

Lecture 4. Linear filtering in continuous time

4.1

4.1. The Kalman-Bucy filter: scalar case

Consider the following system of SDEs:

$$(4.1) \quad dX_t = a_t X_t dt + b_t dW_t$$

$$(4.2) \quad dY_t = A_t X_t dt + B_t dV_t$$

where W and V are independent Wiener processes and the scalar coefficients are deterministic functions of time, such that (4.1) has a unique strong solution.

We also assume that X_0, Y_0 are independent of (W, V) and are square integrable. Moreover, we will also assume that

$$(4.3) \quad B_t^2 \geq c > 0 \text{ for some constant } c.$$

$\mathcal{H}_t^Y :=$ the closed linear subspace generated by the random variables $Y_s, s \leq t$. (+ constants)

$\hat{E}[\cdot | \mathcal{H}_t^Y] :=$ the orthogonal projection onto $\mathcal{H}_t^Y$.

As discussed in Lecture 2, $\hat{X}_t := \hat{E}[X_t | \mathcal{H}_t^Y]$ is the best linear estimate of X_t , given the observations $\{Y_s, s \leq t\}$.

Theorem 4.1 (Kalman-Bucy filter)

The optimal linear estimate $\hat{X}_t$ and the corresponding mean square error $P_t = E[(X_t - \hat{X}_t)^2]$ satisfy the equations

$$(4.4) \quad \begin{cases} d\hat{X}_t = a_t \hat{X}_t dt + \frac{P_t A_t}{B_t^2} (dY_t - A_t \hat{X}_t dt) \\ \dot{P}_t = 2a_t P_t + b_t^2 - \frac{A_t P_t^2}{B_t^2} \end{cases} \quad (4.2)$$

subject to

$$(4.5) \quad \begin{cases} \hat{X}_0 = E[X_0] + \text{cov}(X_0, Y_0) \text{cov}(Y_0)^{-1} (Y_0 - E[Y_0]) \\ P_0 = \text{cov}(X_0) - \text{cov}(X_0, Y_0)^2 \text{cov}(Y_0)^{-1} \end{cases}$$

Proof

Step 1 (getting rid of $\hat{X}_0$)

It is easier to treat the case $\hat{X}_0 \equiv 0$ and we claim that is enough to prove the theorem under this assumption.

Let

$$X'_t = X_t - \hat{X}_0 \exp\left(\int_0^t a_s ds\right)$$

$$Y'_t = Y_t - \int_0^t A_s \hat{X}_0 \exp\left(\int_0^s a_u du\right) ds$$

The process (X', Y') satisfies

$$\begin{aligned} dX'_t &= dX_t - \hat{X}_0 d\left(\exp\left(\int_0^t a_s ds\right)\right) \\ &= a_t X_t + b_t dW_t - \hat{X}_0 e^{\int_0^t a_s ds} a_t dt \\ &= a_t X'_t + b_t dW_t \end{aligned}$$

$$\begin{aligned} dY'_t &= dY_t - d\left(\int_0^t A_s \hat{X}_0 \exp\left(\int_0^s a_u du\right) ds\right) \\ &= A_t X_t dt + B_t dV_t - A_t \hat{X}_0 e^{\int_0^t a_s ds} dt \\ &= A_t X'_t dt + B_t dV_t \end{aligned}$$

Subject to $X'_0 = X_0 - \hat{X}_0$ and $Y'_0 = Y_0$ 4.3

Clearly $h^Y_t = h^{Y'}_t$ (since $\hat{X}_0$ is an affine transform of Y_0), and hence

$$\begin{aligned}\hat{X}_t &= \hat{E}[X_t | h^Y_t] = \hat{E}[X_t | h^{Y'}_t] \\ &= \hat{E}[X'_t | h^{Y'}_t] + \hat{X}_0 \exp\left(\int_0^t a_s ds\right).\end{aligned}$$

Note that $\hat{E}[X'_0 | Y'_0] = \hat{E}[X_0 - \hat{X}_0 | Y_0] = 0$

Suppose that $\hat{X}'_t = \hat{E}[X'_t | h^{Y'}_t]$ and $P_t = E[(X'_t - \hat{X}'_t)^2]$ satisfy (4.4), subject to $\hat{X}'_0 = 0$ and $P'_0 = E[(X'_0 - \hat{X}'_0)^2]$.

