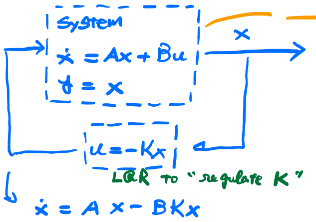
Motivation for Full-state estimation (from Steve Brunton)

Recall

$\dot{x} = Ax + Bu$ $x \in \mathbb{R}^n$, $u \in \mathbb{R}^q$



$\dot{x} = Ax - BKx$

$\dot{x} = (A - BK)x$

(placing eig-values)

I don't necessarily hv all states in real life

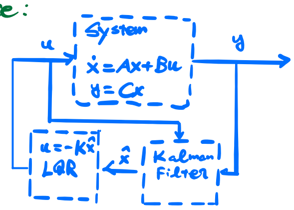
$\dot{x} = Ax + Bu$ (controllability)   ctrb (A, B)

$y = Cx$ (observability)   obsv (A, C)

Main Question here:

Can I estimate any state $\underline{x}$ from measurement $y(t)$

hence:



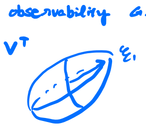
Observability

- Duality exists between AB, AC
- observability matrix

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix} \qquad \mathcal{C} = \begin{bmatrix} B & AB & A^2B & \cdots & A^{n-1}B \end{bmatrix}$$

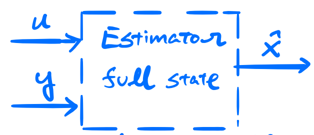
1. observable if >> rank (obsv (A, C)) = n
2. Can estimate x from y
3. >> $[U, \Sigma, V] = \text{svd}(\mathcal{O})$

observability Gramian



In some direction, we hv higher signal to noise

Full State Estimation



dyn. system (linear)

$$\frac{d}{dt}\hat{x} = A\hat{x} + Bu + K_f(y - \hat{y})$$

$$\hat{y} = C\hat{x}$$

($K_f(y - \hat{y})$: update; K_f: filter)

$$\frac{d}{dt}\hat{x} = A\hat{x} + Bu + K_f y - K_f C\hat{x}$$

$$= (A - K_f C)\hat{x} + \begin{bmatrix} B & K_f \end{bmatrix}\begin{bmatrix} u \\ y \end{bmatrix}$$

↓ pick K_f to place the eigen values to hv optimal choice

Error $\varepsilon = x - \hat{x}$

$$\frac{d}{dt}\varepsilon = \frac{d}{dt}x - \frac{d}{dt}\hat{x}$$

$$= Ax + Bu - A\hat{x} + K_f C\hat{x} - K_f y - Bu$$

$$= Ax - A\hat{x} + K_f C\hat{x} - K_f y \quad (y = Cx)$$

$$= A(x - \hat{x}) + K_f C(\hat{x} - x)$$

$$= A(x - \hat{x}) - K_f C(x - \hat{x})$$

$$= (A - K_f C)\varepsilon$$

if observable, then place eigs by choosing K_f: so that error converge eventually

← motivation of Kalman filter

Kalman filter

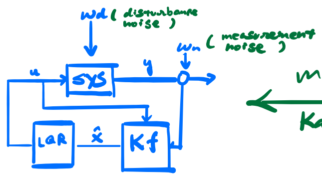
- W_d – Gaussian
  V_d Variance
  W_n – Gaussian
  V_n Variance
- recall
  $\dot{\varepsilon} = (A - K_f C)\varepsilon$
  $\varepsilon = x - \hat{x}$
- cost function
  $J = \mathbb{E}\left((x-\hat{x})^T(x-\hat{x})\right)$
- >>> K_f = lqe (A, C, vd, vn)

← Kalman filter

real system:

$\dot{x} = Ax + Bu + w_d$

$y = Cx + w_n$



↓

get the best K_f to place poles (eigs) based on w_d & w_n

more detailed derivation

Kalman Filtering

$\theta \in R^n$ random vector

$Y_1^k = [y(1), y(2), \ldots, y(k-1), y(k)]$ observation

$y(i)$ & θ : independent

joint pdf of θ & Y_1^k : $p(\theta, Y_1^k)$

conditional pdf of $\theta | Y_1^k$: $p(\theta | Y_1^k)$

pdf of Y_1^k : $p(Y_1^k)$

The Estimation Problem

given $y(1), y(2), \ldots, y(k)$: evaluate θ

(3.3) $\hat{\theta}(k) = f[y(i), i = 1 \ldots, k]$ $\xrightarrow[\text{a criteria}]{\text{optimize}}$
- mean square error
- maximum a posterior

