

Kalman-Bucy Filtering and Duality

We are interested in estimating (reconstructing) the state of a linear system driven by noise, using linear measurements of the state corrupted by additive measurement noise. In discrete-time this is known as Kalman filter(ing). In continuous-time we refer to the Kalman-Bucy filter. The derivation is based on a dual deterministic optimal control problem.

The subject is built on stochastic processes.

Let $(\Omega, \mathcal{A}, P)$ be a probability triple i.e., Ω = sample space, $\mathcal{A}$ = a σ -algebra of subsets (events) of Ω and $P: \mathcal{A} \rightarrow [0, 1]$ is a probability measure satisfying $P(\emptyset) = 0$, $P(\Omega) = 1$, $P(A \cup B) = P(A) + P(B)$ when $A \cap B = \emptyset$.

A random variable $X: \Omega \rightarrow \mathbb{R}$ is a $\mathcal{A}$ -measurable function, i.e. if $B \subset \mathbb{R}$ is a Borel subset then $X^{-1}(B) \in \mathcal{A}$. Then a stochastic process indexed by a (time) set T is a family

$$\{X_t: \Omega \rightarrow \mathbb{R} \mid t \in T\}$$

of random variables defined on $(\Omega, \mathcal{A}, P)$.

For us $T = [t_0, t_f]$ $t_0 < t_f \in \mathbb{R} \cup \{-\infty, \infty\}$.

Distribution function $F_{X_t}(x) = P\{\omega: X_t(\omega) \leq x\}$

Here
we fill
in details

in later
notes

and joint distribution

$$F_{X_{t_1}, X_{t_2}, \dots, X_{t_n}}(x_1, x_2, \dots, x_n)$$

$$= P \{ \omega : X_{t_i} \leq x_i \quad i=1, 2, 3, \dots, n, \}$$

n a positive integer

and $t_i \in [t_0, t_f] = T$

To fully define a stochastic process one needs to define all finite dimensional distributions of the above form fulfilling certain consistency conditions — postponed.

ABUSE

of NOTATION [We will often write $x(t)$ or $X(t)$ when we mean X_t (e.g. in Åström)]

Expectation

$$E(X) = \int_{\mathbb{R}} x dF(x)$$

defines a linear operator on random variables

integral suitably interpreted.

First we need to write stochastic models down. We write

$$dx = Ax dt + dv \quad (1a)$$

$$dy = Cx dt + de \quad (1b)$$

(1a) is a state evolution equation and (1b) is a (noisy) measurement equation. The noise (stochastic process) v drives the evolution. What does (1) mean?

Idea: Interpret it via integrals

$$\begin{aligned} \int_{t_0}^t dx &= x(t) - x(t_0) \\ &= \int_{t_0}^t A(\sigma) x(\sigma) d\sigma + \int_{t_0}^t dv \quad (2) \end{aligned}$$

$$= \int_{t_0}^t A(\sigma) x(\sigma) d\sigma + v(t) - v(t_0)$$

$t_0 < t$

Similarly for (1b). Then we collect

Meaning of (1)

$$\boxed{\begin{aligned} x(t) &= x(t_0) + \int_{t_0}^t A(\sigma) x(\sigma) d\sigma + v(t) - v(t_0) \\ y(t) &= y(t_0) + \int_{t_0}^t C(\sigma) x(\sigma) d\sigma + e(t) - e(t_0) \end{aligned}}$$

— 3 —

(require)
We interpret integrals to fulfill natural axioms (inverse relationship to derivative, linearity) when defined precisely (later)

In this setting $v(t)$ and $e(t)$ denote stochastic processes that generate the state process $x(t)$ and the measurement process $y(t)$ as in the integral formulas above.

Trouble: $x(t)$ is implicit and we need to solve the integral equation for $x(t)$. How?

First draw an analogy with deterministic setting

$$\dot{x} = Ax + u \quad (1a)$$

We can solve (1a) as follows.

Let $\Phi_A(t, t_0)$ be defined to be

the transition matrix (fundamental matrix) satisfying

$$\frac{d}{dt} \Phi_A(t, t_0) = A(t) \Phi_A(t, t_0)$$

$$\Phi_A(t_0, t_0) = \underline{1} = \text{identity matrix}$$

Let $\boxed{z(t)} \triangleq \Phi_A(t, t_0)^{-1} x(t)$ change of variables
 From the definition of Φ_A one can show

$$\Phi_A(t, s) = \Phi_A(s, t)^{-1} \quad (\text{inversion})$$

$$\begin{aligned} \Phi_A(t_2, t_1) \Phi_A(t_1, t_0) &= \Phi_A(t_2, t_0) \quad (\text{composition}) \\ &= \Phi_A(t_2, t_0) \end{aligned}$$

