

Filtrado de Kalman

86.51

Cecilia G. Galarza

FIUBA

Laboratorio de Procesamiento de Señales y Comunicaciones

2do Cuatrimestre 2022

- Usando las ideas de estimación de variables aleatorias a partir de observaciones correlacionadas, obtuvimos el filtro de Wiener, que resuelve el problema óptimo MMSE.
- Un criterio alternativo de optimización es el de Cuadrados Mínimos (LS)
- El RLS implementa una versión iterativa de la solución LS.
- En principio, todos estos métodos resuelven problemas estacionarios.
- Para trabajar con procesos no-estacionarios, se introducen parámetros heurísticos (constante de olvido) de difícil calibrado

Planteo del problema

Considere los procesos en tiempo discreto $\mathbf{x}(k) \in \mathbb{C}^n$, $\mathbf{y}(k) \in \mathbb{C}^p$, $\mathbf{v}(k) \in \mathbb{C}^l$, $\mathbf{w}(k) \in \mathbb{C}^p$, conocidos como *estado*, *observación*, *ruido de estado*, y *ruido de medición* del sistema \mathcal{F} descrito por el modelo lineal variante en el tiempo

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{F}_k \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k) \\ \mathbf{y}(k) &= \mathbf{H}_k \mathbf{x}(k) + \mathbf{w}(k).\end{aligned}$$

Asuma que $\mathbf{x}(0)$ también es un vector aleatorio. Luego, para cada instante k , el filtro de Kalman calcula

$$\hat{\mathbf{x}}[k|k-1] = \mathbb{E} [\mathbf{x}(k) \mid \mathbf{y}(k-1), \dots, \mathbf{y}(0)]$$

o en forma equivalente, resuelve el problema de estimación MMSE

$$\min \mathbb{E} [\|\mathbf{x}(k) - \hat{\mathbf{x}}(k)\|^2 \mid \mathbf{y}(k-1), \dots, \mathbf{y}(0)]$$

Todo sistema dinámico tiene asociado un *estado*, generalmente llamado $\mathbf{x}(k)$. En todo instante k , el vector $\mathbf{x}(k)$ contiene toda la información necesaria para poder determinar el futuro comportamiento del sistema si se conocen las entradas.

En particular, si el sistema tiene una dinámica lineal y evoluciona en tiempo discreto, podemos plantear el modelo

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{F}_k \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k) \\ \mathbf{y}(k) &= \mathbf{H}_k \mathbf{x}(k) + \mathbf{w}(k) \quad k = 0, 1, \dots\end{aligned}\tag{1}$$

donde $\mathbf{x}(k) \in \mathbb{C}^n$ es el estado del sistema en el instante k , $\mathbf{y}(k)$ es la observación o salida del sistema, y $\mathbf{v}(k)$ y $\mathbf{w}(k)$ son dos procesos ruidosos que afectan la dinámica del estado y de la observación respectivamente. Las matrices $\mathbf{F}_k \in \mathbb{C}^{n \times n}$, $\mathbf{G}_k \in \mathbb{C}^{n \times l}$, $\mathbf{H}_k \in \mathbb{C}^{p \times n}$ varían con el tiempo de un modo conocido¹. Las ecuaciones (1) son un modelo de estados de un sistema de dimensión finita.

¹En [3] podrán encontrar un tratamiento metódico de las distintas representaciones en variables de estado.

- Sistema ARMA en tiempo discreto:

$$y(k) = a_1 y(k-1) + a_2 y(k-2) + b_1 v(k-1)$$

Definimos el vector de estado $\mathbf{x} = [x_1 \ x_2 \ x_3]^t$ de modo que para todo k

$$x_1(k) = y(k-1) \quad x_2(k) = y(k-2) \quad x_3(k) = v(k-1)$$

Luego,

$$\underbrace{\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix}}_{\mathbf{x}(k+1)} = \underbrace{\begin{bmatrix} a_1 & a_2 & b_1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}}_{\mathbf{F}_k=\mathbf{F}} \underbrace{\begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}}_{\mathbf{x}(k)} + \underbrace{\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}}_{\mathbf{G}_k=\mathbf{G}} v(k)$$
$$y(k) = \underbrace{\begin{bmatrix} a_1 & a_2 & b_1 \end{bmatrix}}_{\mathbf{H}_k=\mathbf{H}} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}$$

- Sistema en tiempo continuo

$$\ddot{y}(t) = av(t)$$

Consideramos el siguiente vector de estados: $\mathbf{x} = [x_1 \ x_2]$ tal que

$$x_1(t) = y(t) \quad x_2(t) = \dot{y}(t)$$

Luego, el modelo de estado en tiempo continuo resulta

$$\dot{x}_1(t) = \dot{y}(t) = x_2(t)$$

$$\dot{x}_2(t) = \ddot{y}(t) = av(t)$$

$$\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} 0 \\ a \end{bmatrix} v(t)$$

$$y(t) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}$$

Aproximación en tiempo discreto $\dot{x}(t) \simeq \frac{x[(k+1)T] - x[kT]}{T}$. Luego,

$$\dot{x}_1(t) \simeq \frac{x_1[(k+1)T] - x_1[kT]}{T} = x_2(kT)$$

$$\dot{x}_2(t) \simeq \frac{x_2[(k+1)T] - x_2[kT]}{T} = av(kT)$$

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 & T \\ 0 & 1 \end{bmatrix}}_{\mathbf{F}} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \underbrace{\begin{bmatrix} 0 \\ aT \end{bmatrix}}_{\mathbf{G}} v(k)$$

$$y(k) = \underbrace{\begin{bmatrix} 1 & 0 \end{bmatrix}}_{\mathbf{H}} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}$$

Algunos ejemplos (cont)

Vuelvo a sistema ARMA: $y(k) = a_1y(k-1) + a_2y(k-2) + b_1v(k-1)$

- $x_1(k) = y(k-1), x_2(k) = y(k-2), x_3(k) = v(k-1)$

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & b_1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} v(k)$$

$$y(k) = \begin{bmatrix} a_1 & a_2 & b_1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}$$

Algunos ejemplos (cont)

Vuelvo a sistema ARMA: $y(k) = a_1y(k-1) + a_2y(k-2) + b_1v(k-1)$

- $x_1(k) = y(k-1), x_2(k) = y(k-2), x_3(k) = v(k-1)$

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & b_1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} v(k)$$

$$y(k) = \begin{bmatrix} a_1 & a_2 & b_1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}$$

- $x_1(k) = y(k), x_2(k) = y(k-1)$

Algunos ejemplos (cont)

Vuelvo a sistema ARMA: $y(k) = a_1 y(k-1) + a_2 y(k-2) + b_1 v(k-1)$

- $x_1(k) = y(k-1), x_2(k) = y(k-2), x_3(k) = v(k-1)$

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & b_1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} v(k)$$
$$y(k) = \begin{bmatrix} a_1 & a_2 & b_1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}$$

- $x_1(k) = y(k), x_2(k) = y(k-1)$

$$x_1(k+1) = y(k+1) = a_1 \underbrace{y(k)}_{x_1(k)} + a_2 \underbrace{y(k-1)}_{x_2(k)} + b_1 v(k)$$
$$x_2(k+1) = y(k) = x_1(k)$$

Algunos ejemplos (cont)

Vuelvo a sistema ARMA: $y(k) = a_1 y(k-1) + a_2 y(k-2) + b_1 v(k-1)$

- $x_1(k) = y(k-1), x_2(k) = y(k-2), x_3(k) = v(k-1)$

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \\ x_3(k+1) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 & b_1 \\ 1 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} v(k)$$

$$y(k) = \begin{bmatrix} a_1 & a_2 & b_1 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \\ x_3(k) \end{bmatrix}$$

- $x_1(k) = y(k), x_2(k) = y(k-1)$

$$\begin{bmatrix} x_1(k+1) \\ x_2(k+1) \end{bmatrix} = \begin{bmatrix} a_1 & a_2 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix} + \begin{bmatrix} b_1 \\ 0 \end{bmatrix} v(k)$$

$$y(k) = \begin{bmatrix} 1 & 0 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}$$

Retomamos el modelo de estados en tiempo discreto y agregamos algunas hipótesis

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{F}_k \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k) \\ \mathbf{y}(k) &= \mathbf{H}_k \mathbf{x}(k) + \mathbf{w}(k) \quad k = 0, 1, \dots\end{aligned}$$

Asumimos que

- $\mathbf{v}(k)$ y $\mathbf{w}(k)$ son dos procesos de media nula, descorrelacionados en el tiempo, tal que

$$\mathbb{E} \left\{ \begin{bmatrix} \mathbf{v}(i) \\ \mathbf{w}(i) \end{bmatrix} \begin{bmatrix} \mathbf{v}(j)^* & \mathbf{w}(j)^* \end{bmatrix} \right\} = \begin{bmatrix} \mathbf{Q}_i & 0 \\ 0 & \mathbf{R}_i \end{bmatrix} \delta_{ij}$$

- El estado inicial $\mathbf{x}(0)$ es un vector aleatorio de media $\bar{\mathbf{x}}_0$ y matriz de covarianza \mathbf{P}_0 .
- $\mathbf{x}(0)$ está descorrelacionado con $\mathbf{v}(k)$ y $\mathbf{w}(k)$ en todo momento k .

En este tema, la elección de una buena notación es fundamental para evitar errores

- $\hat{\mathbf{x}}[k|k-1]$: estimación *a priori* de $\mathbf{x}(k)$. Es la estimación dadas todas las observaciones $\mathbf{y}(0), \dots, \mathbf{y}(k-1)$.
- $\hat{\mathbf{x}}[k|k]$: estimación *a posteriori* de $\mathbf{x}(k)$. Es la estimación dadas todas las observaciones $\mathbf{y}(0), \dots, \mathbf{y}(k)$.
- $\Sigma_{k|k-1} = \mathbb{E} \{ [\mathbf{x}(k) - \hat{\mathbf{x}}[k|k-1]] [\mathbf{x}(k) - \hat{\mathbf{x}}[k|k-1]]^* \mid \mathbf{y}(0), \dots, \mathbf{y}(k-1) \}$: covarianza *a priori*.

