Kalman filter

- orthogonal projection:

Recall that for $x \in \mathbb{R}^n$,

if we let $x = x_1 e_1 + \cdots + x_n e_n$, where

$$e_i \cdot e_j = \delta_{ij} = \begin{cases} 1 & i=j \\ 0 & i \neq j \end{cases}$$

$$\Rightarrow \quad x \cdot y_i = x_1 y_1^i + \cdots + x_n y_n^i$$

$$y = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \qquad [y] = \text{span}\{y_1, \cdots, y_m\}$$

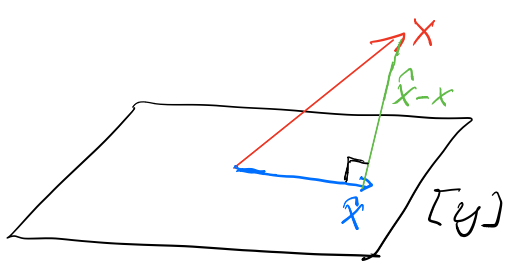
$$\Rightarrow \quad \hat{x} = k^* y \qquad m \in [y]$$

$$\Rightarrow \quad k^* = x \cdot y^T (y \cdot y^T)^{-1} \qquad \Rightarrow \quad m = a_1 y_1 + \cdots + a_m y_m$$

$$\Rightarrow \quad \hat{x} = \underbrace{x \cdot y^T (y \cdot y^T)^{-1}}_{\text{the optimal gain}} y \qquad = [y_1 \cdots y_m] \begin{bmatrix} a_1 \\ \vdots \\ a_m \end{bmatrix} = y^T a$$

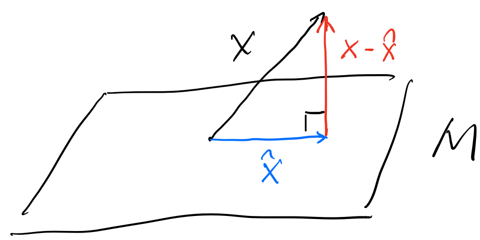
$$(\hat{x} - x) \cdot y^T a = (x \cdot y^T (y \cdot y^T)^{-1} y \cdot y^T a - x \cdot y^T a)$$
$$= 0.$$

$\|x - \hat{x}\| \to \min$



$$\Rightarrow \quad x - \hat{x} \text{ is orthogonal to } [y]\ !$$

Lemma: let M be a subspace of Hilbert space (inner product) H, and let $x \in H$, then $\exists$ a unique $\hat{x} \in M$, such that $\|x - m\|$ ($\|x-m\|^2 = (x-m)\cdot(x-m)$) $\forall\, m \in M$ is minimized. Furthermore, $(x - \hat{x})\cdot m = 0 \quad \forall\, m \in M$.



$\|x - \hat{x}\| \to \min$

$$\hat{x} = \underbrace{x\cdot y^T (y\cdot y^T)^{-1}}_{\text{the optimal gain}} y$$

We denote $\boxed{\hat{x} = E^{M} x}$

Examples: 1. $L_2[a,b]$ all square integrable functions on $[a,b]$

$\forall\, f(t), g(t) \in L_2[a,b]$

$$f\cdot g = \int_a^b f g \, dt$$

$y_i(t)$: $[a,b] = [-\pi, \pi]$

$\sin kt$, $\cos kt$, $k = 0, 1, \cdots, N$

$M = \text{span}\{\sin kt, \cos kt\}$ $k = 0, \cdots, N$

$E^M f(t)$ — Fourier series

2. space of random variables:

given two random variables x, y

$x \cdot y = E\{xy\}$

Properties of orthogonal projection:

1. $E^M(\alpha x_1 + \beta x_2) = \alpha E^M x_1 + \beta E^M x_2$

2. $E^M Ax = AE^M x$

3. If $M \perp N$ (i.e. $m \cdot n = 0$ $\forall m \in M, n \in N$)

$$E^{M \oplus N} x = E^M x + E^N x$$

Kalman filter:

— Discrete systems:

$x \in \mathbb{R}^n$

$$x(t+1) = A(t)\,x(t) + B(t)\,v(t) + (G(t)\,u(t))$$

$$y(t) = C(t)\,x(t) + D(t)\,w(t)$$

where $v(t)$, $w(t)$ are stochastic processes (noises).