Then,

$$d\hat{X}_t = d\hat{X}'_t + a_t \hat{X}_0 \exp\left(\int_0^t a_s ds\right) dt$$

$$= a_t \hat{X}'_t dt + \frac{P_t A_t}{B_t^2} (dY_t - A_t \hat{X}'_t dt)$$

$$+ a_t \hat{X}_0 \exp\left(\int_0^t a_s ds\right) dt$$

$$= a_t \hat{X}_t dt - a_t \hat{X}_0 \exp\left(\int_0^t a_s ds\right) dt$$

$$+ \frac{P_t A_t}{B_t^2} (dY_t - A_t \hat{X}_0 \exp\left(\int_0^t a_s ds\right) - A_t \hat{X}_t dt + A_t \hat{X}_0 \exp\left(\int_0^t a_s ds\right) dt)$$

$$= a_t \hat{X}_t dt + \frac{P_t A_t}{B_t^2} (dY_t - A_t \hat{X}_t dt)$$

Moreover,

$$P_t = E[(X_t - \hat{X}_t)^2]$$

$$= E\left[\left(X'_t + \hat{X}_0 \exp\left(\int_0^t a_s ds\right) - \hat{X}'_t - \hat{X}_0 \exp\left(\int_0^t a_s ds\right)\right)^2\right]$$

$$= E[(X'_t - \hat{X}'_t)^2] = P'_t,$$

that is, P_t satisfies the equation from (4.4)

Step 2 (the general form of the estimate)

From here on $\hat{E}[X_t | Y_0] = 0$ assumed p.a.s.

Let $\pi^n = \{0 = t_0 < t_1 < \dots < t_n = T\}$ be a partition of $[0, T]$ and define

$$\mathcal{L}_t^Y(n) := \overline{\text{span}} \{Y_{t_0}, \dots, Y_{t_n}\}$$

This space coincides with

$$\overline{\text{span}} \{Y_{t_0}, Y_{t_1} - Y_{t_0}, \dots, Y_{t_n} - Y_{t_{n-1}}\}$$

and, therefore,

$$\begin{aligned} \hat{E}[X_t | \mathcal{L}_t^Y(n)] &= \hat{E}[X_t | Y_0] + \sum_{j=0}^{n-1} g_j (Y_{t_{j+1}} - Y_{t_j}) \\ &= \hat{E}[X_t | Y_0] + \int_0^t G^n(t, s) dY_s \end{aligned}$$

where g_j are real numbers and $G(t, s) = \sum_{j=0}^{n-1} g_j 1_{[t_j, t_{j+1})}^{(s)}$.

Since $\mathcal{L}_t^Y$ is a closed subspace (and $\mathcal{L}_t^Y(n) \subseteq \mathcal{L}_t^Y$),

we have that

$$L^2\text{-}\lim_{n \rightarrow \infty} \hat{E}[X_t | \mathcal{L}_t^Y(n)] = \hat{E}[X_t | \mathcal{L}_t^Y]$$

and, hence,

$$E \left[\left(\int_0^t G^n(t, s) dY_s - \int_0^t G^m(t, s) dY_s \right)^2 \right] \xrightarrow{n, m \rightarrow \infty} 0$$

//

$$E \left[\left(\int_0^t \underbrace{(G^n(t, s) - G^m(t, s))}_{\Delta^{n,m}(t, s)} A_s X_s ds + \int_0^t (G^n(t, s) - G^m(t, s)) B_s dV_s \right)^2 \right]$$

$$\begin{aligned} &= E \left[\left(\int_0^t \Delta^{n,m}(t, s) A_s X_s ds \right)^2 \right] + E \left[\left(\int_0^t \Delta^{n,m}(t, s) B_s dV_s \right)^2 \right] \\ &\quad + 2 E \left[\int_0^t \Delta^{n,m}(t, s) A_s X_s ds \right] \cdot E \left[\int_0^t \Delta^{n,m}(t, s) B_s dV_s \right] \end{aligned}$$

Step 2 (the general form of the estimate)

4.9a

From here on $\hat{E}[X_t | Y_0] = x_t$ assumed P-a.s.

Let $\pi^n = \{0 = t_0^{(n)} < t_1^{(n)} < \dots < t_{2^n}^{(n)} = t\}$ be the dyadic partition of $[0, t]$ ($t_k^{(n)} = \frac{k}{2^n} t$) and define

$$\mathcal{L}_t^Y(n) := \overline{\text{span}} \{ Y_{t_0^{(n)}}, \dots, Y_{t_{2^n}^{(n)}} \}$$

This space coincides with

$$\overline{\text{span}} \{ Y_{t_0^{(n)}}, Y_{t_1^{(n)}} - Y_{t_0^{(n)}}, \dots, Y_{t_{2^n}^{(n)}} - Y_{t_{2^n-1}^{(n)}} \}$$

and, therefore, by the normal correlation theorem

$$\begin{aligned} \hat{E}[X_t | \mathcal{L}_t^Y(n)] &= \hat{E}[X_t | Y_0] + \sum_{j=0}^{2^n-1} g_j (Y_{t_{j+1}^{(n)}} - Y_{t_j^{(n)}}) \\ &= \hat{E}[X_t | Y_0] + \int_0^t G^n(t, s) dY_s \end{aligned}$$

where $g_j = g_j(t)$ real numbers and $G(t, s) = \sum_{j=0}^{2^n-1} g_j 1_{[t_j, t_{j+1}^{(n)}]}(s)$.

