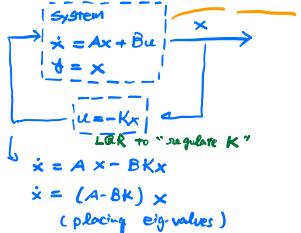


## Motivation for Full-state estimation (from Steve Brunton)

Recall

$$\dot{x} = Ax + Bu \quad x \in \mathbb{R}^n$$

$$u \in \mathbb{R}^m$$



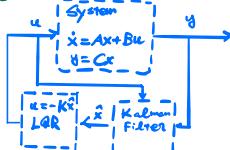
I don't necessarily hv all states in real life

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

$$y = Cx \quad (\text{observability}) \quad \text{obsv}(A, C)$$

Main Question here:  
Can I estimate any state  $\hat{x}$   
from measurement  $y(t)$

hence:



### Observability

- Duality exists between  $\frac{\partial}{\partial u}$  and  $\frac{\partial}{\partial C}$

- observability matrix

$$\mathcal{O} = \begin{bmatrix} C \\ CA \\ CA^2 \\ \vdots \\ CA^{n-1} \end{bmatrix}$$

$$\mathcal{O} = \begin{bmatrix} B & AB & A^2B & \dots & A^{n-1}B \end{bmatrix}$$

1. observable if

$$\Rightarrow \text{rank}(\text{obsv}(A, C)) = n$$

2. can estimate  $x$  from  $y$

$$\Rightarrow [U, Z, V] = \text{svd}(\mathcal{O})$$

observability gramian

$$V^T$$

In some direction,  
we hv higher  
signal to noise

## Full State Estimation



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

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

$$\begin{aligned} \frac{d}{dt} \hat{x} &= A \hat{x} + Bu + K_f y - K_f C \hat{x} \\ &= (A - K_f C) \hat{x} + [B \quad K_f] \begin{bmatrix} u \\ y \end{bmatrix} \end{aligned}$$

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

### Kalman filter

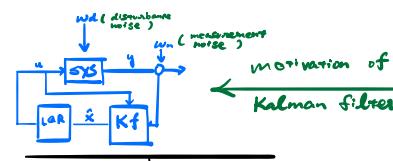
- $w_d$  - Gaussian
- $v_d$  - Variance
- $w_n$  - Gaussian
- $v_n$  - Variance
- recall
- $\hat{E} = (A - K_f C) E$
- $E = x - \hat{x}$
- cost function
- $J = E((x - \hat{x})^T(x - \hat{x}))$
- $\ggg K_f = \text{arg} \min_{K_f} E(A, C, v_d, v_n)$

Kalman filter

real system:

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

$$y = Cx + v_n$$



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

$$\text{Error } E = x - \hat{x}$$

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

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

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

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

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

$$= (A - K_f C) E$$

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

more detailed derivation

Kalman Filtering

P1

 $\theta \in \mathbb{R}^n$  random vector $Y_i^k = [y(1), y(2), \dots, y(k-1), y(k)]$  observation $y(i) \& \theta$  independentJoint pdf of  $\theta$  &  $Y_i^k$ :  $P(\theta, Y_i^k)$ conditional pdf of  $\theta | Y_i^k$ :  $P(\theta | Y_i^k)$ pdf of  $Y_i^k$ :  $P(Y_i^k)$ The Estimation Problemgiven  $y(1), y(2), \dots, y(k)$ : evaluate  $\theta$ 

