]$ is the *finite variation part* of the differential, and (ii) the *multiplication table* given in (19).

Accordingly, equation (20) will be our preferred way of expressing Lemma 3.

Remark on multiplication table heuristics

The first equation in (19) expresses the fact that we discard differential terms that are smaller than dt, and we consider dW_t^i to have size on the order of $\sqrt{dt}$.

The second equation in (19) expresses the fact that variance of the Wiener process increases linearly with time. It also expresses Theorem 6.2.8 in the text on quadratic variation. It consistent with the idea that the size of dW_t^i is on the order of $\sqrt{dt}$.

The third equation in (19) expresses the idea that because of the independence of the mean-zero processes W^i and W^j, the quantity $dW_t^i\, dW_t^j$ should have expectation zero.

Checking the nonanticipating property

It follows from Theorem 6.3.2 in the text that integrating a deterministic function of time against a Wiener processes gives a nonanticipating process in the sense of the definition appearing in the statement of Lemma 3. Using the formula for the solution of a stochastic differential equation given in equation (26) of the next section, it follows that all the processes we have to deal with here are nonanticipating.

It is also an easy argument to show that integrating a nonanticipating process against time gives us another nonanticipating process, if we ever need to know that.

4 Solving stochastic differential equations

Let V be any process that has a differential (for example V might be Wiener process, or V might be some process constructed from a Wiener process by solving a stochastic differential equation).

We sometimes have to solve the following kind of stochastic differential equation.

$$dX_t = u(t)X_t\,dt + g(t)\,dV_t, \tag{21}$$

where u and g are deterministic functions of time, with u continuous and g continuously differentiable. A good way to solve (21) is to use an integrating factor $Z(t)$, defined by

$$Z(t) = e^{-\int_0^t u_s\,ds}. \tag{22}$$

Then $Z'(t) = -u(t)Z(t)$, and after multiplying by Z we can rewrite (21) as

$$Z(t)\,dX_t + X_t\,dZ(t) = Z(t)g(t)\,dV_t. \tag{23}$$

Now we need a lemma.

Lemma 4 (Special product rule for stochastic differentials) *Let W be Wiener process and let*

$$dX_t = f(t)\,dt + g(t)\,dV_t,\ \ dZ_t = h\,(t)\ dt,$$

where f is continuous and possibly random, h is continuous and nonrandom, and g is continuously differentiable and nonrandom. Then

$$d\,(XZ)_t = X_t\,dZ_t + Z_t\,dX_t. \tag{24}$$

Proof Assignment 7.

Finishing the solution

Using Lemma 4, equation (23) becomes

$$d\,(X_tZ(t)) = Z(t)g(t)\,dV_t, \tag{25}$$

and so we find that

$$X(t) = Z(t)^{-1}X_0 + \int_0^t Z(t)^{-1}Z(s)g(s)\,dV_s. \tag{26}$$

Solutions have the Gaussian property

Consider the special situation in which V is Gaussian, for example the situation in which V is a Wiener process. As observed in the text, the property of being a Gaussian process is preserved when taking limits. This means that the function X in (26) is Gaussian. That is, the solution of (21) is Gaussian.

5 Proof for Theorem 1

We recall that we are assuming (without loss of generality) that $EX_0 = 0$.

5.1 Introducing mysterious functions Σ_t, $\hat{X}_t$

Let Σ_t denote the solution of

$$\Sigma_t' = \left(2a - \frac{h^2}{g^2}\Sigma_t\right)\Sigma_t + c^2, \tag{27}$$

$\Sigma_0 = EX_0^2$. (The quantity Σ_t is denoted by $\bar{\Sigma}_t$ in your text.) Define

$$\beta_t = \frac{h}{g^2}\Sigma_t,\ \alpha_t = a - \beta_t h = a - \frac{h^2}{g^2}\Sigma_t. \tag{28}$$

Let $\hat{X}$ denote the solution of the equation

$$d\hat{X}_t = \alpha_t \hat{X}_t\,dt + c\,dW_t - \beta_t g\,dB_t \tag{29}$$

with $\hat{X}_0 = X_0$. By (26) this solution exists and is unique.

5.2 Finding $E\hat{X}_t^2$

Using Lemma 3,

$$\begin{aligned} d_E\hat{X}_t^2 &= 2\hat{X}_t\left[d\hat{X}_t\right] + \langle d\hat{X}_t, d\hat{X}_t\rangle \\ &= 2\alpha_t\hat{X}_t^2\,dt + c^2\,dt + \beta_t^2 g^2\,dt. \end{aligned}$$

Let $\Gamma_t = E\hat{X}_t^2$. We have shown that

$$d\Gamma_t = 2\alpha_t\Gamma_t\,dt + c^2\,dt + \beta_t^2 g^2\,dt.$$

That is,

$$\Gamma_t' = 2\alpha_t\Gamma_t + c^2 + \beta_t^2 g^2.$$

Also $\Gamma_0 = EX_0^2$.