$$\text{Moreover } \hat{E}[X_t | \mathcal{L}_t^Y(n)] = E[X_t | \mathcal{F}_t^Y(n)]$$

$$\text{where } \mathcal{F}_t^Y(n) = \sigma(Y_{t_0^{(n)}}, Y_{t_1^{(n)}} - Y_{t_0^{(n)}}, \dots, Y_{t_{2^n}^{(n)}} - Y_{t_{2^n-1}^{(n)}})$$

$$\text{Since } \mathcal{F}_t^Y(n) \uparrow \mathcal{F}_t^Y,$$

by Lévy's theorem we have that P-a.s.

$$E[X_t | \mathcal{F}_t^Y(n)] \longrightarrow E[X_t | \mathcal{F}_t^Y] \quad (*)$$

Let $\dots, \mathcal{F}_{-2}, \mathcal{F}_{-1}, \mathcal{F}_0, \mathcal{F}_1, \mathcal{F}_2, \dots$ be a nondecreasing sequence of sub- σ -algebras of $\mathcal{F}$. Define

$$\mathcal{F}_\infty = \sigma\left(\bigcup_n \mathcal{F}_n\right), \quad \mathcal{F}_{-\infty} = \bigcap_n \mathcal{F}_n$$

Theorem (Lévy)

4.2.4

Let ξ be a r.v. with $E[|\xi|] < +\infty$. Then, p.a.s.,

$$E[\xi | \mathcal{F}_n] \xrightarrow{n \rightarrow \infty} E[\xi | \mathcal{F}_\infty]$$

$$E[\xi | \mathcal{F}_n] \xrightarrow{n \rightarrow -\infty} E[\xi | \mathcal{F}_{-\infty}]$$

In addition, the sequence $\{E[X_t | \mathcal{F}_t^Y(n)]^2\}_{n \in \mathbb{N}}$ is uniformly integrable and, hence, (*) yields that

$$E[X_t | \mathcal{F}_t^Y(n)] \xrightarrow[n \rightarrow +\infty]{L^2} E[X_t | \mathcal{F}_t^Y]$$

Then,

$$E\left[\left(\int_0^t G^n(t,s) dY_s - \int_0^t G^m(t,s) dY_s\right)^2\right] \xrightarrow{m,n \rightarrow \infty} 0$$

//

$$E\left[\left(\int_0^t \underbrace{(G^n(t,s) - G^m(t,s))}_{\Delta^{n,m}(t,s)} A_s X_s ds + \int_0^t (G^n(t,s) - G^m(t,s)) B_s dV_s\right)^2\right]$$

$$= E\left[\left(\int_0^t \Delta^{n,m}(t,s) A_s X_s ds\right)^2\right] + E\left[\left(\int_0^t \Delta^{n,m}(t,s) B_s dV_s\right)^2\right] \\ + 2 E\left[\int_0^t \Delta^{n,m}(t,s) A_s X_s ds\right] \cdot E\left[\int_0^t \cancel{\Delta^{n,m}(t,s)} B_s dV_s\right] \quad (\text{cf})$$

(*) Independence of X and V and that the Itô 4.5 integral has 0 expectation.

$$= E \left[\left(\int_0^t \Delta^{mm}(t,s) A_s X_s ds \right)^2 \right] + E \left[\int_0^t |\Delta^{mm}(t,s) B_s|^2 ds \right]$$

Itô isometry and B and Δ^{mm} being deterministic.

Hence, due to $B_s^2 \geq C > 0$, we can write

$$\int_0^t |G^n(t,s) - G^m(t,s)|^2 ds \leq \frac{1}{C} \int_0^t |G^n(t,s) - G^m(t,s)|^2 B_s^2 ds$$

$\xrightarrow{n,m \rightarrow \infty} 0$. So G^n is a Cauchy sequence and converges to a limit $G(t,s)$, so that

$$\hat{E}[X_t | \mathcal{F}_t^Y] = \hat{E}[X_t | \mathcal{Y}_0] + \int_0^t G(t,s) dY_s$$

Step 3 (using orthogonality)

Recall that we assume $\hat{E}[X_0 | \mathcal{Y}_0] = 0$, P-a.s.,

so that $E[X_t] = 0$ and $\hat{E}[X_t | \mathcal{Y}_0] = 0$. (Exercise)

Claim: the function $G(t,s)$, satisfies the Wiener-Hopf equation

$$(4.6) \quad K(t,u) A_u = \int_0^t G(t,s) A_s K(s,u) A_u ds + G(t,u) B_u^2, \quad 0 \leq u \leq t.$$

where $K(t,s) = \text{cov}(X_t, Y_s)$.