El procesamiento tiene dos etapas:

- 1 Incorporación de nueva observación: Cálculo de $\hat{\mathbf{x}}[k|k]$

$$\hat{\mathbf{x}}[k|k] = \hat{\mathbf{x}}[k|k-1] + \mathbf{K}_k (\mathbf{y}(k) - \mathbf{H}_k \hat{\mathbf{x}}[k|k-1]) \quad \hat{\mathbf{x}}_{0|-1} = \bar{\mathbf{x}}_0$$

$$\mathbf{K}_k = \Sigma_{k|k-1} \mathbf{H}_k^* [\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}_k]^{-1}$$

$$\Sigma_{k|k} = \Sigma_{k|k-1} - \Sigma_{k|k-1} \mathbf{H}_k^* [\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}_k]^{-1} \mathbf{H}_k \Sigma_{k|k-1} \quad \Sigma_{0|-1} = \mathbf{P}_0$$

- 2 Predicción: Cálculo de $\hat{\mathbf{x}}[k+1|k]$

$$\hat{\mathbf{x}}[k+1|k] = \mathbf{F}_k \hat{\mathbf{x}}[k|k]$$

$$\Sigma_{k+1|k} = \mathbf{F}_k \Sigma_{k|k} \mathbf{F}_k^* + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^*$$

Propiedades del Filtro de Kalman (KF))

- \mathbf{K}_k es la ganancia del filtro de Kalman.
- El KF es un sistema dinámico lineal de tiempo discreto, en general, variante en el tiempo.
- $\hat{\mathbf{x}}[k|k-1] = \mathbb{E}[\mathbf{x}(k)|\mathbf{y}(0), \dots, \mathbf{y}(k-1)]$ es la esperanza condicional de $\mathbf{x}(k)$ dadas las observaciones $\mathbf{y}(0), \dots, \mathbf{y}(k-1)$
- $\Sigma_{k|k-1} = \mathbb{E}[(\mathbf{x}(k) - \hat{\mathbf{x}}[k|k-1])(\mathbf{x}(k) - \hat{\mathbf{x}}[k|k-1])^* | \mathbf{y}(0), \dots, \mathbf{y}(k-1)]$ es la matriz de covarianza condicional del estado.
- $\Sigma_{k|k-1}$ y \mathbf{K}_k no dependen de las observaciones y pueden ser precomputadas.
- Para computar \mathbf{K}_k se debe calcular la inversa de $[\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}_k]$. Para garantizar que ésta exista, se impone $\mathbf{R}_k > 0$, es decir, se asume que en todo momento las observaciones son ruidosas.

Dadas dos variables aleatorias X , Y , cuál es el mejor estimador lineal \hat{X} tal que minimice $\mathbb{E}[\|X - \hat{X}\|^2]$?

Estimador LMMSE

Sean X e Y dos variables aleatorias tal que

$$\mathbb{E}\left\{\begin{bmatrix} X \\ Y \end{bmatrix}\right\} = \begin{bmatrix} m_X \\ m_Y \end{bmatrix} \quad \text{Cov}\left\{\begin{bmatrix} X \\ Y \end{bmatrix}\right\} = \begin{bmatrix} \Sigma_{XX} & \Sigma_{XY} \\ \Sigma_{YX} & \Sigma_{YY} \end{bmatrix}$$

Luego, el estimador lineal que minimiza $\mathbb{E}[\|X - \hat{X}\|^2]$ es

$${}^L\hat{X} = m_X + \Sigma_{XY}\Sigma_{YY}^{-1}(Y - m_Y)$$

Si X e Y son conjuntamente gaussianas, entonces ${}^L\hat{X}$ es también el estimador MMSE.

$${}^L\hat{X}(Y) = m_X + \Sigma_{XY}\Sigma_{YY}^{-1}(Y - m_Y)$$

- El LMMSE es insesgado, es decir

$$\mathbb{E} \left[{}^L\hat{X}(Y) \right] = \mathbb{E} [X]$$

- Principio de ortogonalidad: $X - {}^L\hat{X}$ es ortogonal a Y , es decir

$$\mathbb{E} \left[\left(X - {}^L\hat{X}(Y) \right) Y^* \right] = 0$$

- Sean Y_1, \dots, Y_k v.a. descorrelacionadas entre sí. Luego, el LMMSE de X a partir de Y_1, \dots, Y_k es igual a

$$\begin{aligned} {}^L\hat{X}(Y_1, \dots, Y_k) &= m_X + \sum_{i=1}^k \Sigma_{XY_i} \Sigma_{Y_i Y_i}^{-1} (Y_i - m_{Y_i}) \\ &= \sum_{i=1}^k {}^L\hat{X}(Y_i) - (k-1)m_X \end{aligned}$$

En particular, ${}^L\hat{X}(Y_1, Y_2) = {}^L\hat{X}(Y_1) + {}^L\hat{X}(Y_2) - m_X$

Definición a *mano alzada*: Dada la secuencia de variables aleatorias Z_0, \dots, Z_k , la innovación \tilde{Z}_k contiene la información en Z_k no contenida en Z_0, \dots, Z_{k-1} .

- Caso gaussiano:

$$\tilde{Z}_0 = Z_0 - \mathbb{E}[Z_0]$$

$$\tilde{Z}_1 = Z_1 - \mathbb{E}[Z_1|Z_0]$$

$$\tilde{Z}_2 = Z_2 - \mathbb{E}[Z_2|Z_0, Z_1]$$

$$\tilde{Z}_k = Z_k - \mathbb{E}[Z_k|Z_0, Z_1, \dots, Z_{k-1}]$$

- Caso general:

$$\tilde{Z}_0 = Z_0 - \mathbb{E}[Z_0]$$

$$\tilde{Z}_1 = Z_1 - {}^L\hat{Z}_1(Z_0)$$

$$\tilde{Z}_2 = Z_2 - {}^L\hat{Z}_2(Z_0, Z_1)$$

$$\tilde{Z}_k = Z_k - {}^L\hat{Z}_k(Z_0, \dots, Z_{k-1})$$

\tilde{Z}_k resulta el error del estimador LMMSE de Z_k dados Z_0, \dots, Z_{k-1} .

- $\{\tilde{Z}_0, \tilde{Z}_1, \dots, \tilde{Z}_k\}$ son v.a. de media nula descorrelacionadas entre sí.
- Si $\{Z_0, Z_1, \dots, Z_k\}$ son conjuntamente gaussianas, \tilde{Z}_k es gaussiana
- Por el principio de ortogonalidad, $\tilde{Z}_k \perp \{Z_0, Z_1, \dots, Z_{k-1}\}$
- \tilde{Z}_k es una función lineal de Z_0, Z_1, \dots, Z_k
- Z_k es una función lineal de $\tilde{Z}_0, \tilde{Z}_1, \dots, \tilde{Z}_k$
- Si X es una variable aleatoria de media nula,

$$\mathbb{E}[X|Z_0, Z_1, \dots, Z_k] = \mathbb{E}[X|\tilde{Z}_0, \tilde{Z}_1, \dots, \tilde{Z}_k] = \sum_{i=0}^k \mathbb{E}[X|\tilde{Z}_i]$$

El problema a resolver es la estimación del vector de estados $\mathbf{x}(k)$ a partir de las observaciones $\mathbf{y}(0), \dots, \mathbf{y}(k-1)$. Usando el criterio MMSE, calculamos

$$\hat{\mathbf{x}}[k|k-1] = \mathbb{E}[\mathbf{x}(k)|\mathbf{y}(0), \dots, \mathbf{y}(k-1)] \quad \text{caso gaussiano}$$

$$\hat{\mathbf{x}}[k|k-1] = {}^L\hat{\mathbf{x}}(\mathbf{y}(0), \dots, \mathbf{y}(k-1)) \quad \text{caso general}$$

Haciendo un abuso de notación, vamos a hacer todo el desarrollo con $\mathbb{E}[\mathbf{x}(k)|\mathbf{y}(0), \dots, \mathbf{y}(k-1)]$, teniendo en cuenta que para el caso general no-gaussiano el estimador corresponde al estimador LMMSE ${}^L\hat{\mathbf{x}}(\mathbf{y}(0), \dots, \mathbf{y}(k-1))$.

