RECURSIVE OPTIMAL ESTIMATORS

Kalman Filter

Discrete time sys.

$$x(k+1) = A(k)x(k) + B(k)U(k) + \tilde{F}(k)N'_k$$

$$Y(k+1) = C(k)x(k) + D(k)U(k) + \tilde{G}(k)N''_k$$

$x \in \mathbb{R}^n$, $U \in \mathbb{R}^p$, $Y \in \mathbb{R}^m$, $\{N\}$ sequences of random vecs
↳ deterministic inputs $\quad E\{N\} = 0$

Aim: estimate $x(j)$ from observations $\{Y(0), \dots Y(k)\}$

$$\begin{cases} \text{predicting} & j > k \\ \text{filtering} & j = k \\ \text{interpolating} & j < k \end{cases}$$

First step: split $x(k)$ into deterministic and stochastic

$$x(k) = \underbrace{x_d(k)}_{\text{depends on } U} + \underbrace{x_s(k)}_{\text{depends on } N}$$

$x_s(k) = x(k) - E\{x(k)\}$

$x_d(k) = E\{x(k)\}$

The evolution of x_d, x_s:

$$x_d(k+1) = E\{x(k+1)\} = E\{A(k)x(k) + B(k)U(k) + \tilde{F}(k)N'_k\}$$
$$= A(k)x_d(k) + B(k)U(k)$$

$$x_d(0) = E\{x(0)\}$$

$$x_s(k+1) = x(k+1) - E\{x(k+1)\} = A(k)x_s(k) + \tilde{F}(k)N'_k$$
$$x_s(0) = x(0) - E\{x(0)\}$$

Similarly for $Y(k) = Y_d(k) + Y_s(k)$

$$Y_d(k) = C(k)x_d(k) + D(k)U(k)$$

$$Y_s(k) = C(k)x_s(k) + \tilde{G}(k)N''_k \qquad E\{Y_s(k)\} = 0$$

It is sufficient to estimate $x_s(k)$ and sum with $x_d(k)$

Stochastic system:
$$\begin{cases} x_s(k+1) = A(k)x_s(k) + \tilde{F}(k)N'_k \\ x_s(0) = x(0) - E\{x(0)\} \\ Y_s(k) = C(k)x_s(k) + \tilde{G}(k)N''_k \end{cases}$$

We define $N_k = \begin{pmatrix} N_k' \\ N_k'' \end{pmatrix}$ $\quad F(k) = (\tilde{F}(k) \mid 0)$, $\quad G(k) = (0 \mid \tilde{G}(k))$

So that the previous sys becomes

$$\begin{cases} x_s(k+1) = A(k)\, x_s(k) + F(k) N_k \\ x_s(0) = x(0) - E\{x(0)\} \\ y_s(k) = C(k)\, x_s(k) + G(k) N_k \end{cases}$$

$\{N_k\}$ white and gaussian $\rightarrow \begin{cases} E\{N_k N_j^T\} = \delta_{kj} I \\ E\{N_k\} = 0 \end{cases} \quad \forall k, j$

$F N_k$, $G N_k$ are uncorrelated.

$(\Omega, \mathcal{F}, P)$ probability space,

$\mathcal{Y}_k^{y_s}$ the σ-algebra generated by $Y_{s,k} = \begin{pmatrix} y_s(0) \\ \vdots \\ y_s(k) \end{pmatrix}$

$\mathcal{Y}_k^{y_s} = \sigma(Y_{s,k}) \subset \mathcal{F}$, $\quad \mathcal{Y}_k^{y_s} \subseteq \mathcal{Y}_{k+1}^{y_s}$ (i.e. a filtration)

$\hat{x}_s(k) = f(Y_{s,k})$

$f(\cdot)$ can be such that to minimize the error variance

$J(f(Y_{s,k})) = E\{\| f(Y_{s,k}) - x_s(k) \|^2\}$

then $\boxed{\hat{x}_s(k) = \underbrace{E\{x_s(k) \mid \mathcal{Y}_k^{y_s}\}}}$

to compute this quantity we need a recursive solution (possible because N_k white gaussian)

Assumptions: $\{N_k\}$ white gaussian
$E\{N_k\} = 0$, $\quad E\{N_k N_j^T\} = \delta_{kj} I$, $\quad k \neq j$
$x(0)$ gaussian and independent from $\{N_k\}$ (uncorrelated)

Recursive expression of the covariance $\Psi_{x_s}(k)$

$$\begin{aligned}
\Psi_{x_s}(k+1) &= E\{x_s(k+1)\, x_s^T(k+1)\} = \\
&= E\{(A(k)x_s(k) + F(k)N_k)(A(k)x_s(k) + F(k)N_k)^T\} \\
&= A(k)\, E\{x_s(k)\, x_s^T(k)\} A^T(k) + A(k) E\{x_s(k) N_k^T\} F_k^T + \\
&\quad + F(k) E\{N_k\, x_s^T(k)\} A^T(k) + F(k) E\{N_k N_k^T\} F^T(k)
\end{aligned}$$

$E\{x_s(k) N_k^T\} = 0 \longrightarrow \boxed{\Psi_{x_s}(k+1) = A(k)\, \Psi_{x_s}(k) A^T(k) + F(k)\, F^T(k)}$