Consider

mean-square error for General Estimation of Random Parameters

(3.4) cost function — $\theta - \hat{\theta}_k$ → this is a distribution

$$J[\tilde{\theta}(k)] = E[\tilde{\theta}(k)^T \tilde{\theta}(k)]$$

$$\tilde{\theta}(k) \triangleq \theta - \hat{\theta}(k)$$ — ① we want to minimize error

(3.6) $\hat{\theta}(k) = \text{argmin } E[(\theta - \hat{\theta}(k))^T (\theta - \hat{\theta}(k))]$

FACT $\underset{\hat{\theta}(k)}{\text{minimize}} \; E[\tilde{\theta}(k)^T \tilde{\theta}(k)] \Leftrightarrow \underset{\hat{\theta}(k)}{\text{minimize}} \; E[\tilde{\theta}(k)^T \tilde{\theta}(k) | Y_1^k]$

② (as $\tilde{\theta}(k) = \theta - \hat{\theta}(k)$, & $\hat{\theta}(k)$ is a function of y. Hence, $E[\tilde{\theta}(k)]$ is related to the joint pdf of θ & y

proof △ $E[\tilde{\theta}(k)^T \tilde{\theta}(k)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{\theta}(k)^T \tilde{\theta}(k) \, p(\theta, Y_1^k) \, d\theta \, dY_1^k$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{\theta}(k)^T \tilde{\theta}(k) \, p(\theta, Y_1^k) \underbrace{d\theta_1 d\theta_2 \ldots d\theta_n}_{\infty} \underbrace{dy_1 dy_2 \ldots dy_k}_{\infty}$$

△ Bayes Law: $p(\theta, Y_1^k) = p(\theta | Y_1^k) \, p(Y_1^k)$

△ $E[\tilde{\theta}(k)^T \tilde{\theta}(k)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{\theta}(k)^T \tilde{\theta}(k) \, p(\theta, Y_1^k) \, d\theta_1 d\theta_2 \ldots d\theta_2 \, dy_1 dy_2 \ldots dy_k$

$$= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} \tilde{\theta}(k)^T \tilde{\theta}(k) \, p(\theta | Y_1^k) \, d\theta_1 d\theta_2 \ldots d\theta_2 \right] p(Y_1^k) \, dy_1 dy_2 \ldots dy_k$$

$$= \int_{-\infty}^{\infty} \underbrace{E[\tilde{\theta}(k)^T \tilde{\theta}(k) | Y_1^k]}_{\text{minimize this}} \, p(Y_1^k) \, dY_1^k$$

conclusion #1 as in (3.6)

$$\hat{\theta}(k) = \text{argmin } E[\tilde{\theta}(k)^T \tilde{\theta}(k) | Y_1^k]$$

← so we minimize this, which we can then start to cook up the relations between $\hat{\theta}$ & $y(i)$

$\ell(\theta, \hat{\theta}_k)$

$\iint (\theta - \hat{\theta}_k)(\theta - \hat{\theta}_k)$ $\iint$

$P(x|y) = \dfrac{P(x,y)}{P(y)}$

No

Date

FACT (3.9) for the minimization problem

$$\hat{\theta}(k) = \text{argmin } E[\tilde{\theta}(k)^T\tilde{\theta}(k) | Y_1^k]$$

we have

$$\hat{\theta}(k) = E[\theta | Y_1^k]$$

$E[\hat{\theta}(k)^T\hat{\theta}(k) | Y_1^k]$
$= \int_{-\infty}^{\infty} \hat{\theta}(k)^T\hat{\theta}(k) P(Y_1^k)\, dY_1^k$
$= \int_{-\infty}^{\infty} P(Y_1^k)\, dY_1^k \cdot \hat{\theta}(k)^T\hat{\theta}(k)$
$= 1 \cdot \hat{\theta}(k)^T\hat{\theta}(k)$
$= \hat{\theta}(k)^T\hat{\theta}(k)$