Vamos a trabajar con la secuencia de innovaciones de las mediciones

$$\begin{aligned}\tilde{\mathbf{y}}(k) &= \mathbf{y}(k) - \mathbb{E}[\mathbf{y}(k)|\mathbf{y}(0), \dots \mathbf{y}(k-1)] \\ &= \mathbf{y}(k) - \mathbb{E}[\mathbf{H}_k \mathbf{x}(k) + \mathbf{w}(k)|\mathbf{y}(0), \dots \mathbf{y}(k-1)] \\ &= \mathbf{y}(k) - \mathbf{H}_k \mathbb{E}[\mathbf{x}(k)|\mathbf{y}(0), \dots \mathbf{y}(k-1)] \\ &= \mathbf{y}(k) - \mathbf{H}_k \hat{\mathbf{x}}[k|k-1] \\ &= \mathbf{H}_k \mathbf{x}(k) + \mathbf{w}(k) - \mathbf{H}_k \hat{\mathbf{x}}[k|k-1] \\ &= \mathbf{H}_k \underbrace{(\mathbf{x}(k) - \hat{\mathbf{x}}[k|k-1])}_{\tilde{\mathbf{x}}(k)} + \mathbf{w}(k)\end{aligned}$$

y con la secuencia de innovaciones del vector de estados o error de estimación del estado

$$\tilde{\mathbf{x}}(k) = \mathbf{x}(k) - \hat{\mathbf{x}}[k|k-1]$$

Vamos a calcular $\hat{\mathbf{x}}[k+1|k]$ en forma recursiva. Para ello, observamos que

$$\begin{aligned}\mathbb{E}[\mathbf{x}(k+1) | \mathbf{y}(0), \dots, \mathbf{y}(k)] &= \mathbb{E}[\mathbf{x}(k+1) | \tilde{\mathbf{y}}(0), \dots, \tilde{\mathbf{y}}(k)] \\ &= \mathbb{E}[\mathbf{x}(k+1) | \tilde{\mathbf{y}}(k)] + \mathbb{E}[\mathbf{x}(k+1) | \tilde{\mathbf{y}}(0), \dots, \tilde{\mathbf{y}}(k-1)] - \mathbb{E}[\mathbf{x}(k+1)]\end{aligned}$$

En la última igualdad, recordamos que para dos v.a. independientes Y_1, Y_2 ,

$$\mathbb{E}[X|Y_1, Y_2] = \mathbb{E}[X|Y_1] + \mathbb{E}[X|Y_2] - m_X$$

.

Como $\mathbf{x}(k+1)$ y $\tilde{\mathbf{y}}(k)$ tienen una distribución conjunta (gaussiana o no) conocida, tenemos

$$\mathbb{E}[\mathbf{x}(k+1) | \tilde{\mathbf{y}}(k)] = \mathbb{E}[\mathbf{x}(k+1)] + \underbrace{\mathbf{Cov}[\mathbf{x}(k+1), \tilde{\mathbf{y}}(k)]}_1 \underbrace{\mathbf{Cov}[\tilde{\mathbf{y}}(k), \tilde{\mathbf{y}}(k)]^{-1}}_2 \tilde{\mathbf{y}}(k)$$

Calculemos los términos 1 y 2 por separado

$$\mathbf{Cov}[\mathbf{x}(k+1), \tilde{\mathbf{y}}(k)] = \mathbf{Cov}[\mathbf{F}_k \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k), \mathbf{H}_k \tilde{\mathbf{x}}(k) + \mathbf{w}(k)] \quad (2)$$

$$= \mathbb{E}[(\mathbf{F}_k \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k)) (\mathbf{H}_k \tilde{\mathbf{x}}(k) + \mathbf{w}(k))^*] \quad (3)$$

$$= \mathbf{F}_k \mathbb{E}[\mathbf{x}(k) \tilde{\mathbf{x}}^*(k)] \mathbf{H}_k^* + \mathbf{G}_k \mathbb{E}[\mathbf{v}(k) \mathbf{w}(k)^*] \quad (4)$$

$$= \mathbf{F}_k \mathbb{E}[(\hat{\mathbf{x}}[k|k-1] + \tilde{\mathbf{x}}(k)) \tilde{\mathbf{x}}(k)^*] \mathbf{H}_k^* + \mathbf{G}_k \mathbf{S}_k \quad (5)$$

$$= \mathbf{F}_k \mathbb{E}[\tilde{\mathbf{x}}(k) \tilde{\mathbf{x}}(k)^*] \mathbf{H}_k^* + \mathbf{G}_k \mathbf{S}_k \quad (6)$$

$$= \mathbf{F}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{G}_k \mathbf{S}_k \quad (7)$$

- (2) a (3): Definición de **Cov**
- (3) a (4): Por hipótesis, $\mathbf{w}(k)$ está descorrelacionado con $\mathbf{x}(0)$ y con $\mathbf{v}(j)$ para todo k, j . Por ende, $\mathbf{w}(k)$ y $\mathbf{x}(k)$ están descorrelacionados. Por otro lado, $\mathbf{v}(k)$ y $\tilde{\mathbf{x}}(k)$ son independientes y de media nula. Por ende están descorrelacionados.
- (4) a (5): Reemplazo definición de innovación $\tilde{\mathbf{x}}(k) = \mathbf{x}(k) - \hat{\mathbf{x}}[k|k-1]$.
- (5) a (6): Por definición de innovación $\tilde{\mathbf{x}}(k)$ y $\hat{\mathbf{x}}[k|k-1]$ están descorrelacionadas
- Reemplazo definición de $\Sigma_{k|k-1}$

$$\begin{aligned}\mathbf{Cov}[\tilde{\mathbf{y}}(k), \tilde{\mathbf{y}}(k)] &= \mathbf{Cov}[\mathbf{H}_k \tilde{\mathbf{x}}(k) + \mathbf{w}(k), \mathbf{H}_k \tilde{\mathbf{x}}(k) + \mathbf{w}(k)] \\ &= \mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}_k\end{aligned}$$

Reemplazando estas dos expresiones arriba, tenemos:

$$\mathbb{E}[\mathbf{x}(k+1) | \tilde{\mathbf{y}}(k)] = \mathbb{E}[\mathbf{x}(k+1)] + [\mathbf{F}_x \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{G}_k \mathbf{S}_k] [\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}_k]^{-1} \tilde{\mathbf{y}}(k)$$

Por otro lado,

$$\begin{aligned}\mathbb{E}[\mathbf{x}(k+1) | \tilde{\mathbf{y}}(0), \dots, \tilde{\mathbf{y}}(k-1)] &= \mathbb{E}[\mathbf{F}_k \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k) | \tilde{\mathbf{y}}(0), \dots, \tilde{\mathbf{y}}(k-1)] \\ &= \mathbf{F}_k \mathbb{E}[\mathbf{x}(k) | \tilde{\mathbf{y}}(0), \dots, \tilde{\mathbf{y}}(k-1)] \\ &= \mathbf{F}_k \hat{\mathbf{x}}[k|k-1]\end{aligned}$$

Finalmente,

$$\hat{\mathbf{x}}[k+1|k] = \mathbf{F}_k \hat{\mathbf{x}}[k|k-1] + \left\{ \mathbf{F}_k \mathbf{K}_k + \mathbf{G}_k \mathbf{S}_k \left[\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}_k \right]^{-1} \right\} \cdot (\mathbf{y}(k) - \mathbf{H}_k \hat{\mathbf{x}}[k|k-1])$$

donde

$$\mathbf{K}_k = \Sigma_{k|k-1} \mathbf{H}_k^* \left[\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}_k \right]^{-1}$$

Sea

$$\Gamma_k = \mathbf{F}_k \mathbf{K}_k + \mathbf{G}_k \mathbf{S}_k \left[\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}_k \right]^{-1}$$

Luego,

$$\hat{\mathbf{x}}[k+1|k] = \mathbf{F}_k \hat{\mathbf{x}}[k|k-1] + \Gamma_k (\mathbf{y}(k) - \mathbf{H}_k \hat{\mathbf{x}}[k|k-1]).$$

(Observación: En la presentación inicial, tomé la covarianza entre $\mathbf{v}(k)$ y $\mathbf{w}(k)$ nula para simplificar las ecuaciones.)

Ahora le toca el turno a la matriz de covarianza

$$\Sigma_{k+1|k} = \mathbf{Cov} [\tilde{\mathbf{x}}(k+1), \tilde{\mathbf{x}}(k+1)] .$$

Es claro que

$$\begin{aligned}\tilde{\mathbf{x}}(k+1) &= \mathbf{x}(k+1) - \hat{\mathbf{x}}[k+1|k] \\ &= \mathbf{F}_k \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k) - \mathbf{F}_k \hat{\mathbf{x}}[k|k-1] - \Gamma_k (\mathbf{y}(k) - \mathbf{H}_k \hat{\mathbf{x}}[k|k-1]) \\ &= \mathbf{F}_k \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k) - \mathbf{F}_k \hat{\mathbf{x}}[k|k-1] - \Gamma_k (\mathbf{H}_k \mathbf{x}(k) + \mathbf{w}(k) - \mathbf{H}_k \hat{\mathbf{x}}[k|k-1]) \\ &= (\mathbf{F}_k - \Gamma_k \mathbf{H}_k) \tilde{\mathbf{x}}(k) + \mathbf{G}_k \mathbf{v}(k) - \Gamma_k \mathbf{w}(k) \\ &= (\mathbf{F}_k - \Gamma_k \mathbf{H}_k) \tilde{\mathbf{x}}(k) + [\mathbf{G}_k \quad -\Gamma_k] \begin{bmatrix} \mathbf{v}(k) \\ \mathbf{w}(k) \end{bmatrix}\end{aligned}$$

Recordemos que $\tilde{\mathbf{x}}(k)$ está descorrelacionado con $\mathbf{v}(k)$ y con $\mathbf{w}(k)$ ya que estos son procesos blancos descorrelacionados en el tiempo. Luego,

$$\begin{aligned}\Sigma_{k+1|k} &= (\mathbf{F}_k - \Gamma_k \mathbf{H}_k) \mathbf{Cov}[\tilde{\mathbf{x}}(k), \tilde{\mathbf{x}}(k)] (\mathbf{F}_k - \Gamma_k \mathbf{H}_k)^* \\ &\quad + [\mathbf{G}_k \quad -\Gamma_k] \begin{bmatrix} \mathbf{Q}_k & \mathbf{S}_k \\ \mathbf{S}_k^* & \mathbf{R}_k \end{bmatrix} \begin{bmatrix} \mathbf{G}_k^* \\ -\Gamma_k^* \end{bmatrix}\end{aligned}$$