- Definition of innovation sequences

- State innovations

$\nu = \nu i \doteq \mathcal{Y}_k^{Y_s}$ - measurable random vec "innovative" contribution of Y in $X_s(k)$

$$V_s(0) = \hat{X}_s(0) = 0$$
$$V_s(K) = \hat{X}_s(K) - E\{\hat{X}_s(K) \mid \mathcal{Y}_{K-1}^{Y_s}\}$$

- Output innovations

$$\mu_s(0) = \hat{Y}_s(0) = 0$$
$$\mu_s(K) = \hat{Y}_s(K) - E\{\hat{Y}_s(K) \mid \mathcal{Y}_{K-1}^{Y_s}\}$$

<u>One step prediction</u>: optimal estimate of $X_s(K)$ given the observations $Y_s(K), \dots, Y_s(K-1)$

$$E\{\hat{X}_s(K) \mid \mathcal{Y}_{K-1}^{Y_s}\} = E\{E\{X_s(K) \mid \mathcal{Y}_K^{Y_s}\} \mid \mathcal{Y}_{K-1}^{Y_s}\}$$

$\mathcal{Y}_{K-1}^{Y_s}$ - measurable

$$= E\{X_s(K) \mid \mathcal{Y}_{K-1}^{Y}\} = \boxed{\hat{X}_s(K \mid K-1)}$$

Since $Y_s(K)$ is $\mathcal{Y}_K^{Y_s}$-measurable $\to$ $\boxed{\hat{Y}_s(K) = Y_s(K)}$

Therefore:

$$V_s(K) = \hat{X}_s(K) - E\{X_s(K) \mid \mathcal{Y}_{K-1}^{Y_s}\}$$
$$\mu_s(K) = \hat{Y}_s(K) - E\{Y_s(K) \mid \mathcal{Y}_{K-1}^{Y_s}\}$$

<u>Fact 1</u>: $E\{V_s(K)\} = 0$

$E\{V_s(K) V_s^T(j)\} = 0 \quad K \neq j$

<u>Proof</u>: $$E\{V_s(K)\} = E\{\underbrace{E\{X_s(K) \mid \mathcal{Y}_K^{Y_s}\}}_{\hat{X}_s(K)}\} - E\{E\{X_s(K) \mid \mathcal{Y}_{K-1}^{Y}\}\}$$

$$= E\{X_s(K)\} - E\{X_s(K)\} = 0 \quad ✓$$

For the second we assume $K < j$, therefore $\mathcal{Y}_K^{Y_s} \subseteq \mathcal{Y}_j^{Y_s}$:

$$E\{V_s(K) V_s^T(j)\} = E\{E\{V_s(K) V_s^T(j) \mid \mathcal{Y}_K^{Y_s}\}\}$$
$$= E\{V_s(K) \, E\{V_s^T(j) \mid \mathcal{Y}_K^{Y_s}\}\}$$

Since $\mathcal{Y}_K^{Y_s} \subseteq \mathcal{Y}_{j-1}^{Y_s}$

$$E\{V_s(j) \mid \mathcal{Y}_K^{Y_s}\} = E\{\hat{X}_s(j) - E\{X_s(j) \mid \mathcal{Y}_{j-1}^{Y_s}\} \mid \mathcal{Y}_K^{Y_s}\} =$$
$$= E\{E\{X_s(j) \mid \mathcal{Y}_j^{Y_s}\} - E\{X_s(j) \mid \mathcal{Y}_{j-1}^{Y_s}\} \mid \mathcal{Y}_K^{Y_s}\} =$$
$$= E\{X_s(j) \mid \mathcal{Y}_K^{Y_s}\} - E\{X_s(j) \mid \mathcal{Y}_K^{Y_s}\} = 0 \quad ✓$$

Fact 2 $\quad E\{\nu_s(k)\}=0$

$E\{\nu_s(k)\,\nu_s^T(j)\}=0 \quad \forall k\neq j$

Fact 3 $\quad E\{v_s(k)\,\nu_s^T(j)\}=0 \quad \forall k\neq j$

Proof : $v_s(k)$ is $\mathcal{Y}_k^{Y_s}$-measurable , $\nu_s(k)$ is $\mathcal{Y}_j^{Y_s}$-measurable

For $k>j$, since $\mathcal{Y}_j^{Y_s}\subseteq\mathcal{Y}_{k-1}^{Y_s}\subseteq\mathcal{Y}_k^{Y_s}$

$$E\{v(k)\cdot\nu^T(j)\}=E\{E\{v(k)\,\nu^T(j)\,|\,\mathcal{Y}_j^{Y_s}\}\}=$$

$$=E\{\underbrace{E\{v_s(k)\,|\,\mathcal{Y}_j^{Y_s}\}}\,\mu^T(j)\}$$

same of before substituting j and k

Fact 4 : $v_s(k)=\hat{X}_s(k)-A(k-1)\,\hat{X}_s(k-1)$

Therefore $\hat{x}_s(k|k-1)=A(k-1)\,\hat{x}_s(k-1)$ (optimal one step prediction)

to be compared with $X_s(k)=A(k-1)X_s(k-1)+F(k-1)N_k$

Proof : $v_s(k)=\hat{x}_s(k)-E\{x_s(k)\,|\,\mathcal{Y}_{k-1}^{Y_s}\}$

$$=\hat{x}_s(k)-E\{A(k-1)x_s(k-1)+F(k-1)N_{k-1}\,|\,\mathcal{Y}_{k-1}^{Y_s}\}$$

$$=\hat{x}_s(k)-A(k-1)\,\hat{x}_s(k-1)-\underbrace{E\{F(k-1)N_{k-1}\,|\,\mathcal{Y}_{k-1}^{Y_s}\}}$$