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

Consider

mean-square error for General Estimation of Random Parameters

(3.4) cost function

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

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

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

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

proof  $\Delta E[\tilde{\theta}(k)^T \tilde{\theta}(k)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{\theta}(k)^T \tilde{\theta}(k) P(\theta, Y_i^k) d\theta dY_i^k$  to the joint pdf of  $\theta$  &  $Y_i^k$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{\theta}(k)^T \tilde{\theta}(k) P(\theta, Y_i^k) d\theta_1 d\theta_2 \dots d\theta_n dY_1 dY_2 \dots dY_K$$

$$\Rightarrow \text{Bayes Law: } P(\theta, Y_i^k) = P(\theta | Y_i^k) P(Y_i^k)$$

$$\Delta E[\tilde{\theta}(k)^T \tilde{\theta}(k)] = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \tilde{\theta}(k)^T \tilde{\theta}(k) P(\theta | Y_i^k) P(Y_i^k) d\theta_1 d\theta_2 \dots d\theta_n dY_1 dY_2 \dots dY_K$$

$$= \int_{-\infty}^{\infty} \left[ \int_{-\infty}^{\infty} \tilde{\theta}(k)^T \tilde{\theta}(k) P(\theta | Y_i^k) d\theta_1 d\theta_2 \dots d\theta_n \right] P(Y_i^k) dY_1 dY_2 \dots dY_K$$

$$= \int_{-\infty}^{\infty} E[\tilde{\theta}(k)^T \tilde{\theta}(k) | Y_i^k] P(Y_i^k) dY_K$$

minimize this

Conclusion #1 as in (3.6)

$$\hat{\theta}(k) = \underset{\theta(k)}{\operatorname{argmin}} E[\tilde{\theta}(k)^T \tilde{\theta}(k) | Y_i^k]$$

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

FACT for the minimization problem

$$(3.9) \quad \hat{\theta}(k) = \arg\min E[\hat{\theta}(k)^T \hat{\theta}(k) | Y^k]$$

we have

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

proof for 3.5 & 3.9

$$\begin{aligned} J &= E[\hat{\theta}(k)^T \hat{\theta}(k) | Y^k] \\ &= E[(\theta - \hat{\theta}(k))^T (\theta - \hat{\theta}(k)) | Y^k] \\ \Rightarrow J &= E[\theta^T \theta - \theta^T \hat{\theta}(k) - \hat{\theta}(k)^T \theta + \hat{\theta}(k)^T \hat{\theta}(k) | Y^k] \\ &= E[\theta^T \theta | Y^k] - E[\theta^T | Y^k] \hat{\theta}(k) - \hat{\theta}(k)^T E[\theta | Y^k] + E[\hat{\theta}(k)^T \hat{\theta}(k) | Y^k] \\ &= \underbrace{E[\theta^T \theta | Y^k]}_{\textcircled{1}} - \underbrace{E[\theta^T | Y^k] \hat{\theta}(k)}_{-E[\theta^T | Y^k] \hat{\theta}(k)} - \underbrace{\hat{\theta}(k)^T E[\theta | Y^k]}_{-\hat{\theta}(k)^T E[\theta | Y^k]} + \underbrace{E[\hat{\theta}(k)^T \hat{\theta}(k) | Y^k]}_{E[\hat{\theta}(k)^T \hat{\theta}(k)]} + E[\theta^T | Y^k] E[\theta | Y^k] \\ &= \underbrace{E[\theta^T \theta | Y^k]}_{\textcircled{2}} - E[\theta^T | Y^k] E[\theta | Y^k] + E[\theta^T | Y^k] E[\theta | Y^k] - \hat{\theta}(k)^T E[\theta | Y^k] + \hat{\theta}(k)^T \hat{\theta}(k) \\ &= E[\theta^T \theta | Y^k] - E[\theta^T | Y^k] E[\theta | Y^k] + [\hat{\theta}(k) - E[\theta | Y^k]]^T [\hat{\theta}(k) - E[\theta | Y^k]] \end{aligned}$$

$$\begin{aligned} &= E[\hat{\theta}(k)^T \hat{\theta}(k) | Y^k] \\ &= \int_{-\infty}^{\infty} \hat{\theta}(k)^T \hat{\theta}(k) P(Y^k) dY^k \\ &= \int_{-\infty}^{\infty} P(Y^k) dY^k \cdot \hat{\theta}(k)^T \hat{\theta}(k) \\ &= 1 \cdot \hat{\theta}(k)^T \hat{\theta}(k) \\ &= \hat{\theta}(k)^T \hat{\theta}(k) \end{aligned}$$

$\therefore$  minimization problem becomes:

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

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

conclusion #2  $\hat{\theta}(k) = E[\theta | Y^k]$  (3.2.1) (3) with some formulation, the optimal  $\hat{\theta}(k)$  is this.