Desarrollando,

$$\begin{aligned}\Sigma_{k+1|k} &= (\mathbf{F}_k - \Gamma_k \mathbf{H}_k) \Sigma_{k|k-1} (\mathbf{F}_k - \Gamma_k \mathbf{H}_k)^* \\ &\quad + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^* + \Gamma_k \mathbf{R}_k \Gamma_k^* - \mathbf{G}_k \mathbf{S}_k \Gamma_k^* - \Gamma_k \mathbf{S}_k^* \mathbf{G}_k^*.\end{aligned}$$

Diseñamos un filtro de Kalman para observar el estado del siguiente sistema lineal

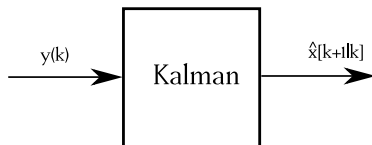
$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{F}_k \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k) \\ \mathbf{y}(k) &= \mathbf{H}_k \mathbf{x}(k) + \mathbf{w}(k)\end{aligned}$$

donde $\mathbf{v}(k)$ y $\mathbf{w}(k)$ son dos procesos blancos de media nula cuyas matrices de covarianza son \mathbf{Q}_k y \mathbf{R}_k respectivamente.

$$\begin{aligned}\hat{\mathbf{x}}[k+1|k] &= \mathbf{F}_k \hat{\mathbf{x}}[k|k-1] + \Gamma_k (\mathbf{y}(k) - \mathbf{H}_k \hat{\mathbf{x}}[k|k-1]) \\ &= \underbrace{(\mathbf{F}_k - \Gamma_k \mathbf{H}_k)}_{\mathbf{F}_{kalman}} \hat{\mathbf{x}}[k|k-1] + \Gamma_k \mathbf{y}(k)\end{aligned}$$

El filtro de Kalman es un sistema lineal cuya dinámica está determinada por \mathbf{F}_{kalman} .

$$\hat{\mathbf{x}}[k+1|k] = (\mathbf{F}_k - \Gamma_k \mathbf{H}_k) \hat{\mathbf{x}}[k|k-1] + \Gamma_k \mathbf{y}(k).$$



Es éste un sistema *estable*?

Analizamos este tema en el caso de un sistema LTI excitado por señales ESA, es decir:

$\mathbf{F}_k = \mathbf{F}$, $\mathbf{G}_k = \mathbf{G}$, $\mathbf{H}_k = \mathbf{H}$, $\mathbf{Q}_k = \mathbf{Q}$, y $\mathbf{R}_k = \mathbf{R}$.

Respuesta impulsiva de un sistema LTI

Sea $g(k)$ la respuesta impulsiva del sistema LTI en tiempo discreto.

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{F}\mathbf{x}(k) + \mathbf{G}\mathbf{v}(k) \\ \mathbf{y}(k) &= \mathbf{H}\mathbf{x}(k).\end{aligned}$$

Si $\mathbf{x}(0) = 0$, $g(k) = 0, k \leq 0$, $g(k) = \mathbf{H}\mathbf{F}^{k-1}\mathbf{G}, k \geq 1$.

Estabilidad en tiempo discreto

El sistema LTI en tiempo discreto

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{F}\mathbf{x}(k) + \mathbf{G}\mathbf{v}(k) \\ \mathbf{y}(k) &= \mathbf{H}\mathbf{x}(k) + \mathbf{w}(k)\end{aligned}$$

es BIBO estable si y sólo si todos los autovalores de \mathbf{F} caen dentro del círculo unitario, es decir, $|\lambda_i(\mathbf{F})| < 1, i = 1, \dots, n$.

Este concepto explora la capacidad de la señal de salida $\mathbf{y}(k)$ para estimar $\hat{\mathbf{x}}[k+1|k]$.

Observabilidad

El sistema en tiempo discreto

$$\mathbf{x}(k+1) = \mathbf{F}\mathbf{x}(k)$$

$$\mathbf{y}(k) = \mathbf{H}\mathbf{x}(k)$$

es **observable** si en todo momento se puede determinar el vector de estados $\mathbf{x}(k)$ a partir de la observación de la salida $\mathbf{y}(s)$ en los instantes siguientes $k \leq s$.

Sistema observable

Los siguientes puntos son equivalentes:

- El par (\mathbf{F}, \mathbf{H}) es observable;

-

$$\rho(\mathcal{O}) = \rho \left(\begin{bmatrix} \mathbf{H} \\ \mathbf{HF} \\ \mathbf{HF}^2 \\ \vdots \\ \mathbf{HF}^{n-1} \end{bmatrix} \right) = n,$$

donde $\mathcal{O} \in \mathbb{C}^{np \times n}$ es la matriz de observabilidad asociada al par $[\mathbf{F}, \mathbf{H}]$.

- Existe $\Gamma \in \mathbb{C}^{n \times p}$ tal que $|\lambda_i(\mathbf{F} - \Gamma\mathbf{H})| < 1, i = 1, \dots, n$.

Controlabilidad

El sistema en tiempo discreto

$$\mathbf{x}(k+1) = \mathbf{F}\mathbf{x}(k) + \mathbf{G}\mathbf{v}(k)$$

es **controlable** en tiempo finito, si a partir de cualquier estado inicial $\mathbf{x}(0)$ se puede alcanzar cualquier estado $\mathbf{x}(k)$ eligiendo la secuencia $\mathbf{v}(s)$, $s \leq k$.

Esto se cumple si y sólo si

$$\rho(\mathcal{C}) = \rho\left(\begin{bmatrix} \mathbf{G} & \mathbf{F}\mathbf{G} & \mathbf{F}^2\mathbf{G} & \dots & \mathbf{F}^{n-1}\mathbf{G} \end{bmatrix}\right) = n.$$

$\mathcal{C} \in \mathbb{C}^{n \times nl}$ es la matriz de controlabilidad asociada al par $[\mathbf{F}, \mathbf{G}]$.

Un sistema BIBO estable mantiene su salida acotada para toda entrada acotada. Qué sucede con los estados del sistema?

Teorema de Lyapunov (formulación reducida)

El sistema $\mathbf{x}(k+1) = \mathbf{F}\mathbf{x}(k)$ es estable (i.e., $|\lambda_i(\mathbf{F})| < 1, i = 1, \dots, n$) si y sólo si existen dos matrices $\mathbf{P} \geq 0$ y $\mathbf{V} \geq 0$ tal que

$$\mathbf{P} = \mathbf{F}\mathbf{P}\mathbf{F}^* + \mathbf{V}$$

Cuándo la salida de un sistema lineal es ESA? En Procesos Estocásticos vimos que la salida de todo sistema LTI impulsado por una entrada ESA es ESA. Es esta condición suficiente?

Supongamos un sistema LTI

$$\begin{cases} \mathbf{x}(k+1) = \mathbf{F}\mathbf{x}(k) + \mathbf{G}\mathbf{v}(k) \\ \mathbf{y}(k) = \mathbf{H}\mathbf{x}(k) + \mathbf{w}(k) \end{cases}$$

con $\mathbb{E}[\mathbf{v}(k)] = 0$, y $\mathbb{E}[\mathbf{v}(k)\mathbf{v}(r)] = \mathbf{Q}\delta_{kr}$, $\mathbb{E}[\mathbf{w}(k)] = 0$, y $\mathbb{E}[\mathbf{w}(k)\mathbf{w}(r)] = \mathbf{R}\delta_{kr}$.

A diferencia de Procesos Estocásticos, acá vamos a considerar que en el tiempo inicial k_0 , el sistema comienza a evolucionar. En ese instante, la condición inicial es un vector aleatorio $\mathbf{x}(k_0) = \mathbf{x}_0$ con matriz de covarianza \mathbf{P}_0 .

Estacionariedad de un sistema lineal II

Sea \mathbf{P}_k la matriz de covarianza de $\mathbf{x}(k)$. \mathbf{P}_k satisface la recursión

$$\mathbf{P}_{k+1} = \mathbf{F}\mathbf{P}_k\mathbf{F}^* + \mathbf{G}\mathbf{Q}\mathbf{G}^*.$$

Si \mathbf{F} es estable, entonces de acuerdo a Lyapunov, existe $\mathbf{P} \geq 0$ tal que

$$\mathbf{P} = \mathbf{F}\mathbf{P}\mathbf{F}^* + \underbrace{\mathbf{G}\mathbf{Q}\mathbf{G}^*}_{\mathbf{V}}.$$

Pregunta: $\mathbf{P}_k \rightarrow \mathbf{P}$?

$$\begin{aligned}\mathbf{P}_{k+1} - \mathbf{P} &= \mathbf{F}[\mathbf{P}_k - \mathbf{P}]\mathbf{F}^* \\ &= \mathbf{F}^k[\mathbf{P}_0 - \mathbf{P}](\mathbf{F}^k)^*\end{aligned}$$

Como \mathbf{F} es estable, sus autovalores están dentro del círculo unitario y por ende

$$\mathbf{F}^k \xrightarrow[k \rightarrow \infty]{} 0 \quad \text{luego,} \quad \mathbf{P}_k \xrightarrow[k \rightarrow \infty]{} \mathbf{P}$$

Un sistema LTI estable tiene estados y salidas asintóticamente ESA.