Proof of claim: By orthogonality property of the orthogonal projection, for any fixed $t \in [0, T]$ and any measurable and bounded deterministic function λ we have

$$E \left[(X_t - \hat{E}[X_t | \mathcal{F}_t^Y]) \int_0^t \lambda_s dY_s \right] = E \left[(X_t - \int_0^t G(t,s) dY_s) \int_0^t \lambda_s dY_s \right] = 0$$

Then, (4.6) holds since

11.6

$$\begin{aligned} E\left[X_t \int_0^t \lambda_u dY_u\right] &= E\left[X_t \int_0^t \lambda_u A_u X_u du\right] \\ &\quad + E\left[X_t \int_0^t \lambda_u B_u dV_u\right] \\ &= \int_0^t \lambda_u A_u K(t, u) du \end{aligned}$$

and

$$\begin{aligned} E\left[\int_0^t G(t, s) dY_s \int_0^t \lambda_u dY_u\right] &= E\left[\int_0^t G(t, s) A_s X_s ds \int_0^t \lambda_u dY_u\right] \\ &\quad + E\left[\int_0^t G(t, s) B_s dV_s \int_0^t \lambda_u dY_u\right] \\ &= E\left[\int_0^t G(t, s) A_s X_s ds \int_0^t \lambda_u A_u X_u du\right] \\ &\quad + E\left[\int_0^t G(t, s) A_s X_s ds \int_0^t \lambda_u B_u dV_u\right] \\ &\quad + E\left[\int_0^t G(t, s) B_s dV_s \int_0^t \lambda_u A_u X_u du\right] \\ &\quad + E\left[\int_0^t G(t, s) B_s dV_s \int_0^t \lambda_u B_u dV_u\right] = \textcircled{1} + \textcircled{2} + \textcircled{3} + \textcircled{4} \end{aligned}$$

$$\textcircled{1} \text{ Fubini: } \int_0^t \int_0^t G(t, s) A_s K(s, u) A_u \lambda_u du ds$$

② and ③. Independence of X and V and Itô stochastic integral being centered $\Rightarrow = 0$.

$$\textcircled{4} \text{ It\^o isometry: } \int_0^t G(t, s) \lambda_s B_s^2 ds$$

for arbitrary λ .

Under assumption $B_t^2 \geq C > 0$, the Wiener-Hopf equation has a unique solution. 4.7

Suppose it is not the case, i.e., both $G_2(t, 1)$ and $G_2(t, 1)$ satisfy (4.6) and let $\Delta(t, s) = G_2(t, s) - G_2(t, 1)$ then $\Delta(t, 1)$ satisfies

$$\int_0^t \Delta(t, s) A_s K(s, u) A_u ds + \Delta(t, u) B_u^2 = 0 \quad 0 \leq u \leq t.$$

Multiplying this equation by $\Delta(t, u)$ and integrating with respect to u we get

$$\int_0^t \int_0^t \Delta(t, u) A_u K(s, u) \Delta(t, s) A_s ds du + \int_0^t \Delta^2(t, u) B_u^2 du = 0$$

Since $K(s, u)$ is a nonnegative definite function the term with the double integral is always nonnegative. Thus for $t \in [0, T]$ we have

$$\int_0^t \Delta^2(t, u) B_u^2 du = 0 \Rightarrow \Delta^2(t, u) = 0 \text{ a.e.}$$

Step 4 (Solving the Wiener-Hopf equation)

We will look for differentiable solutions. Thanks to the uniqueness this will be the solution.

Taking derivative with respect to t in (4.6) we get

$$\begin{aligned} (*) \quad \frac{\partial}{\partial t} K(t, u) A_u &= G(t, t) A_t K(t, u) A_u \\ &+ \int_0^t \frac{\partial}{\partial t} G(t, s) A_s K(s, u) A_u ds + \frac{\partial}{\partial t} G(t, u) B_u^2 \end{aligned}$$

Thanks to Exercise 10, Chapter 9 in Chigovsky's notes we have that

$$\frac{\partial}{\partial t} K(t, u) = a_t K(t, u), \quad K(u, u) = E[X_u^2]$$

and, hence, $(*)$ reads

$$K(t, u) A_u (a_t - G(t, t) A_t) - \int_0^t \frac{\partial}{\partial t} G(t, s) A_s K(s, u) A_u ds \\ - \frac{\partial}{\partial t} G(t, u) B_u^2 = 0$$

