Linear Gaussian Estimation

assumption: discrete, linear, time varying

- motion model:

$$x_k = A_{k-1} x_{k-1} + v_k + w_k$$

(x_{k-1}: random; v_k: input, deterministic; w_k: process noise, random)

$$y_k = C_k x_k + n_k$$

$$x_0 \in \mathbb{R}^N \sim \mathcal{N}(\check{x}_0, \check{P}_0)$$
$$w_k \in \mathbb{R}^N \sim \mathcal{N}(0, Q_k)$$
$$n_k \in \mathbb{R}^M \sim \mathcal{N}(0, R_k)$$

- batch linear-Gaussian estimation problem
  - Bayesian
  - Maximum a Posteriori

Maximum a Posteriori

- $$\hat{x} = \arg\max_x P(x \mid v, y)$$

$x = x_{0:K} = (x_0, x_1, \dots x_K)$

$v = (\check{x}_0, v_{1:K}) = (\check{x}_0, v_1, \dots v_K)$ for prior

$y = y_{0:K} = (y_0, \dots y_K)$ for posterior

- Bayes' rule:

$$\hat{x} = \arg\max_x P(x \mid v, y)$$
$$= \arg\max_x \frac{P(y \mid x, v) P(x \mid v)}{P(y \mid v)}$$

v does not affect y; $P(y \mid v)$ does not affect x

$$= \arg\max_x P(y \mid x) P(x \mid v)$$

$$P(A \mid B) = \frac{P(A \cap B)}{P(B)}$$
$$P(B \mid A) = \frac{P(A \cap B)}{P(A)}$$
$$\therefore P(A \cap B) = P(B \mid A) P(A)$$
$$\therefore P(A \mid B) = \frac{P(B \mid A) P(A)}{P(B)}$$

$$P(A \mid BC) = \frac{P(A \cap B \cap C)}{P(B \cap C)}$$
$$P(C \mid AB) = \frac{P(A \cap B \cap C)}{P(A \cap B)}$$
$$\therefore P(A \cap B \cap C) = P(A \cap B) P(C \mid A \cap B) = P(A \mid B) P(B) P(C \mid A \cap B)$$
$$\therefore P(A \mid BC) = \frac{P(A \mid B) P(B) P(C \mid A \cap B)}{P(B) P(C \mid B)} = \frac{P(A \mid B) P(C \mid A \cap B)}{P(C \mid B)}$$

- assume each w_k, n_k are NOT correlated:

$$P(y \mid x) = \prod_{k=0}^{K} P(y_k \mid x_k)$$

- Bayes' Rule:

$$P(x \mid v) = P(x_0 \mid \check{x}_0) \prod_{k=0}^{K} P(x_k \mid x_{k-1}, v_k)$$

$$\therefore P(x_0 \mid \check{x}_0) = \frac{1}{\sqrt{(2\pi)^N \det \check{P}_0}} \times \exp\left(-\frac{1}{2}(x_0 - \check{x}_0)^T \check{P}_0^{-1} (x_0 - \check{x}_0)\right)$$

$$\therefore P(x_k \mid x_{k-1}, v_k) = \frac{1}{\sqrt{(2\pi)^N \det Q_k}} \times \exp\left(-\frac{1}{2}(x_k - (A_{k-1}x_{k-1} + v_k))^T Q_k^{-1} (x_k - (A_{k-1}x_{k-1} + v_k))\right)$$

$$\therefore P(y_k \mid x_k) = \frac{1}{\sqrt{(2\pi)^M \det R_k}} \times \exp\left(-\frac{1}{2}(y_k - C_k x_k)^T R_k^{-1} (y_k - C_k x_k)\right)$$

- logarithm

$$\ln(P(y \mid x) P(x \mid v))$$
$$= \ln P(x_0 \mid \check{x}_0) + \sum_{k=1}^{K} \ln P(x_k \mid x_{k-1}, v_k) + \sum_{k=1}^{K} \ln P(y_k \mid x_k)$$

where

$$\ln P(x_0 \mid \check{x}_0) = -\frac{1}{2}(x_0 - \check{x}_0)^T \check{P}_0^{-1} (x_0 - \check{x}_0) + \cancel{-\frac{1}{2}\ln((2\pi)^N \det \check{P}_0)}$$
$$\ln P(x_k \mid x_{k-1}, v_k) = -\frac{1}{2}(x_k - A_{k-1}x_{k-1} - v_k)^T Q_k^{-1} (x_k - A_{k-1}x_{k-1} - v_k) + \cancel{-\frac{1}{2}\ln((2\pi)^N \det Q_k)}$$
$$\ln P(y_k \mid x_k) = -\frac{1}{2}(y_k - C_k x_k)^T R_k^{-1} (y_k - C_k x_k) + \cancel{-\frac{1}{2}\ln((2\pi)^M \det R_k)}$$