$v(t)$ and $w(t)$ are uncorrelated, i.e.

$$E\{v(t)\,w^T(s)\} = 0 \quad \forall\, t, s.$$

and we also assume

$$E\{w(t)\,w^T(s)\} = R\,\delta(t-s)$$

$$E\{v(t)\,v^T(s)\} = Q\,\delta(t-s)$$

where $\delta(t-s) = \begin{cases} 1 & t = s \\ 0 & t \neq s \end{cases}$

$R > 0, \quad Q \geq 0.$

Problem: Find the "best" estimator $\hat{x}(t)$ of $x(t)$ based on measurement $y(0), \cdots, y(t)$.

Let us assume $y(t) = \begin{bmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{bmatrix}$

More precisely $\hat{x}(t)$ based on

$y_1(0), \cdots y_m(0), \cdots\cdots, y_1(t), \cdots, y_m(t)$.

Denote $H_t(y) = \text{span}\{y_1(0), \cdots, y_m(0), \cdots, y_1(t), \cdots, y_m(t)\}$

$\Rightarrow \quad H_{t-1}(y) = \text{span}\{y_1(0), \cdots, y_m(t-1)\}$

Now we want find $\hat{x}$, s.t.

$\| x_i(t) - \hat{x}_i(t) \| \to \min \quad (\| x_i(t) - \hat{x}_i(t) \|^2$

$= E\{(x_i - \hat{x})^2\})$

where $\quad x(t) = \begin{bmatrix} x_1(t) \\ \vdots \\ x_n(t) \end{bmatrix}$

(different from $\| x(t) - \hat{x}(t) \| \to \min$)

$$\Rightarrow \quad \hat{x}_i(t) = \hat{E}^{H_{t-1}(y)} x_i(t)$$

$$\Rightarrow \quad \hat{x}(t) = \hat{E}^{H_{t-1}(y)} x(t)$$

Given $\quad H_{t-1}(y) = \text{Span}\{y_1(0), \cdots, y_m(t-1)\}$

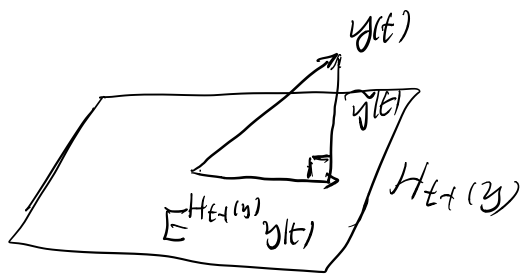
Now comes $\quad y(t) = \begin{bmatrix} y_1(t) \\ \vdots \\ y_m(t) \end{bmatrix}$

Denote $[y] = \text{span}\{y_1(t), \cdots, y_m(t)\}$

We do $\overrightarrow{E}^{\mathcal{H}_{t-1}(y)} y(t)$

$$\tilde{y}(t) := y(t) - E^{H_{t-1}(y)} y(t)$$

Denote $[\tilde{y}(t)] = \mathrm{span}\{\tilde{y}_1(t), \cdots, \tilde{y}_m(t)\}$