Then $\frac{dz}{dt} = \frac{d}{dt} (\Phi_A(t, t_0)^{-1}) x(t) + \Phi_A(t, t_0)^{-1} \frac{dx}{dt}$

$$= -\Phi_A(t, t_0)^{-1} \frac{d\Phi_A(t, t_0)}{dt} \Phi_A(t, t_0)^{-1} x(t)$$

$$+ \Phi_A(t, t_0)^{-1} (A(t)x(t) + u(t))$$

$$= -\cancel{\Phi_A(t, t_0)^{-1} A(t) \Phi_A(t, t_0)} \Phi_A(t, t_0)^{-1} x(t)$$

$$+ \cancel{\Phi_A(t, t_0)^{-1} A(t)} x(t)$$

$$+ \Phi_A(t, t_0)^{-1} u(t)$$

$$\Rightarrow \frac{dz}{dt} = \Phi_A(t, t_0)^{-1} u(t)$$

$$= \Phi_A(t_0, t) u(t)$$

$$\Rightarrow z(t) = z(t_0) + \int_{t_0}^t \Phi_A(t_0, \sigma) u(\sigma) d\sigma$$

$$= x(t_0) + \int_{t_0}^t \Phi_A(t_0, \sigma) u(\sigma) d\sigma$$

$$\Rightarrow x(t) = \Phi_A(t, t_0) \underline{\tilde{z}}(t)$$

$$= \Phi_A(t, t_0) x(t_0)$$

$$+ \Phi_A(t, t_0) \int_{t_0}^t \Phi_A(t_0, \sigma) u(\sigma) d\sigma$$

$$= \Phi_A(t, t_0) x(t_0)$$

$$+ \int_{t_0}^t \Phi_A(t, \sigma) u(\sigma) d\sigma \quad \left(\begin{array}{l} \text{using} \\ \text{composition} \end{array} \right)$$

This is Lagrange's variation of constants formula for $(\tilde{A})$, a deterministic equation

In the above calculation, everything works for integrals interpreted as Riemann integrals under the hypothesis that $u(\cdot)$ is piecewise continuous

For the STOCHASTIC setting (1a) we need to take a longer route since we do not know if it is legal to write $\frac{dx}{dt}$ as a process. So we

need to proceed via the integral equation interpretation as in (2), writing,

$$x(t) = x(t_0) + \int_{t_0}^t A(\sigma) x(\sigma) d\sigma + v(t) - v(t_0)$$

and using the same change of variables as before

$$\begin{aligned} z(t) &= \Phi_A^{-1}(t, t_0) x(t) \\ &= \Phi_A(t_0, t) x(t) \end{aligned}$$

as in the deterministic setting,

$$\begin{aligned}
Z(t) &= \Phi_A(t, t_0) \left[X(t_0) + \int_{t_0}^t A(\sigma) X(\sigma) d\sigma \right] \\
&\quad + V(t) - V(t_0) \\
&= \Phi_A(t, t_0) \left[Z(t_0) + \int_{t_0}^t A(\sigma) \Phi_A(\sigma, t_0) Z(\sigma) d\sigma \right. \\
&\quad \left. + V(t) - V(t_0) \right] \\
&= \Phi_A(t, t_0) \left[Z(t_0) + \int_{t_0}^t \frac{d}{d\sigma} \Phi_A(\sigma, t_0) Z(\sigma) d\sigma \right. \\
&\quad \left. + V(t) - V(t_0) \right]
\end{aligned}$$

Now we hypothesize that the notion of integral used in the above expression is also 'nice enough' to fulfill conditions for integration by parts to hold i.e.

$$\begin{aligned}
\int_a^b \frac{d}{d\sigma} \phi(\sigma) Z(\sigma) d\sigma &= \phi(\sigma) Z(\sigma) \Big|_a^b \\
&\quad - \int_a^b \phi(\sigma) dZ(\sigma)
\end{aligned}$$

where Z is a stochastic process but ϕ is a deterministic function. [This is checkable later when we define the integrals (more) precisely]

It follows that

$$\begin{aligned}
 Z(t) = & \Phi_A(t_0, t) \left[Z(t_0) + \Phi_A(t, t_0) Z(t_0) \right. \\
 & - \Phi_A(t_0, t_0) Z(t_0) \\
 & - \int_{t_0}^t \Phi_A(\sigma, t_0) dZ(\sigma) \\
 & \left. + V(t) - V(t_0) \right]
 \end{aligned}$$

$$\begin{aligned}
 = & Z(t) + \Phi_A(t_0, t) \left[\cancel{Z(t_0)} - \cancel{Z(t_0)} \right. \\
 & - \int_{t_0}^t \Phi_A(\sigma, t_0) dZ(\sigma) \\
 & \left. + V(t) - V(t_0) \right]
 \end{aligned}$$

$$\left\{ \begin{array}{l} \text{where we have used} \\ \Phi_A(t_0, t) \Phi_A(t, t_0) = \mathbb{1} \text{ and} \\ \Phi_A(t_0, t_0) = \mathbb{1} \end{array} \right.$$

$$\Rightarrow V(t) - V(t_0) = \int_{t_0}^t \Phi_A(\sigma, t_0) dZ(\sigma)$$

$$\Rightarrow dV(t) = \Phi_A(t, t_0) dZ(t) \quad (\text{integral property})$$

$$\Rightarrow dZ(t) = \Phi_A(t_0, t) dV(t)$$

$$\Rightarrow Z(t) = \underbrace{Z(t_0)}_{X(t_0) \leftarrow \text{recall}} + \int_{t_0}^t \Phi_A(t, \sigma) dV(\sigma)$$

$$\Rightarrow X(t) = \Phi_A(t, t_0) Z(t)$$

$$= \Phi_A(t, t_0) X(t_0)$$

$$+ \Phi_A(t, t_0) \int_{t_0}^t \Phi_A(t, \sigma) dV(\sigma)$$

$$= \Phi_A(t, t_0) X(t_0) + \int_{t_0}^t \Phi_A(t, \sigma) dV(\sigma)$$

We have arrived at a stochastic variation of constants formula to solve the stochastic differential equation (1) by interpreting it in the integral equation form (2). In the process we needed only certain natural properties of integrals (equivalently, assert certain axioms for integrals).