- J, cost

$$J_{v,k}(x) = \begin{cases} \frac{1}{2}(x_0 - \check{x}_0)^T \check{P}_0^{-1} (x_0 - \check{x}_0) \\ \frac{1}{2}(x_k - A_{k-1}x_{k-1} - v_k)^T Q_k^{-1} (x_k - A_{k-1}x_{k-1} - v_k) \end{cases}$$

$$J_{y,k}(x) = \frac{1}{2}(y_k - C_k x_k)^T R_k^{-1} (y_k - C_k x_k)$$

Mahalanobis distance

$$J(x) = \sum_{k=0}^{K} (J_{v,k}(x) + J_{y,k}(x))$$

$$\therefore x = \arg\min_x J(x)$$

- from previous

$$J_{v,k}(x) = \begin{cases} \frac{1}{2}(x_0 - \check{x}_0)^T \check{P}_0^{-1} (x_0 - \check{x}_0) \\ \frac{1}{2}(x_k - A_{k-1}x_{k-1} - v_k)^T Q_k^{-1} (x_k - A_{k-1}x_{k-1} - v_k) \end{cases}$$

$$J_{y,k}(x) = \frac{1}{2}(y_k - C_k x_k)^T R_k^{-1} (y_k - C_k x_k)$$

let

$$z = \begin{bmatrix} \check{x}_0 \\ v_1 \\ \vdots \\ v_K \\ \hline y_0 \\ y_1 \\ \vdots \\ y_K \end{bmatrix}, \quad x = \begin{bmatrix} x_0 \\ \vdots \\ x_K \end{bmatrix}$$

&

$$H = \begin{bmatrix} 1 & & & \\ -A_0 & 1 & & \\ & \ddots & \ddots & \\ & & -A_{K-1} & 1 \\ \hline C_0 & & & \\ & C_1 & & \\ & & \ddots & \\ & & & C_K \end{bmatrix}$$

$$W = \begin{bmatrix} \check{P}_0 & & & & & & \\ & Q_1 & & & & & \\ & & \ddots & & & & \\ & & & Q_K & & & \\ & & & & R_0 & & \\ & & & & & R_1 \ddots & \\ & & & & & & R_K \end{bmatrix}$$

$$\therefore J(x) = \frac{1}{2}(z - Hx)^T W^{-1} (z - Hx)$$

$$P(z \mid x) = \eta \exp\left(-\frac{1}{2}(z - Hx)^T W^{-1} (z - Hx)\right)$$

η: normalization constant

$$\therefore \left.\frac{\partial J(x)}{\partial x^T}\right|_{\hat{x}} = -H^T W^{-1}(z - H\hat{x}) = 0$$

$$\Rightarrow (H^T W^{-1} H)\hat{x} = H^T W^{-1} z$$

Bayesian Inference

compute full Bayesian Posterior

$P(x \mid v, y)$

- $$x_k = A_{k-1} x_{k-1} + v_k + w_k$$

$$\xrightarrow{\text{lift}} x = A(v + w)$$

where

$$A = \begin{bmatrix} 1 & & & & \\ A_0 & 1 & & & \\ A_1 A_0 & A_1 & 1 & & \\ \vdots & \vdots & \vdots & \ddots & \\ A_{K-2}\cdots A_0 & A_{K-2}\cdots A_1 & A_{K-2}\cdots A_2 & \cdots\ 1 & \\ A_{K-1}\cdots A_0 & A_{K-1}\cdots A_1 & A_{K-1}\cdots A_2 & \cdots\ A_{K-1} & 1 \end{bmatrix}$$

- $$\check{x} = E[x] = E[A(v + w)] = E[Av + Aw] = Av$$

- $$\check{P} = E[(x - E[x])(x - E[x])^T] = AQA^T$$

$Q = E[ww^T] = \mathrm{diag}(\check{P}_0, Q_1, \dots Q_K)$

- prior $\therefore \Rightarrow$

$$P(x \mid v) = \mathcal{N}(\check{x}, \check{P}) = \mathcal{N}(Av, AQA^T)$$

- $$y_k = C_k x_k + n_k$$

$$\xrightarrow{\text{lift}} y = Cx + n$$

where

$$C = \mathrm{diag}(C_0, C_1, \dots, C_K)$$

- joint density of prior lifted state & measurements is now:

$$P(x, y \mid v) = \mathcal{N}\left(\begin{bmatrix} \check{x} \\ C\check{x} \end{bmatrix}, \begin{bmatrix} \check{P} & \check{P}C^T \\ C\check{P} & C\check{P}C^T + R \end{bmatrix}\right)$$

$$R = E[nn^T] = \mathrm{diag}(R_0, R_1, \dots, R)$$

- $$P(x, y \mid v) = \underline{P(x \mid y, v)}\, P(y \mid v)$$

we only care about this

- $$P(x, y) = \mathcal{N}\left(\begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}, \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}\right)$$