But

$$Y_s(k-1)=C(k-1)\underbrace{X_s(k-1)}_{A(k-2)X_s(k-2)+F(k-2)N_{k-2}}+G(k-1)N_{k-1}$$

$F(k-1)N_{k-1}$, $F(j-2)N_{j-2}$, $G(j-1)N_{j-1}$ are independent for $j\leq k$

$$\Rightarrow E\{F(k-1)N_{k-1}\,|\,\mathcal{Y}_{k-1}^{Y_s}\}=E\{F(k-1)N_{k-1}\}=0 \quad ✓$$

Fact 5 : $\nu_s(k)=Y_s(k)-C(k)\,A(k-1)\,\hat{X}_s(k-1)$

Proof . $\nu_s(k)=Y_s(k)-E\{Y_s(k)\,|\,\mathcal{Y}_{k-1}^{Y_s}\}$

$$=Y_s(k)-E\{C(k)X_s(k)+G(k)N_k\,|\,\mathcal{Y}_{k-1}^{Y_s}\}$$

$$=Y_s(k)-C(k)\hat{x}_s(k|k-1)-E\{G(k)N_k\,|\,\mathcal{Y}_{k-1}^{Y_s}\}=$$

$$=Y_s(k)-C(k)A(k-1)\hat{x}_s(k-1)$$

Fact 6 : $\{v_s(k)\}$, $\{\nu_s(k)\}$ are gaussian random vectors

Proof : $x_s(k)$ is a linear combination of gaus. vecs with zero mean
therefore is a gauss. vec.

$y_s(j)$ is gaussian with zero mean

$\hat{x}_s(k) = E\{x_s(k) | \mathcal{Y}_k^{y_s}\}$ affine function of the gaussian observations $y_s(k)$, therefore is gaussian

$\hat{x}_s(k|k-1) = E\{x_s(k) | \mathcal{Y}_{k-1}^{y_s}\}$ same

Same for $E\{y_s(k) | \mathcal{Y}_{k-1}^{y_s}\} \rightarrow$ $v_s(k)$, $\nu_s(k)$ are gaussian vecs

---

Fact 7 $\mathcal{Y}_k^{\nu_s} = \mathcal{Y}_k^{y_s}$ where $\mathcal{Y}_k^{\nu_s} = \sigma\{\nu_s(j) : 0 \le j \le k\}$

Proof : since $\nu_0(0) = y_s(0)$
$\nu_s(k) = y_s(k) - E\{y_s(k) | \mathcal{Y}_{k-1}^{y_s}\}$

It follows that $\nu_s(j)$, $0 \le j \le k$, is $\mathcal{Y}_k^{y_s}$-measurable and $\mathcal{Y}_j^{y_s} \subseteq \mathcal{Y}_k^{y_s}$

$\mathcal{Y}_k^{\nu_s}$ is the smallest σ-algebra for which $\nu_s(j)$ is $\mathcal{Y}_k^{\nu_s}$-measurable

Therefore $\mathcal{Y}_k^{\nu_s} \subseteq \mathcal{Y}_k^{y_s}$

And also : $\mathcal{Y}_k^{y_s} \subseteq \mathcal{Y}_k^{\nu_s}$

$y_s(j)$ is a measurable function of $\nu_s(i)$ $\forall i \le j$

$y_s(0) = \nu_s(0) \rightarrow \mathcal{Y}_0^{y_s} = \mathcal{Y}_0^{\nu_s}$

$y_s(1) = \nu_s(1) + E\{y_s(1) | \mathcal{Y}_0^{y_s}\} \rightarrow$ meas. fun. of $\nu_s(1)$, $\nu_s(0)$

$y_s(2) = \nu_s(2) + E\{y_s(2) \, \mathcal{Y}_1^{y_s}\} \rightarrow$ meas. fun. of $\nu_s(2)$, $\nu_s(1)$, $\nu_s(0)$

By induction $y_s(k)$ meas. fun. of $\nu_s(k), \dots, \nu_s(0)$

$\Rightarrow \mathcal{Y}_k^{y_s} \subseteq \mathcal{Y}_k^{\nu_s}$

- Fact 8 $\quad \hat{v}_s(k) = E\{v_s(k) \mid \mathcal{Y}_k^{Y_s}\} = v_s(k)$

$= E\{v_s(k) \mid \mathcal{Y}_k^{\mu_s}\}$ (from fact 7)

Since $\begin{bmatrix} v_s(k) \\ \mu_s(0) \\ \vdots \\ \mu_s(k) \end{bmatrix}$ is gaussian with 0 mean it follows that