Vimos que el filtro de Kalman podía ser visto como un sistema lineal. Para el caso de un sistema LTI,

$$\hat{\mathbf{x}}[k+1|k] = (\mathbf{F} - \Gamma_k \mathbf{H}) \hat{\mathbf{x}}[k|k-1] + \Gamma_k \mathbf{y}(k).$$

En el caso que $\mathbf{v}(k)$ y $\mathbf{w}(k)$ estén descorrelacionado, $\mathbf{S} = 0$ y tenemos que

$$\Gamma_k = \mathbf{F} \mathbf{K}_k = \mathbf{F} \Sigma_{k|k-1} \mathbf{H}^* [\mathbf{H} \Sigma_{k|k-1} \mathbf{H}^* + \mathbf{R}]^{-1}$$

Observaciones:

- Que el sistema $(\mathbf{F}, \mathbf{G}, \mathbf{H})$ sea LTI y los procesos $\mathbf{v}(k)$ y $\mathbf{w}(k)$ sean ESA no garantiza que Γ_k o \mathbf{K}_k sean invariantes.
- Para tener un filtro invariante, se requiere que $\Sigma_{k|k-1} \rightarrow \Sigma$ para $k \rightarrow \infty$.
- La ganancia de Kalman asintótica resulta $\Gamma = \mathbf{F} \Sigma \mathbf{H}^* [\mathbf{H} \Sigma \mathbf{H}^* + \mathbf{R}]^{-1}$.
- Vamos a asumir que $[\mathbf{H} \Sigma \mathbf{H}^* + \mathbf{R}]^{-1}$ existe. Para ello, se requiere que $\Sigma \geq 0$ y \mathbf{R} sea definida positiva.
- Para tener una estimación ESA, necesitamos que el filtro sea estable, es decir, $|\lambda_i(\mathbf{F} - \Gamma \mathbf{H})| < 1, i = 1, \dots, n$.

Por qué nos interesa un filtro estable? Cómo es la dinámica del error de estimación?

$$\begin{aligned}\tilde{\mathbf{x}}(k+1) &= \mathbf{x}(k+1) - \hat{\mathbf{x}}[k+1|k] \\ &= \mathbf{F}\mathbf{x}(k) + \mathbf{G}\mathbf{v}(k) - \{\mathbf{F}\hat{\mathbf{x}}[k|k-1] + \Gamma[\mathbf{H}\mathbf{x}(k) + \mathbf{w}(k) - \mathbf{H}\hat{\mathbf{x}}[k|k-1]]\} \\ &= [\mathbf{F} - \Gamma\mathbf{H}]\tilde{\mathbf{x}}(k) + \mathbf{G}\mathbf{v}(k) - \Gamma\mathbf{w}(k)\end{aligned}$$

- La dinámica del filtro de Kalman está dada por $(\mathbf{F} - \Gamma\mathbf{H})$.
- Ésta es la misma dinámica que sigue el error de estimación $\tilde{\mathbf{x}}(k)$.
- Si el filtro es estable y asintóticamente LTI, entonces las innovaciones son ESA. Más aún, $\tilde{\mathbf{x}}(k) \rightarrow 0$ para toda condición inicial $\hat{\mathbf{x}}[0|-1]$.

- La ecuación de actualización de la matriz de covarianza del error de estimación

$$\Sigma_{k+1|k} = (\mathbf{F}_k - \Gamma_k \mathbf{H}_k) \Sigma_{k|k-1} (\mathbf{F}_k - \Gamma_k \mathbf{H}_k)^* + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^* + \Gamma_k \mathbf{R}_k \Gamma_k^*$$

- En régimen asintótico o de estado estacionario, esperamos que

$$\Sigma_{k+1|k} = \Sigma_{k|k-1} = \Sigma.$$

- Por otro lado, la matriz $\Gamma_k = \Gamma = \mathbf{F} \Sigma \mathbf{H}^* [\mathbf{H} \Sigma \mathbf{H}^* + \mathbf{R}]^{-1}$

- Reemplazando arriba tenemos,

$$\begin{aligned} \Sigma &= \mathbf{F} \Sigma \mathbf{F}^* - \mathbf{F} \Sigma \mathbf{H}^* \Gamma^* - \Gamma \mathbf{H} \Sigma \mathbf{F}^* + \underbrace{\Gamma \mathbf{H} \Sigma \mathbf{H}^* \Gamma^* + \Gamma \mathbf{R} \Gamma^*}_{\Gamma [\mathbf{H} \Sigma \mathbf{H}^* + \mathbf{R}] \Gamma^*} + \mathbf{G} \mathbf{Q} \mathbf{G}^* \\ &= \mathbf{F} \Sigma \mathbf{F}^* - \mathbf{F} \Sigma \mathbf{H}^* \Gamma^* - \Gamma \mathbf{H} \Sigma \mathbf{F}^* + \underbrace{\Gamma [\mathbf{H} \Sigma \mathbf{H}^* + \mathbf{R}] \Gamma^*}_{\mathbf{F} \Sigma \mathbf{H}^* \Gamma^* + \mathbf{G} \mathbf{Q} \mathbf{G}^*} + \mathbf{G} \mathbf{Q} \mathbf{G}^* \\ &= \mathbf{F} \Sigma \mathbf{F}^* - \mathbf{F} \Sigma \mathbf{H}^* \Gamma^* - \Gamma \mathbf{H} \Sigma \mathbf{F}^* + \mathbf{F} \Sigma \mathbf{H}^* \Gamma^* + \mathbf{G} \mathbf{Q} \mathbf{G}^* \end{aligned}$$

Discrete-time Algebraic Ricatti Equation(DARE)

- La ecuación de actualización de la matriz de covarianza del error de estimación

$$\Sigma_{k+1|k} = (\mathbf{F}_k - \Gamma_k \mathbf{H}_k) \Sigma_{k|k-1} (\mathbf{F}_k - \Gamma_k \mathbf{H}_k)^* + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^* + \Gamma_k \mathbf{R}_k \Gamma_k^*$$

- En régimen asintótico o de estado estacionario, esperamos que

$$\Sigma_{k+1|k} = \Sigma_{k|k-1} = \Sigma.$$

- Por otro lado, la matriz $\Gamma_k = \Gamma = \mathbf{F} \Sigma \mathbf{H}^* [\mathbf{H} \Sigma \mathbf{H}^* + \mathbf{R}]^{-1}$

- Reemplazando arriba tenemos,

$$\begin{aligned}\Sigma &= \mathbf{F} \Sigma \mathbf{F}^* - \mathbf{F} \Sigma \mathbf{H}^* \Gamma^* - \Gamma \mathbf{H} \Sigma \mathbf{F}^* + \underbrace{\Gamma \mathbf{H} \Sigma \mathbf{H}^* \Gamma^* + \Gamma \mathbf{R} \Gamma^*}_{\mathbf{F} \Sigma \mathbf{H}^* \Gamma^* + \Gamma \mathbf{H} \Sigma \mathbf{F}^*} + \mathbf{G} \mathbf{Q} \mathbf{G}^* \\ &= \mathbf{F} \Sigma \mathbf{F}^* - \mathbf{F} \Sigma \mathbf{H}^* \Gamma^* - \Gamma \mathbf{H} \Sigma \mathbf{F}^* + \underbrace{\Gamma [\mathbf{H} \Sigma \mathbf{H}^* + \mathbf{R}] \Gamma^*}_{\mathbf{F} \Sigma \mathbf{H}^* \Gamma^* + \Gamma \mathbf{H} \Sigma \mathbf{F}^*} + \mathbf{G} \mathbf{Q} \mathbf{G}^* \\ &= \mathbf{F} \Sigma \mathbf{F}^* - \mathbf{F} \Sigma \mathbf{H}^* \Gamma^* - \Gamma \mathbf{H} \Sigma \mathbf{F}^* + \mathbf{F} \Sigma \mathbf{H}^* \Gamma^* + \mathbf{G} \mathbf{Q} \mathbf{G}^*\end{aligned}$$

DARE:Discrete-time Algebraic Ricatti Equation

$$\Sigma = \mathbf{F} \Sigma \mathbf{F}^* - \mathbf{F} \Sigma \mathbf{H}^* (\mathbf{H} \Sigma \mathbf{H}^* + \mathbf{R})^{-1} \mathbf{H} \Sigma \mathbf{F}^* + \mathbf{G} \mathbf{Q} \mathbf{G}^* \quad (\text{DARE})$$

Ecuación algebraica de Ricatti

$$\Sigma = \mathbf{F}\Sigma\mathbf{F}^* - \mathbf{F}\Sigma\mathbf{H}^* (\mathbf{H}\Sigma\mathbf{H}^* + \mathbf{R})^{-1} \mathbf{H}\Sigma\mathbf{F}^* + \mathbf{G}\mathbf{Q}\mathbf{G}^* \quad (\text{DARE})$$

- DARE es un sistema de ecuaciones que puede tener varias soluciones (o ninguna).
- Obtener las soluciones de la DARE es complejo, pero existen rutinas numéricas adecuadas para ello.
- Para tener $\Sigma_{k|k-1} \rightarrow \Sigma$, necesitamos que $\Sigma_{k|k-1}$ esté acotada y que la convergencia sea independiente de $\Sigma_{0|-1}$.

Volvemos al sistema de innovaciones del estado:

$$\tilde{\mathbf{x}}(k+1) = [\mathbf{F} - \Gamma\mathbf{H}] \tilde{\mathbf{x}}(k) + [\mathbf{G} \quad -\Gamma] \begin{bmatrix} \mathbf{v}(k) \\ \mathbf{w}(k) \end{bmatrix}$$

De acuerdo a Lyapunov, este sistema es estable si existe $\Sigma > 0$ $\mathbf{V} > 0$ tal que

$$\Sigma = (\mathbf{F} - \Gamma\mathbf{H}) \Sigma (\mathbf{F} - \Gamma\mathbf{H})^* + \mathbf{V}$$

$$\text{Si } \mathbf{V} = [\mathbf{G} \quad -\Gamma] \begin{bmatrix} \mathbf{Q} & 0 \\ 0 & \mathbf{R} \end{bmatrix} \begin{bmatrix} \mathbf{G}^* \\ -\Gamma^* \end{bmatrix} = \mathbf{G}\mathbf{Q}\mathbf{G}^* + \Gamma\mathbf{R}\Gamma^*, \text{ obtenemos la DARE}$$

$$\Sigma = (\mathbf{F} - \Gamma\mathbf{H}) \Sigma (\mathbf{F} - \Gamma\mathbf{H})^* + \mathbf{G}\mathbf{Q}\mathbf{G}^* + \Gamma\mathbf{R}\Gamma^*$$

Entonces, para tener convergencia del filtro de Kalman, se necesita que $\Sigma > 0$. En ese caso, $|\lambda_i(\mathbf{F} - \Gamma\mathbf{H})| < 1$, $i = 1, \dots, n$. Esto se conoce como una *solución estabilizante*.

$$\Sigma = \mathbf{F}\Sigma\mathbf{F}^* - \mathbf{F}\Sigma\mathbf{H}^* (\mathbf{H}\Sigma\mathbf{H}^* + \mathbf{R})^{-1} \mathbf{H}\Sigma\mathbf{F}^* + \mathbf{G}\mathbf{Q}\mathbf{G}^* \quad (\text{DARE})$$

Teorema: Solución estabilizante

Supongamos que \mathbf{F} es estable. Luego, existe una única solución estabilizante Σ si y sólo si (\mathbf{F}, \mathbf{H}) es observable y $(\mathbf{F}, \mathbf{G}\mathbf{Q}^{1/2})$ es controlable.

