

Estimación y Predicción en Series Temporales

Filtro de Kalman

Departamento de Procesamiento de Señales

Instituto de Ingeniería Eléctrica
Facultad de Ingeniería

2022

Biblio:

- Brown and Hwang, *Introduction to Random Signals and Applied Kalman Filtering*, 3ra edición, 1996.
- Anderson and Moore, *Optimal Filtering*, 1979 (re-edición Dover 2005).

Hasta aquí estudiamos los filtros de Wiener y filtros adaptivos (LMS).

- El filtro de Wiener no se puede escribir en forma recursiva, ni tampoco fácilmente en situaciones de múltiples entradas y múltiples salidas
- El algoritmo LMS tiene un tiempo de convergencia relativamente lento. No utiliza toda la información de la forma eficiente.

En los '60 Kalman formuló el problema de filtrado de mínimo error cuadrático medio (MMSE) utilizando una representación en variables de estado.

- El resultado es un algoritmo recursivo que es óptimo en cada paso.
- Este filtro fue fácilmente implementable en aplicaciones de tiempo real y tuvo aplicaciones inmediatas en predicción, tracking, etc.

Existen variaciones y versiones más sofisticadas. Veremos

Formulación básica y notación

Suponemos que el proceso a estimar se puede modelizar como

$$\mathbf{x}(k+1) = \Phi(k)\mathbf{x}(k) + \mathbf{w}(k),$$

y que las observaciones o mediciones del proceso se realizan según

$$\mathbf{y}(k) = \mathbf{C}(k)\mathbf{x}(k) + \mathbf{v}(k).$$

- $\mathbf{x}(k)$ vector $n \times 1$ de estados del proceso.
- $\Phi(k)$ matriz $n \times n$ que relaciona $\mathbf{x}(k)$ a $\mathbf{x}(k+1)$ en ausencia de una fuerza o entrada exterior.
- $\mathbf{w}(k)$ vector $n \times 1$ de ruido blanco con covarianza conocida.
- $\mathbf{y}(k)$ vector $m \times 1$ de observaciones o medidas
- $\mathbf{C}(k)$ matriz $m \times n$ que relaciona los estados (sin ruido) y las observaciones.
- $\mathbf{v}(k)$ vector $m \times 1$ de ruido blanco con covarianza conocida.

Observaciones:

- Básicamente se propone ver al proceso como el resultado

Formulación básica y notación (cont.)

Suponemos además las siguientes covarianzas entre ruidos de estado y de observación:

$$\mathbb{E}[\mathbf{w}(k)\mathbf{w}(i)^T] = \begin{cases} \mathbf{Q}(k) & i = k \\ \mathbf{0} & i \neq k \end{cases}$$

$$\mathbb{E}[\mathbf{v}(k)\mathbf{v}(i)^T] = \begin{cases} \mathbf{R}(k) & i = k \\ \mathbf{0} & i \neq k \end{cases}$$

$$\mathbb{E}[\mathbf{v}(k)\mathbf{w}(i)^T] = \mathbf{0} \quad \forall i, k.$$

Suponemos que tenemos un estimado del proceso en el tiempo k , basado en todo nuestro conocimiento hasta el tiempo $k-1$: $\hat{\mathbf{x}}(k|k-1) = \hat{\mathbf{x}}_k^-$ Suponemos que conocemos la matriz de covarianza del estimado:

$$\mathbf{e}(k|k-1) = \mathbf{x}(k) - \hat{\mathbf{x}}(k|k-1)$$

$$\mathbf{e}_k^- = \mathbf{x}_k - \hat{\mathbf{x}}_k^-$$

$$\mathbf{P}(k|k-1) = \mathbb{E}[\mathbf{e}(k|k-1)\mathbf{e}^T(k|k-1)]$$

$$\mathbf{P}_k^- = \mathbb{E}[\mathbf{e}_k^- \mathbf{e}_k^{-T}]$$

Si comenzamos el problema sin mediciones, y si el proceso tiene media nula, en general tomamos el estimado inicial como $\mathbf{0}$ y como covarianza del error la de \mathbf{x}

Evolución del filtro, Ganancia de Kalman

Consideramos que tenemos el estimado $\hat{\mathbf{x}}_k^-$ y buscamos mejorarlo usando la info medida $\mathbf{y}(k)$ de la siguiente forma:

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{y}_k - \mathbf{C}_k\hat{\mathbf{x}}_k^-)$$

El objetivo del filtro de Kalman es encontrar \mathbf{K}_k para que la actualización sea óptima en el sentido del MSE. Empezamos introduciendo la covarianza del error del estimador actualizado, $\mathbf{x}(k)$:

$$\mathbf{P}_k = \mathbb{E}[\mathbf{e}_k \mathbf{e}_k^T] = \mathbb{E}[(\mathbf{x}_k - \hat{\mathbf{x}}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k)^T].$$

La idea es entonces obtener la matriz \mathbf{K} que minimize la varianza del error del estimador actualizado, es decir $\xi(k) = \text{Tr}[\mathbf{P}_k]$.

Ejercicio:

- 1 Escribir \mathbf{P}_k en función de \mathbf{K}_k .
- 2 Obtener la ganancia de Kalman minimizando $\xi(k)$.