- $$P(x, y) = P(x \mid y) P(y)$$

w/ schur complement

$$\begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix} = \begin{bmatrix} I & \Sigma_{xy}\Sigma_{yy}^{-1} \\ 0 & I \end{bmatrix} \begin{bmatrix} \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} & 0 \\ 0 & \Sigma_{yy} \end{bmatrix} \begin{bmatrix} I & 0 \\ \Sigma_{yy}^{-1}\Sigma_{yx} & I \end{bmatrix}$$

$$\Rightarrow \begin{bmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{bmatrix}^{-1} = \begin{bmatrix} I & 0 \\ -\Sigma_{yy}^{-1}\Sigma_{yx} & I \end{bmatrix} \begin{bmatrix} (\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1} & 0 \\ 0 & \Sigma_{yy}^{-1} \end{bmatrix} \begin{bmatrix} I & -\Sigma_{xy}\Sigma_{yy}^{-1} \\ 0 & I \end{bmatrix}$$

- PDF of $P(x, y)$

recall:

$$P(x) = (2\pi)^{-\frac{N}{2}} \det(\Sigma)^{-\frac{1}{2}} \exp\left(-\frac{1}{2}(x - \mu)^T \Sigma^{-1} (x - \mu)\right)$$

$$\Rightarrow P(x, y) = (2\pi)^{-\frac{N}{2}} \det(\Sigma)^{-\frac{1}{2}} \exp(Z)$$

$$Z = \left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}\right)^T \begin{pmatrix} \Sigma_{xx} & \Sigma_{xy} \\ \Sigma_{yx} & \Sigma_{yy} \end{pmatrix}^{-1} \left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}\right)$$

$$= \left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}\right)^T \begin{bmatrix} I & 0 \\ -\Sigma_{yy}^{-1}\Sigma_{yx} & I \end{bmatrix} \begin{bmatrix} (\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1} & 0 \\ 0 & \Sigma_{yy}^{-1} \end{bmatrix} \begin{bmatrix} I & -\Sigma_{xy}\Sigma_{yy}^{-1} \\ 0 & I \end{bmatrix} \left(\begin{bmatrix} x \\ y \end{bmatrix} - \begin{bmatrix} \mu_x \\ \mu_y \end{bmatrix}\right)$$

$$= [x - \mu_x - \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y)]^T (\Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})^{-1} \times [x - \mu_x - \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y)] + (y - \mu_y)^T \Sigma_{yy}^{-1} (y - \mu_y)$$

sum of two quadratic terms

$$P(x, y) = P(x \mid y) P(y)$$
$$P(x \mid y) = \mathcal{N}(\mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y),\ \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})$$
$$P(y) = \mathcal{N}(\mu_y, \Sigma_{yy})$$

on $(H^T W^{-1} H)\hat{x} = H^T W^{-1} z$

- unique solution:

$$\hat{x} = \underline{(H^T W^{-1} H)^{-1}} H^T W^{-1} z$$

need to be invertible

$W\ \hat{x} \in \mathbb{R}^{K+1}$

$$\mathrm{rank}(H^T W^{-1} H) = K+1$$
$$\Rightarrow \mathrm{rank}(H^T H) = \mathrm{rank}(H^T)$$

- $\mathrm{rank}(H^T)$: we hv $\check{x}_0$ / we don't hv $\check{x}_0$

- we hv $\check{x}_0$

$\mathrm{rank}\ H^T$

$$= \mathrm{rank}\begin{bmatrix} 1 & -A_0^T & & & C_0^T & & & \\ & 1 & -A_1^T & & & C_1^T & & \\ & & \ddots & -A_{K-1}^T & & & C_2^T \ddots & \\ & & & 1 & & & & C_K^T \end{bmatrix}$$

$\Rightarrow$ full rank. $\therefore$ $Q_k \succ 0, R_k \succ 0 \Rightarrow$ hv sol.

- we don't hv $\check{x}_0$

$\mathrm{rank}\ H^T$

$$= \mathrm{rank}\begin{bmatrix} -A_0^T & & & & C_0^T & & & \\ \hline 1 & -A_1^T & & & C_1^T & & & \\ & 1 & \ddots & & & C_2^T & & \\ & & \ddots & -A_{K-1}^T & & & \ddots & \\ & & & 1 & & & & C_K^T \end{bmatrix}$$

$$\Rightarrow \mathrm{rank}\begin{bmatrix} 1 & -A_1^T & & & C_1^T & & \\ & 1 & \ddots & & & C_2^T & \\ & & \ddots & -A_{K-1}^T & & & \ddots \\ & & & 1 & & & C_K^T \\ \hline -A_0^T & & & & C_0^T & & \end{bmatrix}$$