Next, using the expression for $K(t, u) A_u$ from (4.6), one gets

$$\left(\int_0^t G(t, s) A_s K(s, u) A_u ds + G(t, u) B_u^2 \right) (a_t - G(t, t) A_t) \\ - \int_0^t \frac{\partial}{\partial t} G(t, s) A_s K(s, u) A_u ds - \frac{\partial}{\partial t} G(t, u) B_u^2 = 0$$

or

$$\int_0^t \left\{ G(t, s) (a_t - G(t, t) A_t) - \frac{\partial}{\partial t} G(t, s) \right\} A_s K(s, u) A_u ds \\ + \left\{ G(t, u) (a_t - G(t, t) A_t) - \frac{\partial}{\partial t} G(t, u) B_u^2 \right\} = 0$$

Multiplying the last equality by

$$\psi(t, u) := G(t, u) (a_t - G(t, t) A_t) - \frac{\partial}{\partial t} G(t, u) B_u^2$$

and integrating with respect to u we get

$$\int_0^t \int_0^t \underbrace{\psi(t, s) A_s K(s, u) \psi(t, u) A_u}_{\geq 0} ds du + \int_0^t \psi^2(t, u) B_u^2 du = 0,$$

$\checkmark$ $K(s, u)$ is a nonnegative definite function

$$\Rightarrow \int_0^t \psi^2(t, u) B_u^2 du = 0 \Rightarrow \psi(t, u) = 0 \quad \forall t \geq 0, u \text{ a.e.} \quad (4.9)$$

Hence, $G(t, s)$ satisfies

$$(4.7) \quad \frac{\partial}{\partial t} G(t, s) = G(t, s) (a_t - G(t, t) A_t)$$

With $u = t$ in (4.6), we get

$$0 = K(t, t) A_t - A_t \int_0^t G(t, s) A_s K(s, t) ds - G(t, t) B_t^2,$$

which implies

$$0 = A_t E \left[X_t \left(X_t - \int_0^t G(t, s) A_s X_s ds \right) \right] - G(t, t) B_t^2$$

$$dY_s = B_s dV_s$$

$$= A_t E \left[X_t \left(X_t - \int_0^t G(t, s) dY_s \right) \right] + A_t E \left[X_t \int_0^t G(t, s) B_s dV_s \right] - G(t, t) B_t^2$$

$\int_0^t$ X and V indep.
and zero expectation of
its integral

$$= A_t E \left[\left\{ \int_0^t G(t, s) dY_s + X_t - \int_0^t G(t, s) dY_s \right\} \left\{ X_t - \int_0^t G(t, s) dY_s \right\} \right] - G(t, t) B_t^2$$

$$= A_t E \left[\int_0^t G(t, s) dY_s \left(X_t - \int_0^t G(t, s) dY_s \right) \right] + A_t E \left[\left(X_t - \int_0^t G(t, s) dY_s \right)^2 \right] - G(t, t) B_t^2$$

$\int_0^t$ orthogonality

$$= A_t P_t - G(t, t) B_t^2 \quad \left(P_t = E[(X_t - \hat{X}_t)^2] = E[(X_t - \int_0^t G(t, s) dY_s)^2] \right)$$

Hence, $G(t, t) = \frac{A_t P_t}{B_t^2}$ and (4.7) reads

$$(4.8) \quad \frac{\partial}{\partial t} G(t, s) = G(t, s) \left(a_t - \frac{A_t^2 P_t}{B_t^2} \right)$$

Since (4.8) is linear, it admits the following solution

$$G(t, s) = \Phi(s, t) G(s, s) \quad 0 \leq s \leq t, \text{ where } \Phi(s, t)$$

is the fundamental solution of (4.8)

14.10

Then,

$$\begin{aligned}\hat{X}_t &= \int_0^t G(t,s) dY_s = \int_0^t \Phi(s,t) G(s,s) dY_s \\ &= \Phi(0,t) \int_0^t \Phi^{-1}(0,s) G(s,s) dY_s\end{aligned}$$

and applying Itô formula one gets

$$d\hat{X}_t = \int_0^t \Phi^{-1}(0,s) G(s,s) dY_s \frac{d}{dt} \Phi(0,t) dt$$

$$+ \underbrace{\Phi(0,t) \Phi^{-1}(0,t) G(t,t) dY_t}_{\text{I.I.}} + d\left[\Phi(0,\cdot), \int_0^\cdot \Phi^{-1}(0,s) G(s,s) dY_s\right]_t$$

(0)
 $\Phi(0,t)$ has continuous path and finite variation.

$$= \int_0^t \Phi^{-1}(0,s) G(s,s) dY_s \left(\alpha_t - \frac{A_t^2 P_t}{B_t^2} \right) \Phi(0,t) dt + G(t,t) dY_t$$

$$= \alpha_t \hat{X}_t dt - \frac{A_t^2 P_t}{B_t^2} \hat{X}_t dt + G(t,t) dY_t$$

$$= \alpha_t \hat{X}_t dt + \frac{A_t P_t}{B_t^2} (dY_t - A_t \hat{X}_t dt) \quad \left(G(t,t) = \frac{A_t P_t}{B_t^2} \right)$$

which is the first equation in (4.4).