### Gaussian Particularization

FACT  $\theta, Y^k$  are jointly Gaussian

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

$$\downarrow \quad E[(\theta - E(\theta))(Y^k - E(Y^k))^T]$$

$$\downarrow \quad E[(Y^k - E(Y^k))(Y^k - E(Y^k))^T]$$

$$\text{proof } E[X_a | X_b] = E[Z + CX_b | X_b] \quad \text{remark: we want } Z \text{ to be independent } \rightarrow X_b \text{ so that } E[Z | X_b] = E[Z].$$

$$\downarrow \quad = E[Z | X_b] + E[ZX_b | X_b]$$

$$\downarrow \quad = E[Z | X_b] + CX_b$$

$$\text{let } Z = X_a - CX_b \quad \downarrow \quad = E[X_a - CX_b] + CX_b$$

to be independent / uncorrelated

to  $X_b$ :

$$= E[X_a] + C(X_b - E[X_b]) = E[X_a] + Z_{ab} Z_{bb}^{-1} (X_b - E[X_b]) \quad \text{④ prior to this, ③ applies to any distribution.}$$

$$\begin{aligned} 0 &= \text{cov}(Z, X_b) = \text{cov}(X_a - CX_b, X_b) \\ &= \text{cov}(X_a, X_b) - C \text{cov}(X_b, X_b) \\ &= Z_{ab} - C Z_{bb} \quad \therefore C = Z_{ab} Z_{bb}^{-1} \quad \text{for Gaussian, it's:} \\ &= E[\theta] + \sum_{\theta Y^k} Z_{bb}^{-1} (Y^k - E[Y^k]) \end{aligned}$$

## Kalman Filtering

FACT consider general filtering problem:

⑤ A linear dynamics (time-varying) P3  
assumption: LINEAR QUADRATIC GAUSSIAN

$$x(k+1) = f(x(k), u(k), w(k)) \quad (4.1)$$

$$y(k) = h(x(k), v(k)) \quad (4.2)$$

particularizes to

$$x_{k+1} = A_k x_k + B_k u_k + G w_k$$

$$y_k = C_k x_k + V_k v_k$$

obtains minimum mean-square state error estimate

$$E[x_0] = \bar{x}_0 \quad \text{initial state}$$

• joint covariance matrix

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

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

$w_k \leftrightarrow v_k$  independent.

$u_k$  is deterministic

$$\text{FACT} \cdot p(x_k | Y_i^k, U_o^{k-1}) \sim N(\hat{x}_{k|k}, P_{k|k})$$

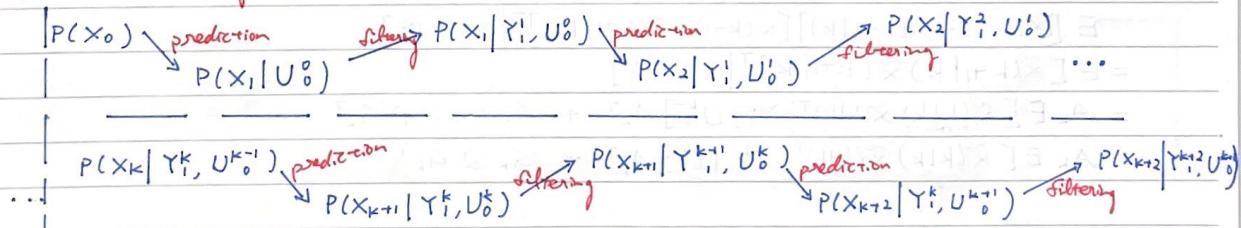
where

$$\hat{x}_{k|k} = E[x_k | Y_i^k, U_o^{k-1}]$$

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

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

## Kalman Filter dynamics



FACT estimates & variance of Kalman filter:

$$(4.8) \cdot p(x_k | Y_i^k, U_o^{k-1}) \sim N(\hat{x}_{k|k}, P_{k|k})$$

(4.12) (4.13)