proof for 3.5 & 3.9

$$J = E[\tilde{\theta}(k)^T\tilde{\theta}(k) | Y_1^k]$$
$$= E[(\theta - \hat{\theta}(k))^T(\theta - \hat{\theta}(k)) | Y_1^k]$$
$$\Rightarrow J = E[\theta^T\theta - \theta^T\hat{\theta}(k) - \hat{\theta}(k)^T\theta + \hat{\theta}(k)^T\hat{\theta}(k) | Y_1^k]$$
$$= E[\theta^T\theta | Y_1^k] - E[\theta^T | Y_1^k]\hat{\theta}(k) - \hat{\theta}(k)^T E[\theta | Y_1^k] + E[\hat{\theta}(k)^T\hat{\theta}(k) | Y_1^k]$$
$$= \underbrace{E[\theta^T\theta | Y_1^k]}_{①} - E[\theta^T|Y_1^k]\hat{\theta}(k) - \hat{\theta}(k)^T E[\theta|Y_1^k] + E[\hat{\theta}(k)^T\hat{\theta}(k)|Y_1^k] + E[\theta^T|Y_1^k]E[\theta|Y_1^k] \underbrace{- E[\theta^T|Y_1^k]E[\theta|Y_1^k]}_{②}$$

Rk

$$= E[\theta^T\theta | Y_1^k] - E[\theta^T|Y_1^k]E[\theta|Y_1^k] \quad ②$$
$$+ E[\theta^T|Y_1^k]E[\theta|Y_1^k] - E[\theta^T|Y_1^k]\hat{\theta}(k) - \hat{\theta}(k)^T E[\theta|Y_1^k] + \hat{\theta}(k)^T\hat{\theta}(k)$$
$$= E[\theta^T\theta | Y_1^k] - E[\theta^T|Y_1^k]E[\theta|Y_1^k] + [\hat{\theta}(k) - E[\theta|Y_1^k]]^T[\hat{\theta}(k) - E[\theta|Y_1^k]]$$

∴ minimization problem becomes:

$$\underset{\hat{\theta}(k)}{\text{minimize}} \; [\hat{\theta}(k) - E[\theta|Y_1^k]]^T[\hat{\theta}(k) - E[\theta|Y_1^k]] + E[\theta^T\theta|Y_1^k] - E[\theta^T|Y_1^k]E[\theta|Y_1^k]$$

$$\therefore \; \hat{\theta}(k) = E[\theta|Y_1^k]$$

③ with some formulation, the optimal $\hat{\theta}(k)$ is this.

conclusion #2 $\hat{\theta}(k) = E[\theta | Y_1^k]$ (3.2.1)

Gaussian Particularization.

FACT θ, Y^k are jointly Gaussian

$$E[\theta | Y_1^k] = E[\theta] + R_{\theta Y_1^k} R_{Y_1^k}^{-1} [Y_1^k - E[Y_1^k]]$$

$R_{\theta Y_1^k} \rightarrow E[(\theta - E(\theta))(Y_1^k - E(Y_1^k)^T]$

$R_{Y_1^k} \rightarrow E[(Y_1^k - E(Y_1^k))(Y_1^k - E(Y_1^k)^T]$

$Z := X_a - CX_b$

proof
$$E[X_a|X_b] = E[Z + CX_b | X_b]$$
remark: we want Z to be independent → X_b so that $E[Z|X_b] = E[Z]$.
$$= E[Z|X_b] + E[CX_b|X_b]$$
$$= E[Z|X_b] + CX_b$$
$$= E[X_a - CX_b] + CX_b$$
$$= E[X_a] + C(X_b - E[X_b]) = E[X_a] + \Sigma_{ab}\Sigma_{bb}^{-1}(X_b - E[X_b])$$

let $Z = X_a - CX_b$ to be independent/uncorrelated to X_b:

$$0 = cov(Z, X_b) = cov(X_a - CX_b, X_b)$$
$$= cov(X_a, X_b) - C\, cov(X_b, X_b)$$
$$= \Sigma_{ab} - C\Sigma_{bb} \quad \therefore C = \Sigma_{ab}\Sigma_{bb}^{-1}$$