$$\Rightarrow \mathrm{rank}\begin{bmatrix} 1 & -A_1^T & & & C_1^T & & \\ & 1 & \ddots & & & C_2^T & \\ & & \ddots & -A_{K-1}^T & & & \ddots \\ & & & 1 & & & C_K^T \\ \hline & & & & C_0^T & A_0^T C_1^T\ A_0^T A_1^T C_2^T \cdots & A_0^T \cdots A_{K-1}^T C_K^T \end{bmatrix}$$

make sure this CLR

- $$\mathrm{rank}[C_0^T\ A_0^T C_1^T\ A_0^T A_1^T C_2^T \cdots A_0^T \cdots A_{K-1}^T C_K^T] = N$$

from Cayley-Hamilton theorem: any power of $A \in \mathbb{R}^{N \times N}$, power $k \geq N$

$$\Rightarrow (A^T)^{k-1} = a_0 I^T + a_1 A^T + a_2 A^T A^T + \cdots + a_{N-1}(A^T)^{N-1}$$

assume $A_0 = A_1 = A_2 = \cdots = A_{K-1} = A$ & $k >> N$

we first hv

$$(A^T)^{K-1} C^T = a_0 I^T C^T + a_1 A^T C^T + a_2 A^T A^T C^T + \cdots + a_{N-1}(A^T)^{N-1} C^T$$

$$\mathrm{rank}[C_0^T\ A_0^T C_1^T\ A_0^T A_1^T C_2^T \cdots A_0^T \cdots A_{K-1}^T C_K^T]$$
$$\Rightarrow \mathrm{rank}[C^T\ A^T C^T\ A^T A^T C^T \cdots (A^T)^K C^T]$$
$$= \mathrm{rank}[C^T\ A^T C^T \cdots (A^T)^{N-1} C^T]$$

- $$\mathcal{O} = \begin{bmatrix} C \\ CA \\ \vdots \\ CA^{(N-1)} \end{bmatrix}$$

- $\mathrm{rank}\ \mathcal{O} = N$ (observability)

$\therefore$ $Q_k \succ 0$, $R_k \succ 0$, $\mathrm{rank}\ \mathcal{O} = N$

$\Rightarrow$ hv sol.

$$P(x, y \mid v) = P(x \mid y, v) P(y \mid v)$$

$$\therefore P(x \mid y, v)$$

recall $P(x \mid y) = \mathcal{N}(\mu_x + \Sigma_{xy}\Sigma_{yy}^{-1}(y - \mu_y),\ \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx})$

&

$$P(x, y \mid v) = \mathcal{N}\left(\begin{bmatrix} \check{x} \\ C\check{x} \end{bmatrix}, \begin{bmatrix} \check{P} & \check{P}C^T \\ C\check{P} & C\check{P}C^T + R \end{bmatrix}\right)$$

$$= \mathcal{N}\left(\check{x} + \check{P}C^T(C\check{P}C^T + R)^{-1}(y - C\check{x}),\ \check{P} - \check{P}C^T(C\check{P}C^T + R)^{-1}(C\check{P})\right)$$

thru Sherman-Morrison-Woodbury Identity

$$= \mathcal{N}\left((\check{P}^{-1} - C^T R^{-1} C)^{-1}(\check{P}^{-1}\check{x} + C^T R^{-1} y),\ (\check{P}^{-1} + C^T R^{-1} C)^{-1}\right)$$

i.e.,

$$\hat{x} = (\check{P}^{-1} - C^T R^{-1} C)^{-1}(\check{P}^{-1}\check{x} + C^T R^{-1} y)$$
$$\Rightarrow (\check{P}^{-1} - C^T R^{-1} C)\hat{x} = \check{P}^{-1}\check{x} + C^T R^{-1} y$$
$$\Rightarrow ((AQA^T)^{-1} - C^T R^{-1} C)\hat{x} = (AQA^T)^{-1} Av + C^T R^{-1} y$$
$$\Rightarrow (A^{-T}Q^{-1}A^{-1} - C^T R^{-1} C)\hat{x} = A^{-T}Q^{-1}A^{-1}Av + C^T R^{-1} y$$
$$\Rightarrow (A^{-T}Q^{-1}A^{-1} - C^T R^{-1} C)\hat{x} = A^{-T}Q^{-1}v + C^T R^{-1} y$$

$$A^{-1} = \begin{bmatrix} 1 & & & & \\ -A_0 & 1 & & & \\ & -A_1 & 1 & & \\ & & -A_2 & \ddots & \\ & & & & 1 \\ & & & -A_{K-1} & 1 \end{bmatrix}, \quad z = \begin{bmatrix} v \\ y \end{bmatrix}, \quad H = \begin{bmatrix} A^{-1} \\ C \end{bmatrix}, \quad W = \begin{bmatrix} Q \\ R \end{bmatrix}$$

$$\Rightarrow (H^T W^{-1} H)\hat{x} = H^T W^{-1} z$$

solving $(H^T W^{-1} H)\hat{x} = H^T W^{-1} z$

- Cholesky Decomposition

$$H^T W^{-1} H = \underline{LL^T}$$

Cholesky Decomp.