- La condición de observabilidad de (\mathbf{F}, \mathbf{H}) garantiza la existencia de una matriz Γ que resulta en un filtro de Kalman estable.
- La condición de controlabilidad de $(\mathbf{F}, \mathbf{G}\mathbf{Q}^{1/2})$ permite obtener una solución semidefinida positiva que puede ser una matriz de covarianza. De hecho, Σ corresponde a la matriz de covarianza asintótica.

- Cuando trabajamos con un sistema LTI, podemos utilizar la ganancia asintótica Γ resolviendo la DARE.
- Si (\mathbf{F}, \mathbf{H}) es observable y $(\mathbf{F}, \mathbf{G}\mathbf{Q}^{1/2})$ es controlable, el filtro de Kalman asintótico resulta

$$\hat{\mathbf{x}}[k+1|k] = [\mathbf{F} - \Gamma\mathbf{H}] \hat{\mathbf{x}}[k|k-1] + \Gamma\mathbf{y}(k)$$

donde

$$\Gamma = \mathbf{F}\Sigma\mathbf{H}^* [\mathbf{H}\Sigma\mathbf{H}^* + \mathbf{R}]^{-1}$$

y $\Sigma \geq 0$ es la única solución de

$$\Sigma = \mathbf{F}\Sigma\mathbf{F}^* - \mathbf{F}\Sigma\mathbf{H}^* (\mathbf{H}\Sigma\mathbf{H}^* + \mathbf{R})^{-1} \mathbf{H}\Sigma\mathbf{F}^* + \mathbf{G}\mathbf{Q}\mathbf{G}^*$$

- Para tiempo finito, ésta es una solución subóptima, pero más conveniente de implementar que la solución óptima que varía en el tiempo. La ganancia se precalcula y no es necesario actualizar en forma recursiva la covarianza.

Más allá del desarrollo teórico para obtener las expresiones del filtro de Kalman, el mismo puede ser visto como un sistema lineal parametrizado por \mathbf{Q}_k y \mathbf{R}_k .

- Qué representa \mathbf{Q}_k ? En algunos casos, el proceso $\mathbf{v}(k)$ modela la incertidumbre del sistema. En estos casos, la matriz \mathbf{Q} puede ser usada como "perilla de ajuste" del sistema.

Más allá del desarrollo teórico para obtener las expresiones del filtro de Kalman, el mismo puede ser visto como un sistema lineal parametrizado por \mathbf{Q}_k y \mathbf{R}_k .

- Qué representa \mathbf{Q}_k ? En algunos casos, el proceso $\mathbf{v}(k)$ modela la incertidumbre del sistema. En estos casos, la matriz \mathbf{Q} puede ser usada como "perilla de ajuste" del sistema.
- Qué representa \mathbf{R}_k ? En general es la covarianza del ruido de medición. Por otro lado, $\mathbf{R} > 0$ asegura que la covarianza de las innovaciones, $(\mathbf{H}\Sigma\mathbf{H}^* + \mathbf{R})$, sea definida positiva.

Cómo influye la elección de las matrices de covarianza en la performance del filtro?
Podemos ganar cierta intuición analizando cómo varía la ganancia del filtro en ciertas circunstancias.

- Supongamos por un momento que \mathbf{H}_k es invertible para todo k y que $\mathbf{R}_k \rightarrow 0$. Esto último implica que las mediciones tienen poco ruido y resultan más confiables. En ese caso, $\Gamma_k \rightarrow \mathbf{F}_k \mathbf{H}_k^{-1}$. Decimos que la estimación confía fuertemente en la nueva información aportada por las innovaciones y aumenta la ganancia para darles mayor relevancia en el cómputo de $\hat{\mathbf{x}}[k+1|k]$.
- Por otro lado, si $\Sigma_{k|k-1} \rightarrow 0$, entonces $\Gamma_k \rightarrow 0$. En este caso, la ganancia confía en la estimación del estado y su evolución temporal dada por el modelo de estados. El filtro "cierra los ojos" ante nuevas observaciones y evoluciona siguiendo el modelo solamente.
- Un modo de prevenir que $\Sigma_{k|k-1} \rightarrow 0$ es asegurar que $\mathbf{Q}_k > 0$, es decir, mantener siempre un cierto nivel de ruido de proceso.

Y si el sistema es no-lineal?

Hasta el momento, analizamos sistemas lineales modelados por sus ecuaciones de estado

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{F}_k \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k) \\ \mathbf{y}(k) &= \mathbf{H}_k \mathbf{x}(k) + \mathbf{w}(k).\end{aligned}$$

Pero el caso más general es cuando la dinámica y/o las mediciones tienen una relación no-lineal, es decir

$$\begin{aligned}\mathbf{x}(k+1) &= f_k(\mathbf{x}(k)) + g_k(\mathbf{x}(k))\mathbf{v}(k) \\ \mathbf{y}(k) &= h_k(\mathbf{x}(k)) + \mathbf{w}(k),\end{aligned}$$

donde $f_k : \mathbb{C}^n \rightarrow \mathbb{C}^n$, $g_k : \mathbb{C}^n \rightarrow \mathbb{C}^{n \times l}$, $h_k : \mathbb{C}^n \rightarrow \mathbb{C}^p$, son funciones del vector de estado $\mathbf{x}(k)$.

Ejemplo

Supongamos un móvil que se mueve sobre un plano, cuyas coordenadas cartesianas son (p_x, p_y) .

El movimiento del móvil en tiempo discreto está determinado por

$$p_x(k+1) = f_1 p_x(k) + v_1(k) \quad p_y(k+1) = f_2 p_y(k) + v_2(k)$$

donde $f_1, f_2, \in \mathbb{R}$ y $v_1(k), v_2(k)$ son dos procesos blancos gaussianos de media nula y varianzas conocidas.

Supongamos que en lugar de observar a cada instante las coordenadas cartesianas, sólo tenemos acceso a las coordenadas polares, es decir, a cada instante disponemos de

$$\rho(k) = \sqrt{p_x(k)^2 + p_y(k)^2} \quad \phi(k) = \text{atan} \left(\frac{p_y(k)}{p_x(k)} \right)$$

En este caso el estado es $\mathbf{x}(k) = \begin{bmatrix} p_x(k) \\ p_y(k) \end{bmatrix}$, las mediciones $\mathbf{y}(k) = \begin{bmatrix} \rho(k) \\ \phi(k) \end{bmatrix}$. Por fin,

$$f_k(\mathbf{x}(k)) = \begin{bmatrix} f_1 & 0 \\ 0 & f_2 \end{bmatrix} \begin{bmatrix} x_1(k) \\ x_2(k) \end{bmatrix}, \quad h_k(\mathbf{x}(k)) = \begin{bmatrix} \sqrt{x_1(k)^2 + x_2(k)^2} & \text{atan} \left(\frac{x_2(k)}{x_1(k)} \right) \end{bmatrix}$$

Sabemos que el estimador MMSE para el estado $\mathbf{x}(k)$ es la esperanza condicional

$$\mathbb{E}[\mathbf{x}(k)|\mathbf{y}(k), \dots, \mathbf{y}(0)].$$

Pero es muy difícil de computar en el caso general!!!!

Tenemos que recurrir a aproximaciones para poder estimar \hat{p}_x y \hat{p}_y .

Linearización sobre una trayectoria

Supongamos que sabemos que $\mathbf{x}(k)$ evoluciona siguiendo aproximadamente una trayectoria conocida $\bar{\mathbf{x}}(k)$.

$$\Delta \mathbf{x}(k) = \mathbf{x}(k) - \bar{\mathbf{x}}(k)$$

Para todo instante, asumimos que $|\Delta \mathbf{x}(k)|$ es pequeño. Luego, planteamos un desarrollo de Taylor en cada punto de la trayectoria $\bar{\mathbf{x}}$.

$$f_k(\mathbf{x}(k)) \simeq f_k(\bar{\mathbf{x}}(k)) + \mathbf{F}_k \Delta \mathbf{x}(k) \quad h_k(\mathbf{x}(k)) \simeq h_k(\bar{\mathbf{x}}(k)) + \mathbf{H}_k \Delta \mathbf{x}(k),$$

donde

$$\mathbf{F}_k = \left. \frac{\partial f_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\bar{\mathbf{x}}(k)} \quad \mathbf{H}_k = \left. \frac{\partial h_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\bar{\mathbf{x}}(k)}.$$

Para tener un sistema lineal, tomamos orden 0 para el ruido de proceso, es decir,

$$g_k(\mathbf{x}(k)) \simeq g_k(\bar{\mathbf{x}}(k)) = \mathbf{G}_k$$

- Modelo dinámico:

$$\begin{cases} \mathbf{x}(k+1) &= f_k(\mathbf{x}(k)) + g_k(\mathbf{x}(k))\mathbf{v}(k) \\ \mathbf{y}(k) &= h_k(\mathbf{x}(k)) + \mathbf{w}(k) \end{cases}$$

- Trayectoria: $\bar{\mathbf{x}}(k+1) = f_k(\bar{\mathbf{x}}(k))$.
- Taylor estado: $\mathbf{x}(k+1) \simeq f_k(\bar{\mathbf{x}}(k)) + \mathbf{F}_k \Delta \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k)$.
- Taylor observaciones: $\mathbf{y}(k) \simeq h_k(\bar{\mathbf{x}}(k)) + \mathbf{H}_k \Delta \mathbf{x}(k) + \mathbf{w}(k)$.