④ prior to this, ③ applies to any distribution. for Gaussian, it's:

$$\therefore E[\theta | Y_1^k]$$
$$= E[\theta] + \Sigma_{\theta Y_1^k}\Sigma_{Y_1^k Y_1^k}^{-1}(Y_1^k - E[Y_1^k])$$

$E[x|z]$ $y = x - Cz$
$= E[x - Cz + Cz | z]$

C is s.t.h that let $x - Cz \perp z$

$= E[y + C\mathfrak{z} \mid \mathfrak{z}]$

$= E[y \mid \mathfrak{z}] + C\mathfrak{z}$

$\therefore E[x \mid \mathfrak{z}] = E[y \mid \mathfrak{z}] + C\mathfrak{z}$

$= E[x] + \Sigma_{x,\mathfrak{z}} \Sigma_{\mathfrak{z}\mathfrak{z}}^{-1} (\mathfrak{z} - E[\mathfrak{z}])$

$\mathrm{cov}(y, \mathfrak{z}) = 0$

$= \mathrm{cov}(x - C\mathfrak{z}, \mathfrak{z})$

$= \mathrm{cov}(x, \mathfrak{z}) - C\,\mathrm{cov}(\mathfrak{z}, \mathfrak{z})$

$= \mathrm{cov}(x, \mathfrak{z}) - C\,\mathrm{cov}(\mathfrak{z}, \mathfrak{z})$

$\therefore C = \Sigma_{x,\mathfrak{z}} \Sigma_{\mathfrak{z}\mathfrak{z}}^{-1}$

Kalman Filtering $E[\hat{x} \mid \check{x}, \mathfrak{z}]$

⑤ A linear dynamics (time-varying) assumption : LINEAR QUADRATIC GAUSSIAN

FACT consider general filtering problem:

$$x(k+1) = f(x(k), u(k), w(k)) \quad (4.1)$$
$$y(k) = h(x(k), v(k)) \quad (4.2)$$

} obtains minimum mean-square state error estimate

↓ particularizes to

- $x_{k+1} = A_k x_k + B_k u_k + G w_k$ ($x_k \in \mathbb{R}^n$, $u_k \in \mathbb{R}^m$, $w_k \in \mathbb{R}^n$)

  $y_k = C_k x_k + v_k$ ($y_k \in \mathbb{R}^n$, $x_k \in \mathbb{R}^n$, $v_k \in \mathbb{R}^n$)

  • } w_k, v_k : white/zero mean // Gaussian noise

  $E[w_k] = E[v_k] = 0$

- $E[x_0] = \bar{x}_0$ initial state
- $E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = \Sigma_0$ covariance matrix

• joint covariance matrix

$$E\left[\begin{pmatrix} w_k \\ v_k \end{pmatrix} (w_k^T \; v_k^T)\right] = \begin{bmatrix} Q_k & 0 \\ 0 & R_k \end{bmatrix}$$

$w_k \leftrightarrow v_k$ independent.

• u_k is deterministic

FACT • $p(x_k \mid Y_1^k, U_0^{k-1}) \sim \mathcal{N}(\hat{x}_{k|k}, P_{k|k})$

where

$$\hat{x}_{k|k} = E[x_k \mid Y_1^k, U_0^{k-1}]$$
$$P_{k|k} = E[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T \mid Y_1^k, U_0^{k-1}]$$

• Kalman filter only propagates 1^{st} & 2^{nd} moment of the distribution

Kalman Filter dynamics

$P(x_0)$ —prediction→ $P(x_1 \mid U_0^0)$ —filtering→ $P(x_1 \mid Y_1^1, U_0^0)$ —prediction→ $P(x_2 \mid Y_1^1, U_0^1)$ —filtering→ $P(x_2 \mid Y_1^2, U_0^1)$ ...