- first decompose this
- $Ax = b$

$LL^T x = b$

$\Rightarrow Ly = b$

- $L^T x = y$
- solve x!

$$\therefore (H^T W^{-1} H)\hat{x} = H^T W^{-1} z$$
$$\Rightarrow LL^T\hat{x} = H^T W^{-1} z$$
$$\Rightarrow Ld = H^T W^{-1} z$$
$$\Rightarrow L^T\hat{x} = d$$

- let LL^T

$$L = \begin{bmatrix} L_0 & & & & & \\ L_{10} & L_1 & & & & \\ & L_{21} & L_2 & & & \\ & & \ddots & \ddots & & \\ & & & L_{K-1,K-2} & L_{K-1} & \\ & & & & L_{K,K-1} & L_K \end{bmatrix}$$

$$L_0 L_0^T = \underbrace{\check{P}_0^{-1} + C_0^T R_0^{-1} C_0}_{I_0} + A_0^T Q_1^{-1} A_0$$
$$L_{10} L_0^T = -Q_1^{-1} A_0$$
$$L_1 L_1^T = -L_{10} L_{10}^T + Q_1^{-1} + C_1^T R_1^{-1} C_1 + A_1^T Q_2^{-1} A_1$$
$$L_{21} L_1^T = -Q_2^{-1} A_1$$
$$\vdots$$
$$L_{K-1} L_{K-1}^T = \underbrace{-L_{K-1,K-2} L_{K-1,K-2}^T + Q_{K-1}^{-1} + C_{K-1}^T R_{K-1}^{-1} C_{K-1}}_{I_{K-1}} + A_{K-1}^T Q_K^{-1} A_{K-1}$$
$$L_{K,K-1} L_{K-1}^T = -Q_K^{-1} A_{K-1}$$
$$L_K L_K^T = \underbrace{-L_{K,K-1} L_{K,K-1}^T + Q_K^{-1} + C_K^T R_K^{-1} C_K}_{I_K}$$

$I_0, I_1, \dots, I_{K-1}, I_K \Rightarrow$ Information Matrix

Basically solve L_0 → solve L_{10} → solve L_1 ⋮ $\quad O(N(K+1))$

forward

- $Ld = H^T W^{-1} z$ get $d = [d_0\ d_1 \dots d_K]^T$

$$L_0 d_0 = \underbrace{\check{P}_0^{-1}\check{x}_0 + C_0^T R_0^{-1} y_0}_{q_0} - A_0^T Q_1^{-1} v_1$$
$$L_1 d_1 = \underbrace{-L_{10} d_0 + Q_1^{-1} v_1 + C_1^T R_1^{-1} y_1}_{q_1} - A_1^T Q_2^{-1} v_2$$
$$\vdots$$
$$L_{K-1} d_{K-1} = \underbrace{-L_{K-1,K-2} d_{K-2} + Q_{K-1}^{-1} v_{K-1} + C_{K-1}^T R_{K-1}^{-1} y_{K-1}}_{q_{K-1}} - A_{K-1}^T Q_K^{-1} v_K$$
$$L_K d_K = \underbrace{-L_{K,K-1} d_{K-1} + Q_K^{-1} v_K + C_K^T R_K^{-1} y_K}_{q_K}$$

Similarly, solve d_0 → solve d_1 → solve d_2 ⋮ $\quad O(N(K+1))$

backward

- $L^T\hat{x} = d$

$$L_K^T \hat{x}_K = d_K$$
$$L_{K-1}^T \hat{x}_{K-1} = -L_{K,K-1}^T \hat{x}_K + d_{K-1}$$
$$\vdots$$
$$L_1^T \hat{x}_1 = -L_{21}^T \hat{x}_2 + d_1$$
$$L_0^T \hat{x}_0 = -L_{10}^T \hat{x}_1 + d_0$$

Similarly: solve $\hat{x}_K$ → solve $\hat{x}_{K-1}$ → solve $\hat{x}_{K-2}$ ⋮ $\quad O(N(K+1))$

Forward

$$L_{K-1} L_{K-1}^T = I_{K-1} + A_{K-1}^T Q_K^{-1} A_{K-1}$$
$$L_{K-1} d_{K-1} = q_{K-1} - A_{K-1}^T Q_K^{-1} v_K$$
$$L_{K,K-1} L_{K-1}^T = -Q_K^{-1} A_{K-1}$$
$$I_K = -L_{K,K-1} L_{K,K-1}^T + Q_K^{-1} + C_K^T R_K^{-1} C_K$$
$$q_K = -L_{K,K-1} d_{K-1} + Q_K^{-1} v_K + C_K^T R_K^{-1} y_K$$