For the second equation in (4.4), consider the process

$$D_t = X_t - \hat{X}_t, \text{ which satisfies}$$

$$dD_t = dX_t - d\hat{X}_t$$

$$= \alpha_t X_t dt + b_t dW_t - \alpha_t \hat{X}_t dt - \frac{P_t A_t}{B_t^2} (dY_t - A_t \hat{X}_t dt)$$

$$= \alpha_t D_t dt + b_t dW_t - \frac{A_t P_t}{B_t^2} (A_t X_t + B_t dV_t - A_t \hat{X}_t dt)$$

$$= \left(\alpha_t - \frac{A_t^2 P_t}{B_t^2} \right) D_t dt + b_t dW_t - \frac{A_t P_t}{B_t} dV_t$$

Now, applying formula to D_t^2 one gets (4.12)

$$dD_t^2 = 2 P_t dD_t + d[D, D]_t$$

$$\begin{aligned} &= 2 \left(a_t - \frac{A_t^2 P_t}{B_t^2} \right) D_t^2 dt + 2 b_t D_t dW_t - 2 \frac{A_t P_t D_t}{B_t} dV_t \\ &\quad + \underbrace{\left(b_t^2 + \frac{A_t^2 P_t^2}{B_t^2} \right) dt}_{d[D, D]_t} \end{aligned}$$

and taking expectation we can write

$$\begin{aligned} dP_t &= 2 \left(a_t - \frac{A_t^2 P_t}{B_t^2} \right) P_t dt + \left(b_t^2 + \frac{A_t^2 P_t^2}{B_t^2} \right) dt \\ &= 2a_t P_t dt + b_t^2 dt - \frac{A_t^2 P_t^2}{B_t^2} dt \end{aligned}$$

subject to $P_0 = E[(X_0 - \hat{X}_0)^2]$.

Remark 4.2

- The Kalman-Bucy filter is a linear SDE with time varying coefficients, which depend on P_t .
- P_t satisfies a Riccati equation (independent of X)
- The innovation process

$$\overline{W}_t := \int_0^t \frac{dY_s - A_s \hat{X}_s ds}{B_s}$$

has uncorrelated increments, and in the case of Gaussian (X_0, Y_0) is a Wiener process with respect to the filtration $\mathcal{F}_t^Y$.

Example 4.3

Consider the system

$$dX_t = a X_t + b dW_t$$

$$X_0 \in L^2(P)$$

$$dY_t = A X_t + B dV_t$$

$$Y_0 = 0$$

The KB filter is

$$(4.9) \quad \begin{cases} d\hat{X}_t = a \hat{X}_t dt + \frac{P_t A}{B^2} (dY_t - A \hat{X}_t dt), & \hat{X}_0 = E[X_0] \\ \frac{dP_t}{dt} = 2a P_t + b^2 - \frac{A^2 P_t^2}{B^2}, & P_0 = \text{Var}[X_0]. \end{cases}$$

Consider the quadratic equation

$$(4.10) \quad 2a P + b^2 - \frac{A^2 P^2}{B^2} = 0$$

If $\boxed{A \neq 0 \text{ and } b \neq 0}$ are assumed, then it has two solutions

$$P_{\pm} = \frac{B^2}{A^2} \left(a \pm \sqrt{a^2 + \frac{A^2 b^2}{B^2}} \right)$$

with $P_- < 0$ and $P_+ > 0$.

Consider the suboptimal filter

$$d\tilde{X}_t = a \tilde{X}_t dt + \frac{A P_+}{B^2} (dY_t - A \tilde{X}_t dt), \quad \tilde{X}_0 = 0$$

The error process $\tilde{D}_t = X_t - \tilde{X}_t$, satisfies

$$d\tilde{D}_t = \left(a - \frac{A^2 P_+}{B^2} \right) \tilde{D}_t dt + b dW_t + \frac{A P_+}{B} dV_t, \quad \tilde{D}_0 = X_0$$

Since

$$(4.11) \quad a - \frac{A^2 P_+}{B^2} = a - \left(a + \sqrt{a^2 + \frac{A^2 b^2}{B^2}} \right) = -\sqrt{a^2 + \frac{A^2 b^2}{B^2}} < 0$$

the mean square error of this filter is bounded, i.e., $\sup_{t \geq 0} E[\tilde{D}_t^2] < +\infty$ (Exercise)

Therefore, by the optimality of $\hat{X}_t$ we have that $\sup_{t \geq 0} P_t \leq \sup_{t \geq 0} E[\tilde{D}_t^2] < +\infty$.