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

$$(4.9) \cdot p(x_{k+1} | Y_i^k, U_o^k) \sim N(\hat{x}_{k+1|k}, P_{k+1|k})$$

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

⑥ propagation of

1st & 2nd moment of  
the distribution

$$\hat{x}_{k|k} = E[x_k | Y_i^k, U_o^{k-1}]$$

$$\hat{x}_{k+1|k} = E[x_{k+1} | Y_i^k, U_o^k]$$

(4.10) (4.11)

Prediction ⑦ based on the linear dynamics:  
prediction propagation

P4

P5

### FACT State Prediction

$$P(x_{k+1} | Y_i^k, U_o^k)$$

- recall  $x_{k+1} = A_k x_k + B_k u_k + G_k w_k$

$$(4.14) \quad E[x_{k+1} | Y_i^k, U_o^k] = A_k E[x_k | Y_i^k, U_o^{k-1}] + B_k E[u_k | Y_i^k, U_o^{k-1}] + G_k E[w_k | Y_i^k, U_o^{k-1}]$$

- recall  $P(x_k | Y_i^k, U_o^{k-1}) \sim N(\hat{x}_{k|k}, P_{k|k})$

$$P(x_{k+1} | Y_i^k, U_o^k) \sim N(\hat{x}_{k+1|k}, P_{k+1|k})$$

&  $w_k$  &  $y_i^k$  are independent /  $\text{cov}(w_k, y_i^k) = 0$ , and  $w_k = 0$

$$4.14 \Rightarrow E[x_{k+1} | Y_i^k, U_o^k] = A_k E[x_k | Y_i^k, U_o^{k-1}] + B_k E[u_k | Y_i^k, U_o^{k-1}] + G_k E[w_k | Y_i^k, U_o^{k-1}]$$

$$\hat{x}_{k+1|k} = A_k \hat{x}_{k|k} + B_k u_k + G_k w_k \approx 0$$

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

### FACT Prediction Error

$$\begin{aligned} \text{let } \tilde{x}_{k+1|k} &= x_{k+1} - \hat{x}_{k+1|k} \\ &= A_k x_k + B_k u_k + G_k w_k - (A_k \hat{x}_{k|k} + B_k u_k) \\ &= A_k (x_k - \hat{x}_{k|k}) + G_k w_k \\ &= A_k \tilde{x}_{k|k} + G_k w_k \end{aligned}$$

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

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

### FACT Covariance Matrix

$$\begin{aligned} &E[(\tilde{x}_{k+1} - \hat{x}_{k+1|k})(\tilde{x}_{k+1} - \hat{x}_{k+1|k})^T | Y_i^k, U_o^k] \\ &= E[\tilde{x}_{k+1|k} \tilde{x}_{k+1|k}^T | Y_i^k, U_o^k] \\ &= A_k E[\tilde{x}_{k|k} \tilde{x}_{k|k}^T | Y_i^k, U_o^{k-1}] A_k^T + G_k Q G_k^T \\ &= A_k E[\tilde{x}_{k|k} \tilde{x}_{k|k}^T | Y_i^k, U_o^{k-1}] A_k^T + G_k Q G_k^T \end{aligned}$$

$$\cdot \text{recall } P_{k|k} = E[(x_k - \hat{x}_{k|k})(x_k - \hat{x}_{k|k})^T | Y_i^k, U_o^{k-1}]$$

$$\cdot P_{k+1|k} = E[(x_{k+1} - \hat{x}_{k+1|k})(x_{k+1} - \hat{x}_{k+1|k})^T | Y_i^k, U_o^k]$$