Backward

$$L_{K-1}^T \hat{x}_{K-1} = -L_{K,K-1}^T \hat{x}_K + d_{K-1}$$

Kalman filter

▲ solving $(H^TW^{-1}H)\hat{x} = H^TW^{-1}z$

• **Rauch-Tung-Striebel**

from above:

$$L_{k-1}L_{k-1}^T = I_{k-1} + A_{k-1}^TQ_k^{-1}A_{k-1}$$
$$L_{k,k-1}L_{k-1}^T = -Q_k^{-1}A_{k-1}$$
$$I_k = -L_{k,k-1}L_{k,k-1}^T + Q_k^{-1} + C_k^TR_k^{-1}C_k$$

$$\therefore\ L_{k,k-1} = -Q_k^{-1}A_{k-1}L_{k-1}^{-T}$$

$$L_{k,k-1}L_{k,k-1}^T = Q_k^{-1}A_{k-1}L_{k-1}^{-T}L_{k-1}^{-1}A_{k-1}^TQ_k^{-T}$$
$$= Q_k^{-1}A_{k-1}(L_{k-1}L_{k-1}^T)^{-1}A_{k-1}^TQ_k^{-T}$$

$$\therefore\ I_k = -L_{k,k-1}L_{k,k-1}^T + Q_k^{-1} + C_k^TR_k^{-1}C_k$$
$$= Q_k^{-1} - Q_k^{-1}A_{k-1}(I_{k-1} + A_{k-1}^TQ_k^{-1}A_{k-1})A_{k-1}^TQ_k^{-T} + C_k^TR_k^{-1}C_k$$
$$= \underbrace{(A_{k-1}I_{k-1}^{-1}A_{k-1}^T + Q_k)^{-1}}_{\text{from Sherman-Morrison-Woodbury}} + C_k^TR_k^{-1}C_k$$

$$I_k = (A_{k-1}I_{k-1}^{-1}A_{k-1}^T + Q_k)^{-1} + C_k^TR_k^{-1}C_k$$

$$\hat{P}_{k,f} = I_k^{-1}$$
$$\check{P}_{k,f} = A_{k-1}I_{k-1}^{-1}A_{k-1}^TQ_k$$

$$\hat{P}_{k,f}^{-1} = \check{P}_{k,f}^{-1} + C_k^TR_k^{-1}C_k$$

▲ inducing KF

let us define

$$K_k = \hat{P}_{k,f}C_k^TR_k^{-1}$$
$$= (\check{P}_{k,f}^{-1} + C_k^TR_k^{-1}C_k)^{-1}C_k^TR_k^{-1}$$
SMW
$$= \underbrace{\check{P}_{k,f}C_k^T(C_k\check{P}_{k,f}C_k^T + R_k)^{-1}}_{\text{which is the Kalman gain that we are familiar of}}$$

- $$\check{P}_{k,f}^{-1} = \hat{P}_{k,f}^{-1} - C_k^TR_k^{-1}C_k$$
$$= \hat{P}_{k,f}^{-1}[1 - \hat{P}_{k,f}C_k^TR_k^{-1}C_k)$$
$$= \hat{P}_{k,f}^{-1}(1 - K_kC_k)$$
$$\Rightarrow \hat{P}_{k,f}\ (1 - K_kC_k)\check{P}_{k,f}$$

- now,

from above,

$$\begin{cases} L_{k,k-1}L_{k-1}^T = -Q_k^{-1}A_{k-1} \\ L_{k-1}d_{k-1} = q_{k-1} - A_{k-1}^TQ_k^{-1}v_k \\ L_{k-1}L_{k-1}^T = I_{k-1} + A_{k-1}^TQ_k^{-1}A_{k-1} \end{cases}$$

$$L_{k,k-1} = (-Q_k^{-1}A_{k-1})L_{k-1}^{-T}$$
$$d_{k-1} = L_{k-1}^{-1}(q_{k-1} - A_{k-1}^TQ_k^{-1}v_k)$$

$$\Rightarrow L_{k,k-1}d_{k-1} = (-Q_k^{-1}A_{k-1})L_{k-1}^{-T}L_{k-1}^{-1} \times (q_{k-1} - A_{k-1}^TQ_k^{-1}v_k)$$
$$\Rightarrow L_{k,k-1}d_{k-1} = (-Q_k^{-1}A_{k-1})(L_{k-1}L_{k-1}^T)^{-1}(q_{k-1} - A_{k-1}^TQ_k^{-1}v_k) \quad (1)$$
$$\because\ q_k = -L_{k,k-1}d_{k-1} + Q_k^{-1}v_k + C_k^TR_k^{-1}y_k \quad (2)$$

combining (1) & (2)