$E\{\mu_s \mid \mathcal{Y}_k^{\mu_s}\}$ optimal estimate of μ_s given $\mathcal{Y}_k^{\mu_s}$ is a linear function of $\begin{bmatrix} \mu_s(0) \\ \vdots \\ \mu_s(k) \end{bmatrix}$:

$$E\{v_s(k) \mid \mathcal{Y}_k^{Y_s}\} = \Pi(k) \begin{bmatrix} \mu_s(0) \\ \vdots \\ \mu_s(k) \end{bmatrix}, \quad \Pi(k) = [\Pi_0(k) \mid \cdots \mid \Pi_k(k)]$$

with $\boxed{\Pi_j(k) = 0 \quad \forall j < k}$

Proof: $v_s(k) = E\{v_s(k) \mid \mathcal{Y}_k^{\mu_s}\} = \sum_{i=0}^{k} \Pi_i(k)\,\mu_s(i)$

post multiplying by $\mu_s^T(j)$, $j<k$

$$E\{v_s(k)\,\mu_s^T(j)\} = \sum_{i=0}^{k} \Pi_i(k)\, \underbrace{E\{\mu_s(i)\,\mu_s^T(j)\}}_{>0 \ v_s(j) \text{ is gaussian}}$$

Since $E\{v_s(k)\mu_s^T(j)\} = 0 \quad k \neq j$

$E\{\mu_s(i)\mu_s(j)\} = 0 \quad i \neq j$

therefore $\Pi_j(k) = 0 \quad \forall_j < k$

Fact 9 $\quad v_s(k) = \overbrace{\Pi_k(k)}^{\text{gain matrix}} \underbrace{\mu_s(k)}_{y_s(k) - C(k)A(k-1)\hat{x}_s(k-1)} = \hat{x}_s(k) - A(k-1)\hat{x}_s(k-1)$

Therefore

$$\hat{x}_s(k) = \overbrace{A(k-1)\hat{x}_s(k-1)}^{\hat{x}_s(k|k-1)} + \overbrace{\Pi_k(k)}^{\text{gain}}\Big(y_s(k) - \overbrace{C(k)A(k-1)\hat{x}_s(k-1)}^{\hat{x}_s(k|k-1)}\Big) \quad (*)$$

or

$$\hat{x}_s(k) = (I - \Pi_k(k)C(k))A(k-1)\hat{x}_s(k-1) + \Pi_k(k)y_s(k) \quad (**)$$

III Computation of $\Pi_k(k)$

$$v_s(k) = \Pi_k(k)\,\mu_s(k)$$

postmultiplying by $\mu_s^T(k)$

$$\underbrace{E\{v_s(k)\mu_s^T(k)\}}_{\Psi_{v_s\mu_s}} = \Pi_k(k)\underbrace{E\{\mu_s(k)\mu_s^T(k)\}}_{\Psi_{\mu_s}}$$

$$\boxed{\Pi_k(k) = \Psi_{v_s\mu_s}\Psi_{\mu_s}^{-1}}$$ we need to redefine this formula

We introduce $e_s(k) = x_s(k) - \hat{x}_s(k)$ estimation error

and we compute $E\{e_s(k)e_s^T(k)\}$

- Calculus of $E\{e_s(k)e_s^T(k)\}$ (estimation error covariance)

Using (**)

$$e_s(k) = x_s(k) - \hat{x}_s(k) =$$

$$= x_s(k) - [I - \Pi_k(k)C(k)]A(k-1)\hat{x}_s(k-1) - \Pi_k(k)y_s(k)$$

From sys. eq. :

$$\begin{cases} x_s(k) = A(k-1)x_s(k-1) + F(k-1)N_{k-1} \\ y_s(k) = C(k)A(k-1)x_s(k-1) + C(k)F(k-1)N_{k-1} \\ \qquad + G(k)N_k \end{cases}$$

$$e_s(k) = [I - \Pi_k(k)C(k)]A(k-1)e_s(k-1) \quad \text{I}$$

$$+ [I - \Pi_k(k)C(k)]F(k-1)N_{k-1} \quad \text{II}$$

$$- \Pi_k(k)G(k)N_k \quad \text{III}$$

I, II, III are independent in pairs $E\{\text{II}\ \text{III}^T\} = E\{\text{II}\}E\{\text{III}^T\} = 0$

example $E\{\text{II}\} = E\{\text{III}\} = 0$

Therefore:

$$\boxed{P(k)} = \{e_s(k)e_s^T(k)\} = E\{\text{I}\ \text{I}^T\} + E\{\text{II}\ \text{II}^T\} + E\{\text{III}\ \text{III}^T\}$$

$$= [I - \Pi_k(k)C(k)]\Lambda_p(k)[I - \Pi_k(k)C(k)]^T$$

$$+ \Pi_k G(k)G^T(k)\Pi_k^T$$

$$\Lambda_p(k) = A(k-1)\,P(k-1)\,A^T(k-1) + F(k-1)\,F^T(k-1)$$ error prediction covariance

$$= A(k-1)\,e_s(k) + F(k-1)\,N_{k-1}$$

- Calculus of $E\{\nu_s(k)\,\nu_s(k)^T\}$ and $E\{v_s(k)\,\nu_s(k)^T\}$

$$\boxed{\nu_s(k)} = Y_s(k) - E\{Y_s(k) \mid \mathcal{Y}^{Y_s}_{k-1}\} =$$

$$= C(k)\,x_s(k) + G(k)\,N_k - C(k)\,A(k-1)\,\hat{x}_s(k-1)$$

$$= C(k)\,A(k-1)\,e_s(k-1) + C(k)\,F(k-1)\,N_{k-1} + G(k)\,N_k$$

Since the three elements in the sum are independent

$$\boxed{E\{\nu_s(k)\,\nu_s^T(k)\}} = C(k)\,\Lambda_p(k)\,C^T(k) + G(k)\,G^T(k)$$