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

## (8) measurement prediction propagation

P5

MeasurementFACT recall  $y(k) = h(x(k), v(k)) \Rightarrow \hat{y}_k = C_k x_k + v_k$ 

$$P(y_{k+1} | Y^k, U^k_0) = P(C_{k+1} x_{k+1} + v_{k+1} | Y^k, U^k_0) \quad (4.21)$$

(Gaussian pdf)

$$\text{FACT } \hat{y}_{k+1|k} = E[y_{k+1} | Y^k, U^k_0] = C_{k+1} \hat{x}_{k+1|k} \quad (4.22)$$

(predicted measurement)

$$\begin{aligned} \text{also } \tilde{y}_{k+1|k} &\stackrel{\Delta}{=} y_{k+1} - \hat{y}_{k+1|k} && \text{(measurement prediction error)} \\ &= C_{k+1} x_{k+1} + v_{k+1} - C_{k+1} \hat{x}_{k+1|k} \\ &= C_{k+1} (x_{k+1} - \hat{x}_{k+1|k}) + v_{k+1} \\ &= C_{k+1} \tilde{x}_{k+1|k} + v_{k+1} \end{aligned}$$

FACT ∵ covariance matrix of  $\hat{y}$ 

$$\begin{aligned} P_{\hat{y}} &= E[\tilde{y}_{k+1|k} \tilde{y}_{k+1|k}^T | Y^k, U^k_0] = E[C_{k+1} \tilde{x}_{k+1|k} \tilde{x}_{k+1|k}^T C_{k+1}^T + v_{k+1} v_{k+1}^T | Y^k, U^k_0] \\ &= C_{k+1} E[\tilde{x}_{k+1|k} \tilde{x}_{k+1|k}^T | Y^k, U^k_0] C_{k+1}^T + R_k \\ &= C_{k+1} P_{k+1|k} C_{k+1}^T + R_k \end{aligned}$$

CorrectionFACT  $E[(\theta - \hat{\theta}_k) f^*(Y^k)] = 0$  i.e., estimation error is orthogonal to function of $Y^k$ proof lemma 1  $E[x g(y)] = E[E(x|y) g(y)]$ 

lemma 1 proof

$$\begin{aligned} &E[E(x|y) g(y)] \\ &= \sum_y E[x|Y=y] dP(y) \sum_y g(y) dP(g|y) \end{aligned}$$

$$\begin{aligned} \text{lemma 2} \quad &= E[E(x|y)] \quad \text{our variable which takes on value } E[x|Y=y] \\ E[E[x|y]] &= \sum_y E[x|Y=y] dP(Y=y) \quad \text{when } Y=y \\ = E[x] &= \sum_y \sum_x x dP(x \cap Y=y) dP(Y=y) \\ &= \sum_y \sum_x x dP(x \cap Y=y) \quad (\text{joint distribution } \Leftrightarrow \text{two distributions' intersection}) \\ &= \sum_x x dP(x) \end{aligned}$$

$$= E[x]$$

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

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

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

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

No \_\_\_\_\_

Date \_\_\_\_\_

P6

proof (cont'd)

$$\begin{aligned} E[xg(y)] &= E[E(x|y)g(y)] \\ \therefore E[(\theta - \hat{\theta}_k) f^*(Y_i^k)] &= E[E[\theta - \hat{\theta}_k | Y_i^k] f^*(Y_i^k)] \\ \text{lemma 3 } E[\theta \hat{\theta}_k | Y_i^k] &= 0 \end{aligned}$$