The function $R_t := P_t - P_+$ satisfies

$$dR_t = \left(2a R_t - \frac{A^2}{B^2} (P_t^2 - P_+^2) \right) dt = \left(2a R_t - \frac{A^2}{B^2} R_t (P_t + P_+) \right) dt$$

and, hence,

$$|R_t| = |R_0| \exp \left\{ 2a t - \frac{A^2}{B^2} \int_0^t (P_s + P_+) ds \right\}$$

$$\leq |R_0| \exp \left\{ 2a t - \frac{A^2}{B^2} P_+ t \right\} = |R_0| \exp \left\{ a t - \sqrt{a^2 + \frac{A^2 P_+^2}{B^2}} t \right\} \xrightarrow{t \rightarrow \infty} 0$$

due to (4.12). Therefore, if $A \neq 0$ and $b \neq 0$, the solution of the Riccati equation stabilizes and

$$P_\infty = \lim_{t \rightarrow \infty} P_t = P_+.$$

If $A=0$ and $b \neq 0$ then $\hat{X}_t = E[X_0] e^{at} = E[X_t]$,

$P_t = E[(X_t - E[X_t])^2]$, and the limit P_∞ exists and it is finite if $a < 0$, otherwise P_t grows to infinity

If $A \neq 0$ and $b=0$ then $P_\infty = 0$, either if

$a < 0$ (since $X_t \rightarrow 0$ in L^2) or if $a > 0$ (since $\frac{a}{A} e^{-at} Y_t \rightarrow X_0$ in L^2) or if $a=0$ (since $A^{-1} \frac{Y_t}{t} \rightarrow X_0$ in L^2).

Unlike in the discrete time case, the scalar Riccati equation in (4.9) has an explicit solution,

$$P_t = \frac{\alpha_- - K \alpha_+ \exp((\alpha_+ - \alpha_-) A^2 t / B^2)}{1 - K \exp((\alpha_+ - \alpha_-) A^2 t / B^2)},$$

$$\alpha_{\pm} = \frac{1}{A^2} (a B^2 \pm B \sqrt{a^2 B^2 + A^2 b^2})$$

$$K = \frac{P_0 - \alpha_-}{P_0 - \alpha_+}$$

2. The Kalman-Bucy filter: the general case

Let $X = (X_t)_{t \in [0, T]} \in \mathbb{R}^m$ and $Y = (Y_t)_{t \in [0, T]} \in \mathbb{R}^n$ be the processes generated by the system of linear SDEs

$$dX_t = (a_0(t) + a_L(t)X_t + a_2(t)Y_t)dt + b_L(t)dW_t + b_2(t)dV_t$$

$$dY_t = (A_0(t) + A_L(t)X_t + A_2(t)Y_t)dt + B_L(t)dW_t + B_2(t)dV_t$$

with W and V two independent multidimensional Wiener processes and subject to a square integrable random vector (X_0, Y_0) independent of (W, V) .

The coefficients are deterministic matrix functions of appropriate dimensions, such that the system has a unique strong solution and $(B \circ B) := B_L B_L^T + B_2 B_2^T$ is uniformly nonsingular matrix

Theorem 4.4

The orthogonal projection $\hat{X}_t = \hat{E}[X_t | \mathcal{Y}_t]$ and the corresponding error covariance matrix $P_t = E[(X_t - \hat{X}_t)(X_t - \hat{X}_t)^T]$

satisfy the Kalman-Bucy equations

$$d\hat{X}_t = (a_0 + a_L \hat{X}_t + a_2 Y_t)dt + (b \circ B + P_t A_L^T)(B \circ B)^{-1} \{dY_t - (A_0 + A_L \hat{X}_t + A_2 Y_t)dt\}$$

$$\dot{P}_t = a_L P_t + P_t a_L^T + b \circ b - (b \circ B + P_t A_L^T)(B \circ B)^{-1}(b \circ B + P_t A_L^T)^T$$

subject to

$$\hat{X}_0 = E[X_0] + \text{cov}(X_0, Y_0) \text{cov}(Y_0)^{\oplus} (Y_0 - E[Y_0])$$

$$P_0 = \text{cov}(X_0) - \text{cov}(X_0, Y_0) \text{cov}(Y_0)^{\oplus} \text{cov}(Y_0, X_0)^T$$

and where $b \circ B = b_L B_L^T + b_2 B_2^T$, $b \circ b = b_L b_L^T + b_2 b_2^T$.