$$\boxed{v_s(k)} = \hat{x}_s(k) - E\{x_s(k) \mid \mathcal{Y}^{Y_s}_{k-1}\}$$

$$= \hat{x}_s(k) - A(k-1)\,\hat{x}_s(k-1) - x_s(k) + x_s(k)$$

$$= -e_s(k) + A(k-1)\,e_s(k-1) + \underbrace{F(k-1)\,N_{k-1}}_{E\{F(k-1)N_{k-1} \mid \mathcal{Y}^{Y_s}_{k-1}\} = 0}$$

$$E\{v_s(k)\,\nu_s^T(k)\} = -E\{e_s(k)\,\nu_s^T(k)\} \quad \mathrm{I}$$

$$+ E\{A(k-1)\,e_s(k-1)\,\nu_s^T(k)\} \quad \mathrm{II}$$

$$+ E\{F(k-1)\,N_{k-1}\,\nu_s^T(k)\} \quad \mathrm{III}$$

$$\mathrm{I} = E\{\underbrace{E\{e_s(k) \mid \mathcal{Y}^{Y_s}_{k}\}}_{E\{x_s \mid \mathcal{Y}^{Y_s}_k\} - E\{\hat{x}_s \mid \mathcal{Y}^{Y_s}_k\} = \hat{x}_s(k) - \hat{x}_s(k) = 0}\,\nu_s^T(k)\} = 0$$

$$\mathrm{II} = A(k-1)\,\underbrace{E\{e_s(k-1)\,e_s^T(k-1)\}}_{P(k-1)}\,A^T(k-1)\,C^T(k)$$

$$[I - \Pi_k C_k]\,\Lambda C_k^T - \Pi_k G G^T = 0$$

$$\mathrm{III} = F(k-1)\,F^T(k-1)\,C^T(k)$$

$$\boxed{E\{v_s(k)\,\nu_s^T(k)\} = \Lambda_p(k)\,C^T(k)}$$

$$\boxed{\Pi_k = \Lambda_p(k)\,C^T(k) \cdot \left[C(k)\,\Lambda_p(k)\,C^T(k) + G(k)\,G^T(k)\right]^{-1}}$$

$$P_k = [I - \Pi_k(k)\,C(k)]\,\Lambda_p(k)$$

Kalman filter equations

$$\hat{x}_s(k) = \underbrace{A(k-1)\,\hat{x}_s(k-1)}_{\hat{x}_s(k|k-1)} + \underbrace{\Pi_k(k)}_{\text{gain}}\Big(y_s(k) - C(k)\underbrace{A(k-1)\,\hat{x}_s(k-1)}_{\hat{x}_s(k|k-1)}\Big)$$

1) Initialization $k=0$

$\mathcal{Y}_0^{y_s} = \mathcal{Y}_m$: no observation at time 0

$$\hat{x}_s(0|-1) = E\{x_s(0) \mid \mathcal{Y}_{-1}^{y_s}\} = E\{x_s(0) \mid \mathcal{Y}_m\} = E\{x_s(0)\} = 0$$

$$\Lambda(0) = E\{e_{sp}(0)\,e_{sp}^T(0)\} = E\{(x_s(0) - x_s(0|-1))(x_s(0)-x_s(0|-1))^T\}$$
$$= E\{x_s(0)\,x_s^T(0)\} = \Psi_{x_s(0)}$$

2) Error precision covariance

$$\Lambda_p(k) = A(k-1)\,P(k-1)\,A^T(k-1) + F(k-1)\,F^T(k-1)$$

3) Gain matrix

$q \times q$ matrix

$$\Pi_k(k) = \Lambda_p(k)\,C^T(k) \cdot \left[C(k)\,\Lambda_p(k)\,C^T(k) + G(k)\,G^T(k)\right]^{-1}$$

4) Error covariance

$$P(k) = \left[I - \Pi_k\,C(k)\right]\Lambda_p(k)$$

5) Optimal prediction

$$\hat{x}_s(k|k-1) = A(k-1)\,\hat{x}_s(k-1)$$

6) Optimal estimate:

$$\hat{x}_s(k) = \hat{x}_s(k|k-1) + \Pi_k(k)\left(y_s(k) - C(k)\,\hat{x}_s(k|k-1)\right)$$

7) $k = k+1$. Go to step 2

The Kalman filter gives at each time k the optimal estimate of $x_s(k)$, its optimal prediction together with the error covariance

$\Pi_k(k)$ varies with time even if A, B, F, G are constant

$P(k)$ can be evaluated offline while $\hat{x}_s(k)$ only online

Optimality must be guaranteed also in step 1

Kalman predictor

For computing at each step the prediction. We have

$$\hat{x}_s(k+1|k) = A(k)\,\hat{x}_s(k) \,, \quad \text{with}$$

$$\hat{x}_s(k) = \hat{x}_s(k|k-1) + \Pi_k(k)\left(y_s(k) - C(k)\,\hat{x}_s(k|k-1)\right)$$

$$\bar{\Pi}_k(k) = A(k)\,\Pi_k(k) \quad \text{prediction gain}$$

$$P_k = \left(I - \Pi_k(k)\,C(k)\right)\Lambda_p(k)$$

$$\Lambda_p(k+1) = A(k)\left(I - \Pi_k(k)\,C(k)\right)\Lambda_p(k)A^T(k) + F(k)F^T(k)$$

$$= A(k)\Lambda_p(k)A^T(k) - \bar{\Pi}_k(k)\,C(k)\,\Lambda_p(k)A^T(k) + F(k)F^T(k)$$