... $P(x_k \mid Y_1^k, U_0^{k-1})$ —prediction→ $P(x_{k+1} \mid Y_1^k, U_0^k)$ —filtering→ $P(x_{k+1} \mid Y_1^{k+1}, U_0^k)$ —prediction→ $P(x_{k+2} \mid Y_1^k, U_0^{k+1})$ —filtering→ $P(x_{k+2} \mid Y_1^{k+2}, U_0^{k+1})$

FACT estimates & variance of Kalman filter:

(4.8) • $p(x_k \mid Y_1^k, U_0^{k-1}) \sim \mathcal{N}(\hat{x}_{k|k}, P_{k|k})$

(4.9) $p(x_{k+1} \mid Y_1^k, U_0^k) \sim \mathcal{N}(\hat{x}_{k+1|k}, P_{k+1|k})$

(4.12) (4.13)

$P_{k|k} = E[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T \mid Y_1^k, U_0^{k-1}]$

$P_{k+1|k} = E[(x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^T \mid Y_1^k, U_0^k]$

⑥ propagation of 1^{st} & 2^{nd} moment of the distribution

$\hat{x}_{k|k} = E[x_k \mid Y_1^k, U_0^{k-1}]$

$\hat{x}_{k+1|k} = E[x_{k+1} \mid Y_1^k, U_0^k]$

(4.10) (4.11)

No.
Date

Prediction ⑦ based on the linear dynamics : prediction propagation

FACT State Prediction

$P(X_{k+1} \mid Y_1^k, U_0^k)$

- recall $X_{k+1} = A_k X_k + B_k u_k + G w_k$

(4.14) $E[X_{k+1} \mid Y_1^k, U_0^k] = A_k E[X_k \mid Y_1^k, U_0^{k-1}] + B_k E[u_k \mid Y_1^k, U_0^k] + G E[w_k \mid Y_1^k, U_0^k]$

- recall $P(X_k \mid Y_1^k, U_0^{k-1}) \sim N(\hat{x}_{k|k}, P_{k|k})$

$P(X_{k+1} \mid Y_1^k, U_0^k) \sim N(\hat{x}_{k+1|k}, P_{k+1|k})$

& w_k & Y_1^k are independent / $cov(w_k, Y_1^k) = 0$, and $\bar{w}_k = 0$

4.14 ⇒ $E[X_{k+1} \mid Y_1^k, U_0^k] = A_k E[X_k \mid Y_1^k, U_0^{k-1}] + B_k E[u_k \mid Y_1^k, U_0^k] + G E[w_k \mid Y_1^k, U_0^k]$

$\hat{x}_{k+1|k} = A_k \hat{x}_{k|k} + B_k u_k + G \bar{w}_k$ (→ 0)

(4.15) $\hat{x}_{k+1|k} = A_k \hat{x}_{k|k} + B_k u_k$

prior ← $\bar{X}_{k+1}$

FACT Prediction Error

let $\tilde{x}_{k+1|k} = x_{k+1} - \hat{x}_{k+1|k}$

$= A_k x_k + B_k u_k + G_k w_k - (A_k \hat{x}_{k|k} + B_k u_k)$

$= A_k (x_k - \hat{x}_{k|k}) + G_k w_k$

$= A_k \tilde{x}_{k|k} + G_k w_k$

$\tilde{x}_{k+1|k} = A_k \tilde{x}_{k|k} + G_k w_k$

prior   posterior

$\tilde{x}_{k|k} \triangleq x_k - \hat{x}_{k|k}$

FACT Covariance Matrix

$E[[x_{k+1} - \hat{x}_{k+1|k}][x_{k+1} - \hat{x}_{k+1|k}]^T \mid Y_1^k, U_0^k]$

$= E[\tilde{x}_{k+1|k} \tilde{x}_{k+1|k}^T \mid Y_1^k, U_0^k]$

$= A_k E[\tilde{x}_{k|k} \tilde{x}_{k|k}^T \mid Y_1^k, U_0^{k-1}] A_k^T + G_k w_k w_k^T G_k^T$

$= A_k E[\tilde{x}_{k|k} \tilde{x}_{k|k}^T \mid Y_1^k, U_0^{k-1}] A_k^T + G_k Q G_k^T$

- recall $P_{k|k} = E[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T \mid Y_1^k, U_0^{k-1}]$

$P_{k+1|k} = E[(x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^T \mid Y_1^k, U_0^k]$

∴ $P_{k+1|k} = A_k P_{k|k} A_k^T + G_k Q G_k^T$

No. Date

⑧ measurement prediction propagation

Measurement

FACT recall $y(k) = h(x(k), v(k)) \Rightarrow y_k = C_k x_k + v_k$

$$p(y_{k+1} | Y_1^k, U_0^k) = p(C_{k+1}x_{k+1} + v_{k+1} | Y_1^k, U_0^k) \quad (4.21)$$

(Gaussian pdf)