Definimos $\Delta \mathbf{x}(k) = \mathbf{x}(k) - \bar{\mathbf{x}}(k)$ y $\Delta \mathbf{y}(k) = \mathbf{y}(k) - h_k(\bar{\mathbf{x}}(k))$. Luego,

$$\begin{cases} \Delta \mathbf{x}(k+1) &= \mathbf{F}_k \Delta \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k) \\ \Delta \mathbf{y}(k) &= \mathbf{H}_k \Delta \mathbf{x}(k) + \mathbf{w}(k) \end{cases}$$

es el sistema linearizado alrededor de la trayectoria $\bar{\mathbf{x}}(k)$.

1 Kalman para sistema linearizado

$$\begin{aligned}\mathbf{K}_k &= \Sigma_{k|k-1} \mathbf{H}_k^* [\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}_k]^{-1} \\ \hat{\Delta \mathbf{x}}[k+1|k] &= \mathbf{F}_k \hat{\Delta \mathbf{x}}[k|k-1] + \mathbf{F}_k \mathbf{K}_k \left(\Delta \mathbf{y}(k) - \mathbf{H}_k \hat{\Delta \mathbf{x}}[k|k-1] \right) \\ \Sigma_{k+1|k} &= \mathbf{F}_k \Sigma_{k|k-1} \mathbf{F}_k^* - \mathbf{F}_k \Sigma_{k|k-1} \mathbf{H}_k^* [\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}_k]^{-1} \mathbf{H}_k \Sigma_{k|k-1} \mathbf{F}_k^* + \mathbf{G}_k \mathbf{Q}_k\end{aligned}$$

2 Reconstrucción del estado del sistema no lineal

$$\begin{aligned}\hat{\mathbf{x}}[k+1|k] &= \bar{\mathbf{x}}(k+1) + \hat{\Delta \mathbf{x}}[k+1|k] \\ &= f_k(\bar{\mathbf{x}}(k)) + \hat{\Delta \mathbf{x}}[k+1|k]\end{aligned}$$

Observaciones:

- Las matrices \mathbf{F}_k , \mathbf{H}_k , y \mathbf{G}_k sólo dependen de $\bar{\mathbf{x}}(k)$.
- Para una trayectoria conocida, \mathbf{K}_k puede ser calculada de antemano.
- Parte de las varianzas de $\mathbf{v}(k)$ y $\mathbf{w}(k)$ cubren el error de aproximación de $f_k(\cdot)$ y $h_k(\cdot)$.
- La performance depende mucho de la trayectoria elegida.

Filtro de Kalman Extendido (EKF)

El EKF (*Extended Kalman Filter*) es una opción para diseñar estimadores recursivos al estilo Kalman para sistemas no-lineales.

La idea es similar a la linealización a lo largo de una trayectoria, pero la trayectoria ahora es la del estado estimado

$$\begin{aligned}f_k(\mathbf{x}(k)) &\simeq f_k(\hat{\mathbf{x}}[k|k]) + \mathbf{F}_k (\mathbf{x}(k) - \hat{\mathbf{x}}[k|k]) \\h_k(\mathbf{x}(k)) &\simeq h_k(\hat{\mathbf{x}}[k|k-1]) + \mathbf{H}_k (\mathbf{x}(k) - \hat{\mathbf{x}}[k|k-1]) \\g_k(\mathbf{x}(k)) &\simeq g_k(\hat{\mathbf{x}}[k|k]) = \mathbf{G}_k,\end{aligned}$$

donde

$$\mathbf{F}_k = \left. \frac{\partial f_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}[k|k]} \quad \mathbf{H}_k = \left. \frac{\partial h_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}[k|k-1]}.$$

- Taylor estado alrededor de $\hat{\mathbf{x}}[k|k]$:

$$\mathbf{x}(k+1) \simeq f_k(\hat{\mathbf{x}}[k|k]) + \mathbf{F}_k (\mathbf{x}(k) - \hat{\mathbf{x}}[k|k]) + \mathbf{G}_k \mathbf{v}(k).$$

Reordenando términos, obtenemos un sistema lineal con entrada conocida.

$$\mathbf{x}(k+1) = \mathbf{F}_k \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k) + \underbrace{(f_k(\hat{\mathbf{x}}[k|k]) - \mathbf{F}_k \hat{\mathbf{x}}[k|k])}_{\mathbf{u}(k)}$$

- Taylor observaciones alrededor de $\hat{\mathbf{x}}[k|k-1]$:

$$\mathbf{y}(k) \simeq h_k(\hat{\mathbf{x}}[k|k-1]) + \mathbf{H}_k (\mathbf{x}(k) - \hat{\mathbf{x}}[k|k-1]) + \mathbf{w}(k).$$

Nuevamente, sistema lineal con offset conocido en $k-1$.

$$\mathbf{z}(k) = \mathbf{y}(k) - \underbrace{(h_k(\hat{\mathbf{x}}[k|k-1]) - \mathbf{H}_k \hat{\mathbf{x}}[k|k-1])}_{\mathbf{y}_o(k)} = \mathbf{H}_k \mathbf{x}(k) + \mathbf{w}(k).$$

Filtro de Kalman Extendido (EKF)

- Sistema linealizado

$$\begin{cases} \mathbf{x}(k+1) &= \mathbf{F}_k \mathbf{x}(k) + \mathbf{G}_k \mathbf{v}(k) + \mathbf{u}(k) \\ \mathbf{z}(k) &= \mathbf{H}_k \mathbf{x}(k) + \mathbf{w}(k) \end{cases}$$

donde

$$\mathbf{F}_k = \left. \frac{\partial f_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}[k|k]} \quad \mathbf{H}_k = \left. \frac{\partial h_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}[k|k-1]} \quad \mathbf{G}_k = g_k(\hat{\mathbf{x}}[k|k]).$$

$$\mathbf{z}(k) = \mathbf{y}(k) - (h_k(\hat{\mathbf{x}}[k|k-1]) - \mathbf{H}_k \hat{\mathbf{x}}[k|k-1]) \quad \mathbf{u}(k) = f_k(\hat{\mathbf{x}}[k|k]) - \mathbf{F}_k \hat{\mathbf{x}}[k|k]$$

- Una vez obtenidas las matrices \mathbf{F}_k , \mathbf{H}_k , y \mathbf{G}_k , se calculan $\Sigma_{k+1|k}$ y \mathbf{K}_k .
- Actualización de mediciones:

$$\begin{aligned} \hat{\mathbf{x}}[k|k] &= \hat{\mathbf{x}}[k|k-1] + \mathbf{K}_k (\mathbf{z}(k) - \mathbf{H}_k \hat{\mathbf{x}}[k|k-1]) \\ &= \hat{\mathbf{x}}[k|k-1] + \mathbf{K}_k (\mathbf{y}(k) - h_k(\hat{\mathbf{x}}[k|k-1]) + \mathbf{H}_k \hat{\mathbf{x}}[k|k-1] - \mathbf{H}_k \hat{\mathbf{x}}[k|k-1]) \\ &= \hat{\mathbf{x}}[k|k-1] + \mathbf{K}_k (\mathbf{y}(k) - h_k(\hat{\mathbf{x}}[k|k-1])) \end{aligned}$$

- Actualización de tiempos:

$$\begin{aligned} \hat{\mathbf{x}}[k+1|k] &= \mathbf{F}_k \hat{\mathbf{x}}[k|k] + \mathbf{u}(k) \\ &= \mathbf{F}_k \hat{\mathbf{x}}[k|k] + f_k(\hat{\mathbf{x}}[k|k]) - \mathbf{F}_k \hat{\mathbf{x}}[k|k] \\ &= f_k(\hat{\mathbf{x}}[k|k]) \end{aligned}$$

Para cada instante k ,

1 Linearización:

$$\mathbf{F}_k = \left. \frac{\partial f_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}[k|k]} \quad \mathbf{H}_k = \left. \frac{\partial h_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}[k|k-1]} \quad \mathbf{G}_k = g_k(\hat{\mathbf{x}}[k|k]).$$

2 Covarianzas y ganancia

$$\mathbf{K}_k = \Sigma_{k|k-1} \mathbf{H}_k^* [\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}_k]^{-1}$$
$$\Sigma_{k+1|k} = \mathbf{F}_k \Sigma_{k|k-1} \mathbf{F}_k^* - \mathbf{F}_k \Sigma_{k|k-1} \mathbf{H}_k^* [\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}_k]^{-1} \mathbf{H}_k \Sigma_{k|k-1} \mathbf{F}_k^* + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^*$$

3 Estimación del estado

$$\hat{\mathbf{x}}[k|k] = \hat{\mathbf{x}}[k|k-1] + \mathbf{K}_k (\mathbf{y}(k) - h_k(\hat{\mathbf{x}}[k|k-1]))$$
$$\hat{\mathbf{x}}[k+1|k] = f_k(\hat{\mathbf{x}}[k|k])$$