Therefore

$$\bar{\Pi}_k(k) = A(k)\Lambda_p(k)C^T(k)\left(C(k)\Lambda_p(k)C^T(k) + G(k)G^T(k)\right)^{-1}$$

Steps:

1. Initialization $k=0$

$$\hat{x}_s(0|-1) = E\{x_s(0)\} = 0$$

$$\Lambda_p(0) = \Psi_{x_s}(0)$$

2. Predictor gain

$$\bar{\Pi}_k(k) = A(k)\,\Lambda_p(k)\,C^T(k)\cdot\left(C(k)\Lambda_p(k)C^T(k) + G(k)G^T(k)\right)^{-1}$$

3. Optimal prediction:

$$\hat{x}_s(k+1|k) = A(k)\,\hat{x}_s(k|k-1) + \bar{\Pi}_k(k)\cdot\left(y_s(k) - C(k)\,\hat{x}_s(k|k-1)\right)$$

4. Covariance

$$\Lambda_p(k+1) = A(k)\,\Lambda_p(k)\,A^T(k) - \bar{\Pi}_k(k)\,C(k)\,\Lambda_p(k)A^T(k) + F(k)F^T(k)$$

5. $k \to k+1$. Go to 2

Kalman filter with deterministic input

For optimal estimate $x(k) = x_s(k) + x_d(k)$

$$
\begin{cases}
x_d(k+1) = A(k)\, x_d(k) + B(k)\, U(k) \\
Y_d(k) = C(k)\, x_d(k) + D(k)\, U(k) \\
x_d(0) = E\{x(0)\} \\
x_s(k+1) = A(k)\, x_s(k) + F(k)\, N_k \\
Y_s(k) = C(k) \cdot x_s(k) + G(k)\, N_k \quad (= Y(k) - Y_d(k)) \\
x_s(0) = x(0) - E\{x(0)\}, \quad E\{x_s(0)\, x_s^T(0)\} = \psi_{x_0}
\end{cases}
$$

Kalman for $\hat{x}_s(k)$:

$$
\begin{cases}
\hat{x}_s(k+1) = A(k)\, \hat{x}_s(k) + \Pi_{k+1}(k+1) \cdot \left(Y_s(k+1) - C(k+1) A(k) \hat{x}_s(k)\right) \\
\hat{x}_s(0) = 0
\end{cases}
$$

we obtain with $\hat{x}(k) = \hat{x}_s(k) + x_d(k)$ and $E\{x_d + x_s \mid \mathcal{Y}_k^Y\} = x_d(k) + \hat{x}_s(k)$:

$$
\begin{aligned}
\hat{x}(k+1) &= A(k)\, \hat{x}(k) + B(k)\, U(k) \\
&\quad + \Pi_{k+1}(k+1)\Big(Y(k+1) - C(k+1)\, \underbrace{x_d(k+1)}_{A(k) x_d(k) + B(k) U(k)} - D(k+1)\, U(k+1) \\
&\qquad - C(k+1)\, A(k)\, \hat{x}_s(k)\Big) \\
&= A(k)\, \hat{x}(k) + B(k)\, U(k) \\
&\quad + \Pi_{k+1}(k+1)\Big(Y(k+1) - D(k+1)\, U(k+1) - C(k+1)\big(A(k)\hat{x}(k) + B(k) U(k)\big)\Big)
\end{aligned}
$$

$\hat{x}(0) = E\{x(0)\}$

$$\Rightarrow \hat{x}(k+1) = \hat{x}(k+1|k) + \Pi_{k+1}(k+1)\left(Y(k+1) - \hat{Y}(k+1|k)\right)$$

Where

$$\hat{x}(k+1|k) = E\{x(k+1) \mid \mathcal{Y}_k^Y\}$$

$$\hat{Y}(k+1|k) = E\{Y(k+1) \mid \mathcal{Y}_k^Y\}$$

Steady state Kalman filter

Kalman filter optimal if $\hat{x}_s(0)$, $P(0)$ and $\Psi_{x_s(0)} = \Psi_{x(0)}$.

It is NOT robust wrt initialization errors

Aim: conditions of $\{P(k)\}$ to have steady state value (insensitive wrt $P(0)$)

conditions of the filter to be AS (insensitive wrt $\hat{x}_s(0)$)

System:

$$\begin{cases} x_s(k+1) = A x_s(k) + F N_k \\ y_s(k) = C x_s(k) + G N_k \\ E\{x_s(0)\} = 0, \quad E\{x_s(0) x_s^T(0)\} = \Psi_{x(0)}, \quad F G^T = 0 \end{cases}$$

Kalman filter $\left(P(k) C^T = (I - \Pi_k(k) C) \Lambda_p(k) C^T = \Pi_k(k) \underbrace{G G^T}_{\text{non sing.}} \right.$

① Initialization

$\hat{x}_s(0) = 0$
$P(0) = \Psi_{x(0)}$
$k = 0$

② $\Lambda_p(k+1) = A P(k) A^T + F F^T$

$P(k+1) = \left(I + \Lambda_p(k+1) C^T (G G^T)^{-1} C \right)^{-1} \Lambda_p(k+1)$