$$\Rightarrow \mathcal{H}_t(y) = H_{t-1}(y) \oplus [\tilde{y}(t)]$$

We need to assure

$E\{\tilde{y}(t)\tilde{y}^T(t)\}$ positive definite

$\tilde{y}(t)$ – innovation

Now let $\hat{x}_t(t) = E^{H_t(y)} x(t)$

$$= E^{H_{t-1}(y) \oplus [\tilde{y}(t)]} x(t)$$

$$= E^{H_{t-1}(y)} x(t) + E^{[\tilde{y}(t)]} x(t)$$

$$= \hat{x}(t) + K \tilde{y}(t)$$

$$x(t+1) = A(t)x(t) + B(t)v(t)$$
$$y(t) = C(t)x(t) + D(t)w(t)$$

$$\Rightarrow \quad \hat{x}(t+1) := E^{H_t(y)} x(t+1)$$

$$\Rightarrow \quad \hat{x}(t+1) = E^{H_t(y)} (Ax(t) + B(t)v(t))$$

$$= A\hat{x}_t(t) \qquad (v(t) \perp H_t(y))$$

$$= A\hat{x}(t) + AK\tilde{y}(t)$$

$$E^{[\tilde{y}(t)]} x(t) = \ ?$$

$$E^{[\tilde{y}(t)]} x(t) = x(t) \cdot \tilde{y}^T (\tilde{y} \cdot \tilde{y}^T)^{-1} \tilde{y}(t)$$

$$= E\{x(t)\tilde{y}^T(t)\} (E\{\tilde{y}(t)\tilde{y}^T(t)\}) \tilde{y}(t)$$

$$\tilde{y}(t) = y(t) - E^{H_{t-1}(y)} y(t)$$

$$= y(t) - E^{H_{t-1}(y)} (C(t)x(t) + D(t)w(t))$$

$$= y(t) - C(t)\hat{x}(t) \qquad (w(t) \perp H_{t-1}(y))$$

$$= C(t)x(t) + D(t)w(t) - C(t)\hat{x}(t)$$

$$= C(t)\tilde{x}(t) + D(t)w(t)$$

Where $\tilde{x}(t) = x(t) - \hat{x}(t)$

Denote $P(t) = E\{\tilde{x}(t)\tilde{x}^T(t)\}$

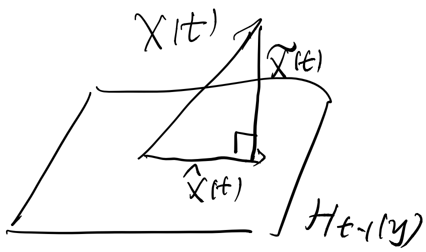
$$E\{x(t)\tilde{y}^T(t)\} = E\{x(t)(C(t)\tilde{x}(t) + D(t)w(t))^T\}$$

$$= E\{x(t)\tilde{x}^T(t)\}C^T(t) + E\{x(t)w^T(t)\}D^T(t)$$

$$= E\{(x(t) - \hat{x}(t) + \hat{x}(t))\tilde{x}^T(t)\}C^T(t)$$

$$= E\{\tilde{x}(t)\tilde{x}^T(t)\}C^T(t) + E\{\hat{x}(t)\tilde{x}^T(t)\}C^T(t)$$

$$= P(t)C^T(t)$$



$$E\{\tilde{y}(t)\tilde{y}^T(t)\} = E\{(C(t)\tilde{x}(t) + D(t)w(t))(\quad)^T\}$$

$$= E\{C(t)\tilde{x}(t)\tilde{x}^T C^T(t)\} + E\{C(t)\tilde{x}(t)w^T(t)D(t)\}$$

$$+ E\{D(t)w(t)\tilde{x}^T(t)C^T(t)\} + E\{D(t)w(t)w^T(t)D^T(t)\}$$

$$= C(t)\,P(t)\,C^T(t) + D(t)\,R\,D^T(t)$$

$$\Rightarrow \quad K^* = P(t)\,C^T(t)\left(C(t)P(t)C^T(t) + D(t)RD^T(t)\right)^{-1}$$

– Kalman gain

$$\hat{x}_t(t) = \hat{x}(t) - K\tilde{y}(t)$$

– measurement update

$$= \hat{x}(t) - K^*\left(y(t) - C(t)\hat{x}(t)\right)$$

$$\hat{x}(t+1) = A\hat{x}(t) - AK^*\left(y(t) - C(t)\hat{x}(t)\right)$$

time update