Observaciones:

- Las matrices \mathbf{F}_k , \mathbf{H}_k , y \mathbf{G}_k dependen de las observaciones en k e instantes anteriores. No pueden ser calculadas de antemano.
- $\hat{\mathbf{x}}[k|k]$ tiene una dependencia no lineal con las observaciones anteriores a través de \mathbf{H}_k .
- Es muy difícil establecer un indicador de performance del EKF de forma analítica
- Un modo de testear la bondad de la estimación es utilizando los residuos $\varepsilon(k) = \mathbf{y}(k) - h_k(\hat{\mathbf{x}}[k|k-1])$
 - ▶ Idealmente, estos residuos deben ser blancos
 - ▶ Idealmente, la matriz de covarianza $\varepsilon(k)$ se debería comportar como $\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^T + \mathbf{R}_k$, que es la covarianza de las innovaciones del sistema linealizado.

las autocorrelaciones de las innovaciones idealmente descorrelacionado en el tiempo

Dinámica del sistema en tiempo discreto:

$$\begin{bmatrix} \lambda(k+1) \\ \theta(k+1) \end{bmatrix} = \underbrace{\begin{bmatrix} 1 - \frac{T}{\beta} & 0 \\ T & 1 \end{bmatrix}}_{\mathbf{F}} \underbrace{\begin{bmatrix} \lambda(k) \\ \theta(k) \end{bmatrix}}_{\mathbf{x}(k)} + \underbrace{\begin{bmatrix} 1 \\ 0 \end{bmatrix}}_{\mathbf{G}} v(k)$$

La dinámica del sistema es lineal. La no linealidad está en la ecuación de salida. Consideramos dos posibles salidas:

Salida escalar:

$$y(k) = \sqrt{2} \sin(\omega_c k + \theta(k)) + w(k)$$

Salida vectorial:

$$\mathbf{y}(k) = \sqrt{2} \begin{bmatrix} \sin(\theta(k)) \\ \cos(\theta(k)) \end{bmatrix} + \mathbf{w}(k)$$

1 Linearización:

$$\mathbf{F}_k = \mathbf{F} \quad \mathbf{G}_k = \mathbf{G}.$$

$$\begin{aligned} \mathbf{H}_k &= \begin{bmatrix} \left. \frac{\partial h_k(\mathbf{x})}{\partial x_1} \right|_{\mathbf{x}=\hat{\mathbf{x}}[k|k-1]} & \left. \frac{\partial h_k(\mathbf{x})}{\partial x_2} \right|_{\mathbf{x}=\hat{\mathbf{x}}[k|k-1]} \end{bmatrix} \\ &= \begin{bmatrix} 0 & \sqrt{2} \cos(\omega_c k + \hat{x}_2[k|k-1]) \end{bmatrix} \end{aligned}$$

2 Covarianzas y ganancia

$$\begin{aligned} \mathbf{K}_k &= \frac{\Sigma_{k|k-1} \mathbf{H}_k^*}{\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + r} \\ \Sigma_{k+1|k} &= \mathbf{F} \Sigma_{k|k-1} \mathbf{F}^* - \frac{\mathbf{F} \Sigma_{k|k-1} \mathbf{H}_k^* \mathbf{H}_k \Sigma_{k|k-1} \mathbf{F}^*}{\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + r} + q \mathbf{G} \mathbf{G}^* \end{aligned}$$

3 Estimación del estado

$$\begin{aligned} \begin{bmatrix} \hat{\lambda}[k|k] \\ \hat{\theta}[k|k] \end{bmatrix} &= \begin{bmatrix} \hat{\lambda}[k|k-1] \\ \hat{\theta}[k|k-1] \end{bmatrix} + \mathbf{K}_k \left(\mathbf{y}(k) - \sqrt{2} \sin(\omega_c k + \hat{\theta}[k|k-1]) \right) \\ \begin{bmatrix} \hat{\lambda}[k+1|k] \\ \hat{\theta}[k+1|k] \end{bmatrix} &= \mathbf{F} \begin{bmatrix} \hat{\lambda}[k|k] \\ \hat{\theta}[k|k] \end{bmatrix} \end{aligned}$$

1 Linearización:

$$\mathbf{F}_k = \mathbf{F} \quad \mathbf{G}_k = \mathbf{G}.$$

$$\mathbf{H}_k = \begin{bmatrix} \left. \frac{\partial h_{1k}(\mathbf{x})}{\partial x_1} \right|_{\mathbf{x}=\hat{\mathbf{x}}[k|k-1]} & \left. \frac{\partial h_{1k}(\mathbf{x})}{\partial x_2} \right|_{\mathbf{x}=\hat{\mathbf{x}}[k|k-1]} \\ \left. \frac{\partial h_{2k}(\mathbf{x})}{\partial x_1} \right|_{\mathbf{x}=\hat{\mathbf{x}}[k|k-1]} & \left. \frac{\partial h_{2k}(\mathbf{x})}{\partial x_2} \right|_{\mathbf{x}=\hat{\mathbf{x}}[k|k-1]} \end{bmatrix} = \begin{bmatrix} 0 & \sqrt{2} \cos(\hat{\theta}[k|k-1]) \\ 0 & -\sqrt{2} \sin(\hat{\theta}[k|k-1]) \end{bmatrix}$$

2 Covarianzas y ganancia

$$\mathbf{K}_k = \Sigma_{k|k-1} \mathbf{H}_k^* [\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}]^{-1}$$

$$\Sigma_{k+1|k} = \mathbf{F} \Sigma_{k|k-1} \mathbf{F}^* - \mathbf{F} \Sigma_{k|k-1} \mathbf{H}_k^* [\mathbf{H}_k \Sigma_{k|k-1} \mathbf{H}_k^* + \mathbf{R}]^{-1} \mathbf{H}_k \Sigma_{k|k-1} \mathbf{F}^* + q \mathbf{G} \mathbf{G}^*$$

3 Estimación del estado

$$\begin{bmatrix} \hat{\lambda}[k|k] \\ \hat{\theta}[k|k] \end{bmatrix} = \begin{bmatrix} \hat{\lambda}[k|k-1] \\ \hat{\theta}[k|k-1] \end{bmatrix} + \mathbf{K}_k \left(\mathbf{y}(k) - \sqrt{2} \begin{bmatrix} \sin(\hat{\theta}[k|k-1]) \\ \cos(\hat{\theta}[k|k-1]) \end{bmatrix} \right)$$

$$\begin{bmatrix} \hat{\lambda}[k+1|k] \\ \hat{\theta}[k+1|k] \end{bmatrix} = \mathbf{F} \begin{bmatrix} \hat{\lambda}[k|k] \\ \hat{\theta}[k|k] \end{bmatrix}$$

Filtro Iterado de Kalman (IKF)

En algunos casos, en particular cuando $h_k(\cdot)$ es una función con fuertes no linealidades, es conveniente utilizar la siguiente adaptación del EKF conocida como IKF (*Iterated Kalman Filter*).

Para cada instante k , computar:

- ❶ Secuencia de matrices $\mathbf{H}_k^0, \dots, \mathbf{H}_k^L$ tal que

► $\mathbf{H}_k^0 = \left. \frac{\partial h_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}[k|k-1]}$

- Para cada $j = 0 \dots L$,

si $L=0$ entonces es un EKF, IKF suele dar mejor performance que el EKF, pero mucha demanda computacional

$$\mathbf{K}_k^j = \Sigma_{k|k-1} (\mathbf{H}_k^j)^* \left[\mathbf{H}_k^j \Sigma_{k|k-1} (\mathbf{H}_k^j)^* + \mathbf{R}_k \right]^{-1}$$

$$\hat{\mathbf{x}}^j[k|k] = \hat{\mathbf{x}}[k|k-1] + \mathbf{K}_k^j (\mathbf{y}(k) - h_k(\hat{\mathbf{x}}[k|k-1]))$$

$$\mathbf{H}_k^{j+1} = \left. \frac{\partial h_k(\mathbf{x})}{\partial \mathbf{x}} \right|_{\mathbf{x}=\hat{\mathbf{x}}^j[k|k]}$$

- ❷ Actualización de la estimación

$$\hat{\mathbf{x}}[k|k] = \hat{\mathbf{x}}^L[k|k] \quad \hat{\mathbf{x}}[k+1|k] = f_k(\hat{\mathbf{x}}[k|k])$$

$$\Sigma_{k|k} = \Sigma_{k|k-1} - \Sigma_{k|k-1} (\mathbf{H}_k^L)^* \left[\mathbf{H}_k^L \Sigma_{k|k-1} (\mathbf{H}_k^L)^* + \mathbf{R}_k \right]^{-1} \mathbf{H}_k^L \Sigma_{k|k-1}$$

$$\Sigma_{k+1|k} = \mathbf{F}_k \Sigma_{k|k} \mathbf{F}_k^* + \mathbf{G}_k \mathbf{Q}_k \mathbf{G}_k^*$$

- Tanto el EKF como el IFK necesitan el cómputo del jacobiano de $f_k()$ y de $h_k()$.
- Una alternativa que soluciona este problema es el llamado filtro UKF (*Unscented Kalman Filter*).
- Lamentablemente, si bien han habido muchos trabajos al respecto, no es claro en qué situaciones hay que utilizar uno u otro enfoque.

- ❶ *Linear Estimation*, Thomas Kailath, Ali H. Sayed, and Babak Hassibi Prentice Hall.
- ❷ *Optimal Filtering*, Anderson, B. D.O., and J. B. Moore, Prentice-Hall: Englewood Cliffs, NJ.
- ❸ *Linear System Theory and Design*, Chi-Tsong Chen , The Oxford Series in Electrical and Computer Engineering
- ❹ *Linear Systems*, Thomas Kailath, Prentice Hall,1980.