③ Gain

$\Pi_{k+1}(k+1) = P(k+1) C^T (G G^T)^{-1}$

④ Optimal estimate

$\hat{x}_s(k+1) = A(k) \hat{x}_s(k) + \Pi_{k+1}(k+1) \left(y_s(k+1) - C A \hat{x}_s(k) \right)$

⑤ $k \to k+1$, go to step ②

- Steady state covariance

Steady state P_∞ and $\Lambda_{p\infty}$ of $P(k)$

$$\begin{cases} P_\infty = (I + \Lambda_{p\infty} C^T (GG^T)^{-1} C)^{-1} \Lambda_{p\infty} \\ \Lambda_{p\infty} = A P_\infty A^T + FF^T \end{cases}$$

$$P_\infty = [I + (A P_\infty A^T + FF^T) C^T (GG^T)^{-1} C]^{-1} (A P_\infty A^T + FF^T)$$

$$\Downarrow$$

$$0 = P_\infty + (A P_\infty A^T + FF^T)(C^T (GG^T)^{-1} C P_\infty - I)$$ Riccati eq. I

And with Kalman filter procedure

$$\Lambda_{p\infty} = A P_\infty A^T + FF^T$$

$$P_\infty = (I - K_\infty C) \Lambda_{p\infty}$$

$$K_\infty = \Lambda_{p\infty} C^T (C \Lambda_{p\infty} C^T (C \Lambda_{p\infty} C^T + GG^T)^{-1}$$

$$0 = P_\infty - A P_\infty A^T - FF^T + $$
$$+ (A P_\infty A^T + FF^T) C^T \cdot (C (A P_\infty A^T + FF^T) C^T + GG^T)^{-1} C \cdot$$
$$\cdot (A P_\infty A^T + FF^T)$$ Riccati eq. II (with symmetry)

Assume $GG^T = I$ with G full row rank

The SS Kalman filter in step 2 is:

$$P(k+1) = (I + \Lambda_p(k) C^T C)^{-1} \Lambda_p(k+1)$$

- Non singularity of $I + \Lambda_p(k) C^T C$

Since $P(k) \geq 0 \quad \forall k \geq 0$

$$\Lambda_p(k+1) = A P(k) A^T + FF^T \geq 0$$

Then $I + \underbrace{\Lambda_p(k+1)}_{\geq 0} \underbrace{C^T C}_{\geq 0} > 0 \quad \forall k \geq -1$, non singular

Positive definiteness of $P(k)$ covariance

When is $P(k) > 0 \quad \forall k \geq 0$?

Fact 1 : If (A,F) reachable pair $\left(\text{rank}\left[F \vdots AF \vdots \cdots \vdots A^{n-1}F\right] = n\right)$

$$P(k) > 0 \Rightarrow P(k+1) > 0 \qquad \forall k \geq 0$$

Proof : $P(k+1)$ non sing. iff $\Lambda_p(k+1)$ nonsing $\left(\begin{matrix}\text{previous}\\ \text{subsection}\end{matrix}\right)$

The nonsingularity of $\Lambda_p(k) = AP(k)A^T + FF^T$ follows because $P(k) > 0$ and $\text{rank}\left[F \vdots AF \vdots \cdots \vdots A^{n-1}F\right] = n$

To guarantee that $P(k) > 0 \Rightarrow P(k+1) > 0$ is sufficient that rank $[F \vdots A] = n$ (weaker than (A,F) reachable) or in alternative the Hautus test

$$\text{rank}\left[\lambda I - A \vdots F\right] = n \qquad \forall \lambda \in \sigma(A)$$

Fact 2 : If (A,F) is reachable pair (if rank $[F \vdots A] = n$) then:

$$P(k) > 0 \Rightarrow P(k+h) > 0 \qquad \begin{matrix}\forall h > 0\\ \forall k \geq 0\end{matrix}$$

If $k = 0 \rightarrow P(0) > 0 \Rightarrow P(k) > 0 \quad \forall k > 0$

Monotonicity of $\{P(k)\}$

Define $\Phi(P) = (I + \Psi(P)C^TC)^{-1}\Psi(P)$

$$\Psi(P) = APA^T + FF^T$$

The Kalman filter equations of step ② are

$$P(k+1) = \Phi(P(k)) \qquad \text{with} \quad P_\infty = \Phi(P_\infty)$$

Fact 3 : P, Q symm and pos. semidef.

$$\Phi(P+Q) \geq \Phi(P)$$

Proof : $\Phi(P+Q) - \Phi(P) = \int_0^1 \frac{d}{d\lambda}\Phi(P+\lambda Q)\,d\lambda$

if $S(\lambda) = I + \Psi(P+\lambda Q)C^TC$:

$$\begin{cases} \frac{\partial S}{\partial \lambda} = AQA^TC^TC \\ \frac{\partial S^{-1}}{\partial \lambda} = -S^{-1}(\lambda)\left(\frac{\partial S}{\partial \lambda}\right)S^{-1}(\lambda) \end{cases}$$

$$\frac{\partial}{\partial \lambda}\Phi(P+\lambda Q) = \frac{\partial}{\partial \lambda}\left(S^{-1}(\lambda)\Psi(P+\lambda Q)\right) = S^{-1}AQA^TS^{-T} \geq 0$$