FACT $$\hat{y}_{k+1|k} = E[y_{k+1} | Y_1^k, U_0^k] = C_{k+1}\hat{x}_{k+1|k} \quad (4.22)$$

(predicted measurement)

also $\tilde{y}_{k+1|k} \triangleq y_{k+1} - \hat{y}_{k+1|k}$ (measurement prediction error)

$$= C_{k+1}x_{k+1} + v_{k+1} - C_{k+1}\hat{x}_{k+1|k}$$

$$= C_{k+1}(x_{k+1} - \hat{x}_{k+1|k}) + v_{k+1}$$

$$= C_{k+1}\tilde{x}_{k+1|k} + v_{k+1}$$

FACT ∴ covariance matrix of $\hat{y}$

$$P_{\hat{y}} = E[\tilde{y}_{k+1|k}\tilde{y}_{k+1|k}^T | Y_1^k, U_0^k] = E[C_{k+1}\tilde{x}_{k+1|k}\tilde{x}_{k+1|k}^T C_{k+1}^T + v_{k+1}v_{k+1}^T | Y_1^k, U_0^k]$$

$$= C_{k+1}E[\tilde{x}_{k+1|k}\tilde{x}_{k+1|k}^T | Y_1^k, U_0^k]C_{k+1}^T + R_k$$

$$= C_{k+1}P_{k+1|k}C_{k+1}^T + R_k$$

Correction

FACT $E[(\theta - \hat{\theta}_k) f^T(Y_1^k)] = 0$ i.e., estimation error is orthogonal to function of Y_1^k

lemma 1

proof $E[x g(y)] = E[E(x|y) g(y)]$

lemma 1 proof

$E[E(x|y) g(y)]$

$= \sum_y E[x | Y=y] \, dP(y) \sum_y g(y) \, dP(g(y))$

lemma 2
$E[E[x|y]] = E[x]$

$= E[E(x|y)]$ → our variable which takes on 「value $E[x|Y=y]$」 when 「$Y=y$」

$= \sum_y E[x | Y=y] dP(Y=y)$

$= \sum_y \sum_x x \, d\frac{P(x \cap Y=y)}{P(Y=y)} \cdot dP(Y=y)$

$= \sum_y \sum_x x \, dP(x \cap Y=y)$ (joint distribution ⟺ two distributions' intersection)

$= \sum_x x \, dP(x)$

$= E[x]$

$= E[x] \sum_y g(y) dP(g(y))$

$= \sum_x \sum_y x \, g(y) dP(x) dP(g(y))$

$= \sum_x \sum_y x \, g(y) dP(x) dP(y)$

$= E[x \, g(y)]$

No

Date

proof (cont'd)

$E[x g(y)] = E[E(x|y) g(y)]$

$\therefore E[(\theta - \hat{\theta}_k) f^T(Y_1^k)]$

$= E[E[\theta - \hat{\theta}_k | Y_1^k] f^T(Y_1^k)]$

lemma 3 $E[\theta - \hat{\theta}_k | Y_1^k] = 0$

lemma 3 proof

$E[\theta - \hat{\theta}_k | Y_1^k]$

$= E[\theta | Y_1^k] - E[\hat{\theta}_k | Y_1^k]$

$= E[\theta | Y_1^k] - \hat{\theta}_k$

lemma 4 $E[\hat{\theta}_k | Y_1^k] = \hat{\theta}_k$

lemma 4 proof

$\hat{\theta}_k = f[y(i), i = 1, \dots, k]$

$\Rightarrow E[\hat{\theta}_k | Y_1^k]$

$= E[g(Y_1^k) | Y_1^k]$

$= \sum_{Y_1^k} g(Y_1^k) \cdot dP(Y_1^k)$

$= g(Y_1^k) = \hat{\theta}_k$

lemma 5 $\hat{\theta}_k = E[\theta | Y_1^k]$ has optimal estimation

$\therefore E[\theta - \hat{\theta}_k | Y_1^k] = 0$

$\therefore E[E[\theta - \hat{\theta}_k | Y_1^k] f^T(Y_1^k)]$

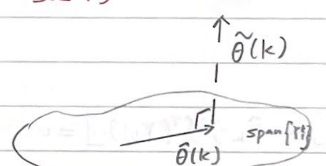
$= E[0 \; f^T(Y_1^k)]$

$= 0$ ☆ (from corollary 3.2.1)

e.g. $\hat{\theta}(k) = f(Y_1^k)$

$\therefore E[(\theta - \hat{\theta}_k)\hat{\theta}_k^T] = 0$

[est. error] [function of y]