$$q_k = \underline{Q_k^{-1}A_{k-1}(I_{k-1} + A_{k-1}^TQ_k^{-1}A_{k-1})^{-1}q_{k-1}}$$
$$+ \underline{(Q_k^{-1} - Q_k^{-1}A_{k-1}(I_{k-1} + A_{k-1}^TQ_k^{-1}A_{k-1})^{-1}A_{k-1}^TQ_k^{-1})}v_k$$
$$+ C_k^TR_k^{-1}y_k$$
thru SMW
$$= (A_{k-1}I_{k-1}^{-1}A_{k-1}^T + Q_k)^{-1}A_{k-1}I_{k-1}^{-1}\ q_{k-1}$$
$$+ (A_{k-1}I_{k-1}^{-1}A_{k-1}^T + Q_k)^{-1}v_k$$
$$+ C_k^TR_k^{-1}y_k$$
$$= (A_{k-1}I_{k-1}^{-1}A_{k-1}^T + Q_k)^{-1}(A_{k-1}I_{k-1}^{-1}q_{k-1} + v_k)$$
$$+ C_k^TR_k^{-1}y_k$$
$$= (\check{P}_{k,f})^{-1}(A_{k-1}\hat{P}_{k-1,f}\hat{P}_{k-1,f}^{-1}\hat{x}_{k-1,f} + v_k)$$
$$+ C_k^TR_k^{-1}y_k$$
let $q_k = \hat{P}_{k,f}^{-1}\hat{x}_{k,f}$

$$\Rightarrow \hat{P}_{k,f}^{-1}\hat{x}_{k,f} = \check{P}_{k,f}^{-1}(A_{k-1}\hat{x}_{k-1,f} + v_k) + C_k^TR_k^{-1}y_k$$
$$\Rightarrow \hat{P}_{k,f}^{-1}\hat{x}_{k,f} = \check{P}_{k,f}^{-1}\check{x}_{k,f} + C_k^TR_k^{-1}y_k$$
$$\Rightarrow \hat{x}_{k,f} = \hat{P}_{k,f}\check{P}_{k,f}\check{x}_{k,f} + \hat{P}_{k,f}C_k^TR_k^{-1}y_k$$
$$= (1 - K_kC_k)\check{x}_{k,f} + K_ky_k$$
$$\Rightarrow \hat{x}_{k,f} = \check{x}_{k,f} + K_k(y_k - C_k\check{x}_{k,f})$$

end of "forward"

• backward

from above

$$L_{k-1}^T\hat{x}_{k-1} = -L_{k,k-1}^T\hat{x}_k + d_{k-1}$$
$$\Rightarrow \hat{x}_{k-1} = (L_{k-1}^T)^{-1}(-L_{k,k-1}^T\hat{x}_k + d_{k-1})$$
$$= L_{k-1}^{-T}L_{k-1}^{-1}L_{k-1}(-L_{k,k-1}^T\hat{x}_k + d_{k-1})$$
$$= (L_{k-1}L_{k-1}^T)^{-1}L_{k-1}(-L_{k,k-1}^T\hat{x}_k + d_{k-1})$$

$\Rightarrow$ from

$$L_{k-1}L_{k-1}^T = I_{k-1} + A_{k-1}^TQ_k^{-1}A_{k-1}$$
$$L_{k-1}d_{k-1} = q_{k-1} - A_{k-1}^TQ_k^{-1}v_k$$
$$L_{k,k-1}L_{k-1}^T = -Q_k^{-1}A_{k-1}$$

$$= (L_{k-1}L_{k-1}^T)^{-1}(-L_{k-1}L_{k,k-1}^T\hat{x}_k + L_{k-1}d_{k-1})$$
$$= (I_{k-1} + A_{k-1}^TQ_k^{-1}A_{k-1})^{-1} \times (-(L_{k,k-1}L_{k-1}^T)^T\hat{x}_k + q_{k-1} - A_{k-1}^TQ_k^{-1}v_k)$$
$$= (I_{k-1} + A_{k-1}^TQ_k^{-1}A_{k-1})^{-1} \times (A_{k-1}^TQ_k^{-1}\hat{x}_k + q_{k-1} - A_{k-1}^TQ_k^{-1}v_k)$$
$$= (I_{k-1} + A_{k-1}^TQ_k^{-1}A_{k-1})^{-1} \times (A_{k-1}^TQ_k^{-1}(\hat{x}_k - v_k) + q_{k-1})$$
$$= \underline{(I_{k-1} + A_{k-1}^TQ_k^{-1}A_{k-1})^{-1}A_{k-1}^TQ_k^{-1}}(\hat{x}_k - v_k)$$
$$+ \underline{(I_{k-1} + A_{k-1}^TQ_k^{-1}A_{k-1})^{-1}}q_{k-1}$$

$$\underline{(I_{k-1} + A_{k-1}^TQ_k^{-1}A_{k-1})^{-1}A_{k-1}^TQ_k^{-1}}(\hat{x}_k - v_k)$$
$$+ \underline{(I_{k-1} + A_{k-1}^TQ_k^{-1}A_{k-1})^{-1}}q_{k-1}$$
thru SMW