The consequence is that if $\exists\, k_0 : P(k_0+1) \geq P(k_0)$

$$\Rightarrow \underbrace{\Phi(P(k_0+1))}_{P(k_0+2)} \geq \underbrace{\Phi(P(k_0))}_{P(k_0+1)}$$

$$\Rightarrow P(k+1) \geq P(k), \quad \forall k \geq k_0$$

Similarly, if $\exists\, k_0 : P(k_0+1) \leq P(k_0)$

$$\Rightarrow P(k+1) \leq P(k), \quad \forall k \geq k_0$$

To know if $\{P(k)\}$ monotone nondecreasing / nonincreasing it is sufficient to see if $P(0) \leq P(1)$ / $P(0) \geq P(1)$

If $P(0)=0$ $\rightarrow$ nondecreasing sequence:

$$P(0)=0 \Rightarrow \Lambda_P(0) = FF^T \geq 0$$

$$\Rightarrow P(1) = (I + FF^T C^T C)^{-1} FF^T \geq 0 \Rightarrow P(1) \geq P(0) = 0$$

- **Boundedness of $\{P(k)\}$**

Fact 4 if (C, A) detectable, $\forall P(0) \geq 0\ \exists \{S(k)\}$ of positive semidefinite matrices s.t.

$$P(k) \leq S(k) \quad \forall k \geq 0 \quad \text{and}$$

$$\lim_{k \to \infty} S(k) = S_\infty \quad \text{with } S_\infty \text{ depending on } P(0)$$

Even in case of both $\{P(k)\}$ nondecreasing or nonincreasing $\lim_{k \to +\infty} P(k) = P_\infty \geq 0$. However P_∞ may depend on $P(0)$.
Under which conditions P_∞ is unique?

Fact 5 Let $P_\infty \geq 0$ symm solution of Riccati eq. I

$$0 = P_\infty + (AP_\infty A^T + FF^T)(C^T(GG^T)^{-1} C P_\infty - I)$$

If (F^T, A^T) detectable, P_∞ is unique and

$$\sigma\big((I - P_\infty C^T C)A\big) \subset S^1 \rightarrow \text{all the eigs in the unit circle} \rightarrow \text{AS}$$

Steady state Kalman filter equations

If (C, A) and (F^T, A^T) detectable:

1) Uniqueness of steady state error covariance:
$\exists$ unique symm $P_\infty \geq 0$ of the RE and $\lim_{k \to +\infty} P(k) = P_\infty$

2) Stability: $\sigma((I - P_\infty C^T C)A) \subset S^1$

3) Asymptotic optimality of SSKF $\hat{z}_s(k+1)$:
$$\hat{z}_s(k+1) = A\hat{z}_s(k) + P_\infty C^T (y_s(k+1) - CA\hat{z}_s(k))$$
optimal in the sense that if $P^{\hat{z}_s}(k) = E\{(\hat{z}_s(k) - x_s(k))(\hat{z}_s - x_s)^T\}$
then:
$$\lim_{k \to +\infty} P^{\hat{z}_s}(k) = P_\infty$$

<u>Proof</u>: 1 & 2 follow from facts 4 & 5 and monotonicity of $\{P(k)\}$

We have only to prove $\lim_{k \to +\infty} P^{\hat{z}_s}(k) = P_\infty$

If $K_\infty = P_\infty C^T$,

$$x_s(k+1) - \hat{z}_s(k+1) = (I - K_\infty C)(A(x_s(k) - \hat{z}_s(k)) + FN_k) - K_\infty G N_{k+1}$$
$$= (\mathrm{I}) + (\mathrm{II}) + (\mathrm{III}) : \text{pairwise independent}$$

Therefore
$$P^{\hat{z}_s}(k+1) = (I - K_\infty C)(AP^{\hat{z}_s}(k)A^T + FF^T) \cdot (I - K_\infty C)^T + K_\infty G G^T K_\infty^T$$

But $P_\infty \geq 0$ is solution of:

$$0 = P_\infty - (I - K_\infty C)(AP_\infty A^T + FF^T)(I - K_\infty C)^T + K_\infty G G^T K_\infty^T$$

Subtracting the two eqs. above:
$$P^{\hat{z}_s}(k+1) - P_\infty = [(I - K_\infty C)A] \cdot (P^{\hat{z}_s}(k) - P_\infty) \cdot [(I - K_\infty C)A]^T$$
solving backwards
$$P^{\hat{z}_s}(k) - P_\infty = [(I - K_\infty C)A]^k \cdot (P^{\hat{z}_s}(0) - P_\infty) \cdot ([(I - K_\infty C)A]^k)^T$$
By 2 since $\sigma((I - K_\infty C)A) \subset S^1$:
$\exists\ N > 0,\ \lambda \in (0,1)$: $\|P^{\hat{z}_s} - P_\infty\| \leq$ $\longrightarrow$ implies $\lim_{k \to +\infty} P^{\hat{z}_s} = P_\infty$

$$\leq \|[(I - K_\infty C)A^k]\| \cdot \|P^{\hat{z}_s}(0) - P_\infty\| \cdot \|([I - K_\infty C)A^k]^T\| \leq N^2 \lambda^{2k} \|P^{\hat{z}_s}(0) - P_\infty\|$$