FACT

$E[x_{k+1} | Y_1^{k+1}, U_0^k] = E[x_{k+1} | Y_1^k, \tilde{y}_{k+1|k}, U_0^k]$

FACT $\hat{x}_{k+1|k+1}$ derivation

$E[\hat{x}_k | Z_1^k, U_1^k]$

$= E[\hat{x} | \tilde{z}_k] = E[E[x_k | Z_1^{k-1}, U_1^k] | \tilde{z}_k]$

prior

proof

$\hat{x}_{k+1|k+1} = \sum_x x_{k+1} dp(x_{k+1} | Y_1^{k+1}, U_0^k)$

$= E[x_{k+1} | Y_1^k, \tilde{y}_{k+1|k}, U_0^k]$

lemma 1 Y_1^k, $\tilde{y}_{k+1|k}$ are independent, from corollary 3.2.1

$\therefore \hat{x}_{k+1|k+1} = E[z | \tilde{y}_{k+1|k}]$ where $z = E[x_{k+1} | Y_1^k, U_0^k]$

(from Gaussian Particularization

$E[x_a | x_b] = E[x_a] + \Sigma_{ab} \Sigma_{bb}^{-1} (x_b - E[x_b])$)

here $x_a = z = E[x_{k+1} | Y_1^k, U_0^k]$

$x_b = \tilde{y}_{k+1|k} = y_{k+1} - \hat{y}_{k+1|k}$

④ innovation / correction: try to get "optimal estimation" given new $y_{k+1|k}$ & model prediction

recall $E[x|z] = E[x] + \Sigma_{xz} \Sigma_{zz}^{-1} (z - E[z])$

$\hat{x}_k = E[x_k \mid Z_1^{k-1}, U_1^k, \tilde{z}_k]$

$= E[\tilde{x}] + \Sigma_{\tilde{x}\tilde{z}} \Sigma_{\tilde{z}\tilde{z}}^{-1} (\tilde{z} - E[\tilde{z}])$

No.

Date

Question why not

△ $\hat{x}_k = E[x_k \mid Z_1^{k-1}, U_1^k, z_k]$?

△ is it becos ($z_k \perp Z_1^{k-1}$) false ?

proof (cont'd)

$$\hat{x}_{k+1|k+1} = E[E[x_{k+1} \mid Y^k, U_0^k]] + E[\tilde{x}_{k+1} \tilde{y}_{k+1|k}^T] \cdot E[\tilde{y}_{k+1|k} \tilde{y}_{k+1|k}^T]^{-1} \cdot (\tilde{y}_{k+1|k} - E[\tilde{y}_{k+1|k}])$$

($E[x_a]$, Σ_{ab}, Σ_{bb}^{-1}, $x_b - E[x_b]$)

$$\Rightarrow \hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + E[(x_{k+1} - \hat{x}_{k+1|k})[C_{k+1}(x_{k+1} - \hat{x}_{k+1|k})]^T] \cdot E[C_{k+1}(x_{k+1} - \hat{x}_{k+1|k})((x_{k+1} - \hat{x}_{k+1|k})C_{k+1})^T + R_k]^{-1} \cdot (y_{k+1} - \hat{y}_{k+1|k} - 0)$$

$$\Rightarrow \hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + \underline{P_{k+1|k} C_{k+1}^T [C_{k+1} P_{k+1|k} C_{k+1}^T + R_k]^{-1}} [y_{k+1} - C_{k+1} \hat{x}_{k+1|k}]$$

let this $= K_{k+1}$

$$\Rightarrow \hat{x}_{k+1|k+1} = \hat{x}_{k+1|k} + K_{k+1} [y_{k+1} - C_{k+1} \hat{x}_{k+1|k}] \quad ※$$