We can easily rewrite (27) as

$$\Sigma_t' = 2\alpha_t\Sigma_t + c^2 + \beta_t^2 g^2.$$

Also $\Sigma_0 = EX_0^2$. It follows that

$$(\Sigma_t - \Gamma_t)' = 2\alpha_t\left(\Sigma_t - \Gamma_t\right),$$

and $\Sigma_0 - \Gamma_0 = 0$. Hence $\Gamma_t = \Sigma_t$ for all t. That is,

$$E\hat{X}_t^2 = \Sigma_t \text{ for all } t. \tag{30}$$

5.3 Finding $E\hat{X}_t X_t$

Using Lemma 3,

$$d_E \hat{X}_t X_t = \hat{X}_t\,[dX_t] + X_t \left[d\hat{X}_t\right] + \langle d\hat{X}_t, dX_t\rangle$$
$$= a\hat{X}_t X_t\,dt + \alpha_t X_t \hat{X}_t\,dt + c^2\,dt.$$

Let $\Lambda_t = E\hat{X}_t X_t$. We have shown that

$$d\Lambda_t = (a + \alpha_t)\,\Lambda_t\,dt + c^2\,dt.$$

That is,

$$\Lambda_t' = (a + \alpha_t)\,\Lambda_t + c^2.$$

Also $\Lambda_0 = EX_0^2$.

Comparing this with (27), we see that

$$(\Sigma_t - \Lambda_t)' = (a + \alpha_t)\,(\Sigma_t - \Lambda_t)\,,$$

and $\Sigma_0 - \Lambda_0 = 0$. Hence $\Lambda_t = \Sigma_t$ for all t. That is,

$$E\hat{X}_t X_t = \Sigma_t \text{ for all} t. \tag{31}$$

5.4 Finding $E\hat{X}_t Y_t$

Using Lemma 3,

$$d_E \hat{X}_t Y_t = \hat{X}_t\,[dY_t] + Y_t \left[d\hat{X}_t\right] + \langle d\hat{X}_t, dY_t\rangle$$
$$= h\hat{X}_t X_t\,dt + \alpha_t Y_t \hat{X}_t\,dt - \beta_t g^2\,dt.$$

Let $\Psi_t = E\hat{X}_t Y_t$. Using the fact that $E\hat{X}_t X_t = \Sigma_t$, we have shown that

$$d\Psi = (h\Sigma_t + \alpha_t)\,\Psi_t\,dt - \beta_t g^2\,dt = \alpha_t \Psi_t\,dt.$$

That is,

$$\Psi_t' = \alpha_t \Psi_t$$

Also $\Psi_0 = 0$ since $Y_0 = 0$.

Hence $\Psi_0 = 0$ for all t. That is,

$$E\hat{X}_t Y_t = 0 \text{ for all } t. \tag{32}$$

5.5 Finding $E\hat{X}_t Y_s$

Using a very easy case of Lemma 3, for $t \geq s$ we have

$$\begin{aligned} d_E \hat{X}_t Y_s &= Y_s \left[d\hat{X}_t \right] \\ &= \alpha_t Y_s \hat{X}_t \, dt. \end{aligned}$$

Let $\Theta_t = E\hat{X}_t Y_s$. We we have shown that

$$d\Theta_t = \alpha_t \Theta_t \, dt$$

That is,

$$\Theta_t' = \alpha_t \Theta_t.$$

Also $\Theta_s = E\hat{X}_s Y_s = 0$.

Hence $\Theta_t = 0$ for all $t \geq s$.

$$E\hat{X}_t Y_s = 0 \text{ for all } t \geq s. \tag{33}$$

5.6 $\hat{X}_t \perp \sigma\left(Y_{[0,t]}\right)$

We see from (33) that for any times $s_1, \ldots, s_\ell \leq t$ we have

$$\hat{X}_t \perp Y_{s_1}, \ldots, Y_{s_\ell}.$$

By the Gaussian property, $\hat{X}_t, Y_{s_1}, \ldots, Y_{s_\ell}$ are independent. By a simple limiting argument, for any element $V \in \sigma\left(Y_{[0,t]}\right)$, $\hat{X}_t$ and V are independent. Since $E\hat{X}_t = 0$, $\hat{X}_t \perp V$. Thus $\hat{X}_t \perp \sigma\left(Y_{[0,t]}\right)$.

5.7 $X_t - \hat{X}_t \in \sigma\left(Y_{[0,t]}\right)$

We have

$$\begin{aligned} d\left(X_t - \hat{X}_t\right) &= dX_t - d\hat{X}_t = aX_t \, dt + c \, dW_t - \alpha_t \hat{X}_t \, dt - c \, dW_t + \beta_t g \, dB_t \\ &= \alpha \left(X_t - \hat{X}_t\right) dt + \beta_t h X_t \, dt + \beta_t g \, dB_t \qquad (34) \\ &= \alpha \left(X_t - \hat{X}_t\right) dt + \beta_t \, dY_t. \end{aligned}$$

Solving this equation using an integrating factor, as in Section 4, we find that

$$X(t) - \hat{X}_t = \int_0^t Z(t)^{-1} Z(s) g(s) \, dY_s. \tag{35}$$

where

$$Z_t = e^{-\int_0^t \alpha_s \, ds}.$$

By Theorem 6.3.3 in the text, the stochastic integral in (35) can be approximated by linear combinations of random variables Y_s, $s \in [0,t]$. Hence

$$X_t - \hat{X}_t \in \sigma\left(Y_{[0,t]}\right).$$

5.8 Finishing the proof

Since $\hat{X}_t \perp \sigma\left(Y_{[0,t]}\right)$ and also $X_t - \hat{X}_t \in \sigma\left(Y_{[0,t]}\right)$, it follows by definition that $X_t - \hat{X}_t = \bar{X}_t$.

By (34), equation (5) holds.

We showed earlier that $\Sigma_t = E\hat{X}_t^2$. Hence by (4) we now know that $\bar{\Sigma}_t = \Sigma_t$. Thus by (27) we have that equation (6) holds.

This completes the proof of Theorem 1.