lemma 3 proof

$$\begin{aligned} E[\theta - \hat{\theta}_k | Y_i^k] &= E[\theta | Y_i^k] - E[\hat{\theta}_k | Y_i^k] \\ &= E[\theta | Y_i^k] - \hat{\theta}_k \quad \text{lemma 4 } E[\hat{\theta}_k | Y_i^k] = \hat{\theta}_k \\ \text{lemma 5 } \hat{\theta}_k &= E[\theta | Y_i^k] \Rightarrow E[\hat{\theta}_k | Y_i^k] = E[g(Y_i^k) | Y_i^k] \\ \text{has optimal estimation} &= \sum_{Y_i^k} g(Y_i^k) \cdot dP(Y_i^k) \\ \therefore E[\theta - \hat{\theta}_k | Y_i^k] &= 0 \quad = g(Y_i^k) = \hat{\theta}_k \\ \therefore E[E[\theta - \hat{\theta}_k | Y_i^k] f^*(Y_i^k)] &= E[\theta f^*(Y_i^k)] \\ &= 0 \quad (\text{from corollary 3.2.1}) \end{aligned}$$

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

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

FACT

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

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

$$\hat{x}_{k+1|k+1} = \sum_x x \cdot dP(x_{k+1} | Y_i^{k+1}, U_o^k)$$

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

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

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

from Gaussian Particularization

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

$$\text{here } x_a = z = E[x_{k+1} | Y_i^k, U_o^k]$$

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

① innovation/correction:

try to get

"optimal estimation"

given new  $y_{k+1|k}$   
& model prediction

proof (cont'd) $E[x_a]$  $\Sigma_{ab}$ 

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

$$\cdot E[\tilde{y}_{k+1|k} \tilde{y}_{k+1|k}^T]^{-1} \Sigma_{bb}^{-1} \\ \cdot (\tilde{y}_{k+1|k} - E[\tilde{y}_{k+1|k}]) - x_b - E[x_b] =$$

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

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

$$\cdot (y_{k+1} - \hat{y}_{k+1|k} = 0)$$

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

let this =  $K_{k+1}$

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

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

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

$$= x_{k+1} - \hat{x}_{k+1|k} - K_{k+1} [y_{k+1} - \hat{y}_{k+1|k}]$$

$$= \tilde{x}_{k+1|k} - K_{k+1} \tilde{y}_{k+1|k}$$

$$= \tilde{x}_{k+1|k} - K_{k+1} [C_{k+1} \tilde{x}_{k+1|k} + v_{k+1}]$$

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

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

$$= P_{k+1|k} - K_{k+1} C_{k+1} P_{k+1|k}$$

$$= [I - K_{k+1} C_{k+1}] P_{k+1|k}$$

(1) final  $\hat{x}_{k+1|k+1}$  &  $P_{k+1|k+1}$ FACT Kalman gain revisit $P_{k+1|k+1}$ 

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

proof

$$[I - K_{k+1} C_{k+1}] P_{k+1|k}$$

$$= P_{k+1|k} - K_{k+1} C_{k+1} P_{k+1|k}$$

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

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

$$= P_{k+1|k} - P_{k+1|k} C_{k+1}^T K_{k+1}^T - K_{k+1} C_{k+1} P_{k+1|k} + K_{k+1} (C_{k+1} P_{k+1|k} C_{k+1}^T + R_k) K_{k+1}^T - \textcircled{4}$$

No \_\_\_\_\_

Date \_\_\_\_\_

P8

proof (contd)

From  $\oplus$ 

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

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

$$\begin{aligned} \frac{d\text{tr}(P(k))}{dK} &= 0 - (H\bar{P})^T - PH^T + 2K(H\bar{P}H^T + R) = 0 \\ \frac{d\text{tr}(K\bar{P})}{dK} &= \frac{d\text{tr}(AB)}{dA} = B^T = (H\bar{P})^T = \bar{P}^T H^T = \bar{P}H^T \\ \frac{d\text{tr}(\bar{P}H^T K^T)}{dK} &= \bar{P}H^T \\ \frac{d\text{tr}(K(H\bar{P}H^T + R)K^T)}{dK} &= 2K(H\bar{P}H^T + R) \\ \Rightarrow \frac{d\text{tr}(P(k))}{dK} &= -2\bar{P}H^T + 2K(H\bar{P}H^T + R) = 0 \end{aligned}$$

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

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