FACT $P_{k+1|k+1}$ derivation

proof define $\tilde{x}_{k+1|k+1} = x_{k+1} - \hat{x}_{k+1|k+1}$

$$\therefore \tilde{x}_{k+1|k+1} = x_{k+1} - (\hat{x}_{k+1|k} + K_{k+1}[y_{k+1} - C_{k+1}\hat{x}_{k+1|k}])$$
$$= x_{k+1} - \hat{x}_{k+1|k} - K_{k+1}[y_{k+1} - \hat{y}_{k+1|k}]$$
$$= \tilde{x}_{k+1|k} - K_{k+1} \tilde{y}_{k+1|k}$$
$$= \tilde{x}_{k+1|k} - K_{k+1}[C_{k+1}\tilde{x}_{k+1|k} + v_{k+1}]$$

$$\therefore P_{k+1|k+1} = E[\tilde{x}_{k+1|k+1} \tilde{x}_{k+1|k+1}^T]$$
$$= E[\tilde{x}_{k+1|k} \tilde{x}_{k+1|k}^T] - K_{k+1} C_{k+1} E[\tilde{x}_{k+1|k} \tilde{x}_{k+1|k}^T]$$
$$= P_{k+1|k} - K_{k+1} C_{k+1} P_{k+1|k}$$
$$= [I - K_{k+1} C_{k+1}] P_{k+1|k}$$

(10) final $\hat{x}_{k+1|k+1}$ & $P_{k+1|k+1}$

FACT Kalman gain revisit

$$K_{k+1} = P_{k+1|k} C_{k+1}^T [C_{k+1} P_{k+1|k} C_{k+1}^T + R_k]^{-1} = \arg\min_K \operatorname{tr}(P(k))$$

proof

$$[I - K_{k+1} C_{k+1}] P_{k+1|k}$$
$$= P_{k+1|k} - K_{k+1} C_{k+1} P_{k+1|k}$$
$$= P_{k+1|k} - P_{k+1|k} C_{k+1}^T K_{k+1}^T - K_{k+1} C_{k+1} P_{k+1|k} + P_{k+1|k} C_{k+1}^T K_{k+1}^T$$
$$= P_{k+1|k} - P_{k+1|k} C_{k+1}^T K_{k+1}^T - K_{k+1} C_{k+1} P_{k+1|k} + P_{k+1} C_{k+1}^T (C_{k+1} P_{k+1|k} C_{k+1}^T + R_k)^{-1} (C_{k+1} P_{k+1|k} C_{k+1}^T + R_k) K_{k+1}^T$$
$$= P_{k+1|k} - P_{k+1|k} C_{k+1}^T K_{k+1}^T - K_{k+1} C_{k+1} P_{k+1|k} + K_{k+1} (C_{k+1} P_{k+1|k} C_{k+1}^T + R_k) K_{k+1}^T \quad — (\#)$$

No

Date

proof (cont'd)

from ㊉

$P_{k+1|k+1} = P(K)$

$= [I - K_{k+1}C_{k+1}]P_{k+1|k} = P_{k+1|k} - P_{k+1|k}C_{k+1}^T K_{k+1}^T - K_{k+1}C_{k+1}P_{k+1|k} + K_{k+1}(C_{k+1}P_{k+1|k}C_{k+1}^T + R_k)K_{k+1}^T$

- consider it as $P(K)$, and we want it to hv minimum variance on each element of x, where it occurs at (for readability $P_{k+1|k} = \bar{P}$, $C_{k+1} = H$, $K_{k+1} = K$.

$$\frac{d\,tr(P(K))}{dK} = 0 - (H\bar{P})^T - \bar{P}H^T + 2K(H\bar{P}H^T + R) = 0$$

$$\frac{d\,tr(KH\bar{P})}{dK} = \frac{d\,tr(AB)}{dA} = B^T = (H\bar{P})^T = \bar{P}^T H^T = \bar{P}H^T$$

$$\frac{d\,tr(\bar{P}H^T K^T)}{dK} = \bar{P}H^T$$

$$\frac{d\,tr(K(H\bar{P}H^T + R)K^T)}{dK} = 2K(H\bar{P}H^T + R)$$

$$\Rightarrow \frac{d\,tr(P(K))}{dK} = -2\bar{P}H^T + 2K(H\bar{P}H^T + R) = 0$$

$$\therefore 2\bar{P}H^T = 2K(H\bar{P}H^T + R)$$

$$K = \bar{P}H^T(H\bar{P}H^T + R)^{-1}$$

※