3. Linear filtering beyond linear diffusions

4.15

Def. 4.5

W_t is a Wiener process in wide sense if $W_0 = 0$, $E[W_t] = 0$ and $E[W_t W_s] = \min(t, s)$, $s, t \geq 0$.

Example 4.6

The stochastic integral $w_t := \int_0^t \frac{X_s}{E[X_s^2]^{1/2}} dW_s$ with a square integrable, positive process $X_t \geq c > 0$ is a Wiener process in the wide sense:

$$E[W_t W_s] = \int_0^{\min(t,s)} E\left[\left(\frac{X_u}{E[X_u^2]^{1/2}}\right)^2\right] du = \min(t, s).$$

Since w_t has uncorrelated increments one can define the stochastic integral L^2

$$I_t(f) = \int_0^t f_s dw_s := \lim_{n \rightarrow \infty} \sum_{i=0}^{n-1} f_{t_i} (w_{t_{i+1}} - w_{t_i})$$

when $f \in L^2[0, T]$ is a deterministic function (Wiener integral)

Since the linear SDE

$$dX_t = a_t X_t dt + b_t dW_t$$

has the explicit solution

$$X_t = \exp\left(\int_0^t a_u du\right) \left(X_0 + \int_0^t \exp\left(-\int_0^s a_u du\right) b_s dW_s\right),$$

analogously one can define the process

$$X_t = \exp\left(\int_0^t a_u du\right) \left(X_0 + \int_0^t \exp\left(-\int_0^s a_u du\right) b_s dw_s\right)$$

to be the solution of

$$dX_t = a_t X_t dt + b_t dw_t$$

With these definitions one can check that the Kalman-Bucy filter equations generate the optimal linear estimate,

if we replace the Wiener process by the Wiener process in the wide sense.

Example 4.7

Consider the SDE system

$$dX_t = -X_t dt + dW_t$$

$$dY_t = X_t^3 dt + dV_t$$

with $X_0 \sim N(0, 1/2)$ and $Y_0 = 0$

By Itô formula

$$dX_t^3 = 3X_t^2 dX_t + 3X_t dt = -3X_t^3 dt + 3X_t dt + 3X_t^2 dW_t.$$

Define $Z_t = X_t^3$ and $w_t = \sqrt{2} \int_0^t X_s^2 dW_s - \frac{W_t}{\sqrt{2}}$

Then w_t is a Wiener process in the wide sense, let $s \leq t$:

$$\begin{aligned} E[w_t w_s] &= E\left[\left(\sqrt{2} \int_0^s X_u^2 dW_u - \frac{W_s}{\sqrt{2}}\right)^2\right] \\ &+ E\left[\left(\sqrt{2} \int_0^s X_u^2 dW_u - \frac{W_s}{\sqrt{2}}\right) E\left[\sqrt{2} \int_s^t X_u^2 dW_u - \frac{W_t - W_s}{\sqrt{2}} \middle| \mathcal{F}_s\right]\right] \\ &= 2 \int_0^s E[X_u^4] du + \frac{E[W_s^2]}{2} - 2 \int_0^s E[X_u^2] du \\ &= 2 \cdot \frac{3}{4} s + \frac{s}{2} - s = s \end{aligned}$$

where we have used that X is a stationary, zero-mean Gaussian process with $E[X_t X_s] = e^{-|t-s|}$. Since

$$X \text{ is Gaussian } E[X_u^4] = 3(E[X_u^2])^2 = \frac{3}{4}.$$

Analogously

$$\begin{aligned} E[w_t w_t] &= E\left[\left(\sqrt{2} \int_0^t X_u^2 dW_u - \frac{W_t}{\sqrt{2}}\right) W_t\right] = \sqrt{2} \int_0^t E[X_u^2] du \\ &- \frac{t}{\sqrt{2}} = \sqrt{2} \cdot t \cdot \frac{1}{2} - \frac{t}{\sqrt{2}} = 0. \end{aligned}$$

Hence (w_t, W_t, V_t) is a 3-dimensional Wiener process in the wide sense. Consider the linear system

$$dX_t = -X_t dt + dW_t$$

(4.12) $dZ_t = -3Z_t dt + 3X_t dt + \frac{3}{\sqrt{2}} dw_t + \frac{3}{2} dW_t$

$$dY_t = Z_t dt + dV_t$$

subject to $(X_0, Z_0) = (X_0, X_0^3)$ (i.e., $E[Z_0] = 0$ 4.17)

$$E[Z_0^2] = E[X_0^6] = \frac{15}{8}, \text{ etc. } \dots$$

the estimate $E[X_t | \mathcal{F}_t^X]$ can be obtained
by means of the Kalman-Bucy equations for
(4.12).