$$= I_{k-1}^{-1}A_{k-1}^T(A_{k-1}I_{k-1}^{-1}A_{k-1}^T + Q_k)^{-1}(\hat{x}_k - v_k)$$
$$+ (I_{k-1}^{-1} - I_{k-1}^{-1}A_{k-1}^T(A_{k-1}I_{k-1}^{-1}A_{k-1}^T + Q_k)^{-1}A_{k-1}I_{k-1}^{-1})q_{k-1}$$

$$= \hat{P}_{k-1,f}A_{k-1}^T\check{P}_{k,f}^{-1}(\hat{x}_k - v_k)$$
$$+ (\hat{P}_{k-1,f} - \hat{P}_{k-1,f}A_{k-1}^T\check{P}_{k,f}^{-1}A_{k-1}\hat{P}_{k-1,f})\hat{P}_{k-1,f}^{-1}\hat{x}_{k-1,f}$$
$$= \hat{P}_{k-1,f}A_{k-1}^T\check{P}_{k,f}^{-1}(\hat{x}_k - v_k)$$
$$+ \hat{x}_{k-1,f} - \hat{P}_{k-1,f}A_{k-1}^T\check{P}_{k,f}^{-1}A_{k-1}\cancel{\hat{P}_{k-1,f}\hat{P}_{k-1,f}^{-1}}\hat{x}_{k-1,f}$$
$$= \hat{P}_{k-1,f}A_{k-1}^T\check{P}_{k,f}^{-1}(\hat{x}_k - v_k)$$
$$+ \hat{x}_{k-1,f} - \hat{P}_{k-1,f}A_{k-1}^T\check{P}_{k,f}^{-1}\check{x}_{k,f}$$
$$= \hat{P}_{k-1,f}A_{k-1}^T\check{P}_{k,f}^{-1}(\hat{x}_k - \check{x}_{k,f} - v_k) + \hat{x}_{k-1,f}$$
$$= \hat{P}_{k-1,f}A_{k-1}^T\check{P}_{k,f}^{-1}(\hat{x}_k - \check{x}_{k,f}) + \hat{x}_{k-1,f}$$

Non-linear Estimations

Recall

J, cost

$$J_{v,k}(x) = \begin{cases} \frac{1}{2}(x_0 - \check{x}_0)^T\check{P}_0^{-1}(x_0 - \check{x}_0) \\ \frac{1}{2}(x_k - A_{k-1}x_{k-1} - v_k)^TQ_k^{-1}(x_k - A_{k-1}x_{k-1} - v_k) \end{cases}$$

$$J_{y,k}(x) = \frac{1}{2}(y_k - C_kx_k)^TR_k^{-1}(y_k - C_kx_k)$$

Mahalanobis distance

$$J(x) = \sum_{k=0}^{K}(J_{v,k}(x) + J_{y,k}(x))$$

$$\therefore\ x = \arg\min_x J(x)$$

• here the system, i.e.,
  - the prediction model
  - the measurement model

is no longer linear

▲
$$\begin{cases} e_{v,k}(x) = \begin{cases} \check{x}_0 - x_0 & ,k=0 \\ f(x_{k-1}, v_k, 0) - x_k & ,k=1\sim K \end{cases} \\ e_{y,k}(x) = y_k - g(x_k, 0), \quad k = 0\sim K \end{cases}$$

$$\begin{cases} J_{v,k}(x) = \frac{1}{2}e_{v,k}(x)^TW_{v,k}^{-1}e_{v,k}(x) \\ J_{y,k}(x) = \frac{1}{2}e_{y,k}(x)^TW_{y,k}^{-1}e_{y,k}(x) \end{cases}$$

$$\Rightarrow J(x) = \sum_{k=0}^{K}(J_{v,k}(x) + J_{y,k}(x))$$

$\Rightarrow$ • let
$$e(x) = \begin{bmatrix} e_v(x) \\ e_y(x) \end{bmatrix}$$
$$e_v(x) = \begin{bmatrix} e_{v,0}(x) \\ \vdots \\ e_{v,K}(x) \end{bmatrix}$$
$$e_y(x) = \begin{bmatrix} e_{y,0}(x) \\ \vdots \\ e_{y,K}(x) \end{bmatrix}$$

• let
$$W = \mathrm{diag}(W_v, W_y)$$
$$W_v = \mathrm{diag}(W_{v,0}, \cdots W_{v,K})$$
$$W_y = \mathrm{diag}(W_{y,0}, \cdots W_{y,K})$$

• $$J(x) = \frac{1}{2}e(x)^TW^{-1}e(x)$$

$\Rightarrow$ let
$$u(x) = Le(x)$$
$$L^TL = W^{-1}$$
$$J(x) = \frac{1}{2}e(x)^TL^TLe(x)$$
$$= \frac{1}{2}[Le(x)]^TLe(x)$$
$$= \frac{1}{2}u(x)^Tu(x)$$
$$\Rightarrow \hat{x} = \arg\min_x J(x)$$

Solving $\hat{x} = \arg\min_x J(x)$