Serán útiles las relaciones siguientes:

$$\frac{d}{d\mathbf{A}} \text{Tr}[\mathbf{A}\mathbf{B}] = \mathbf{B}^T \text{ (A, B cuadradas); } \frac{d}{d\mathbf{A}} \text{Tr}[\mathbf{A}\mathbf{C}\mathbf{A}^T] = 2\mathbf{A}\mathbf{C} \text{ (C simétrica).}$$

Obtención de la ganancia de Kalman

Tenemos:

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_k^-),$$

$$\mathbf{y}_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k.$$

de dónde

$$\begin{aligned}\mathbf{e}_k &= \mathbf{x}_k - \hat{\mathbf{x}}_k = \mathbf{x}_k - \hat{\mathbf{x}}_k^- - \mathbf{K}_k(\mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k - \mathbf{C}_k \hat{\mathbf{x}}_k^-) \\ &= (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k)(\mathbf{x}_k - \hat{\mathbf{x}}_k^-) - \mathbf{K}_k \mathbf{v}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbf{e}_k^- - \mathbf{K}_k \mathbf{v}_k.\end{aligned}$$

Luego, como $\mathbb{E}[\mathbf{e}_k^- \mathbf{v}_k^T] = 0$,

$$\begin{aligned}\mathbf{P}_k &= \mathbb{E}[\mathbf{e}_k \mathbf{e}_k^T] = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbb{E}[\mathbf{e}_k^- (\mathbf{e}_k^-)^T] (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k)^T + \mathbf{K}_k \mathbb{E}[\mathbf{v}_k \mathbf{v}_k^T] \mathbf{K}_k^T \\ &= (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbf{P}_k^- (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \\ &= \mathbf{P}_k^- - \mathbf{K}_k \mathbf{C}_k \mathbf{P}_k^- - (\mathbf{P}_k^-)^T \mathbf{C}_k^T \mathbf{K}_k^T + \mathbf{K}_k (\mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T + \mathbf{R}_k)\end{aligned}$$

Utilizando las reglas de derivación anteriores y observando que

$\text{Tr}[(\mathbf{P}_k^-)^T \mathbf{C}_k^T \mathbf{K}_k^T] = \text{Tr}[\mathbf{K}_k \mathbf{C}_k \mathbf{P}_k^-]$, obtenemos:

$$\frac{\partial \mathbf{P}_k}{\partial \mathbf{K}_k} = 0 \Leftrightarrow \boxed{\mathbf{K}_k = \mathbf{P}_k^- \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T + \mathbf{R}_k)^{-1}}$$

Matriz de covarianza óptima del estimador actualizado

Para esto sustituimos la expresión de la ganancia de Kalman en la expresión de la covarianza del error de estimación \mathbf{P}_k .
Tenemos:

$$\begin{aligned}\mathbf{P}_k &= (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbf{P}_k^- (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k)^T + \mathbf{K}_k \mathbf{R}_k \mathbf{K}_k^T \\ \mathbf{K}_k &= \mathbf{P}_k^- \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T + \mathbf{R}_k)^{-1}\end{aligned}$$

Luego,

$$\begin{aligned}\mathbf{P}_k &= \mathbf{P}_k^- - \mathbf{P}_k^- (\mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T + \mathbf{R}_k)^{-1} \mathbf{C}_k \mathbf{P}_k^- \\ &= \mathbf{P}_k^- - \mathbf{K}_k (\mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T + \mathbf{R}_k) \mathbf{K}_k^T \\ &= (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbf{P}_k^-\end{aligned}$$

Observaciones:

- Estas expresiones son obviamente válidas solo para la ganancia óptima de Kalman.
- Las expresiones son equivalentes aritméticamente, pero no numéricamente.

Actualización de los estimadores

Ahora necesitamos una forma de calcular $\hat{\mathbf{x}}_{k+1}$ y \mathbf{P}_{k+1}^- de forma de tener condiciones iniciales para la próxima observación \mathbf{y}_{k+1} .

Utilizando la dinámica del sistema y por ser \mathbf{w}_k blanco (y por lo tanto de media nula):

$$\hat{\mathbf{x}}_{k+1}^- = \Phi_k \hat{\mathbf{x}}_k.$$

Para calcular la covarianza del error:

$$\begin{aligned}\mathbf{e}_{k+1}^- &= \mathbf{x}_{k+1} - \hat{\mathbf{x}}_{k+1}^- = (\Phi_k \mathbf{x}_k + \mathbf{w}_k) - \Phi_k \hat{\mathbf{x}}_k \\ &= \Phi_k \mathbf{e}_k + \mathbf{w}_k\end{aligned}$$

Luego,

$$\begin{aligned}\mathbf{P}_{k+1}^- &= \mathbb{E}[\mathbf{e}_{k+1}^- (\mathbf{e}_{k+1}^-)^T] = \mathbb{E}[(\Phi_k \mathbf{e}_k + \mathbf{w}_k)(\Phi_k \mathbf{e}_k + \mathbf{w}_k)^T] \\ &= \Phi_k \mathbf{P}_k \Phi_k^T + \mathbf{Q}_k\end{aligned}$$

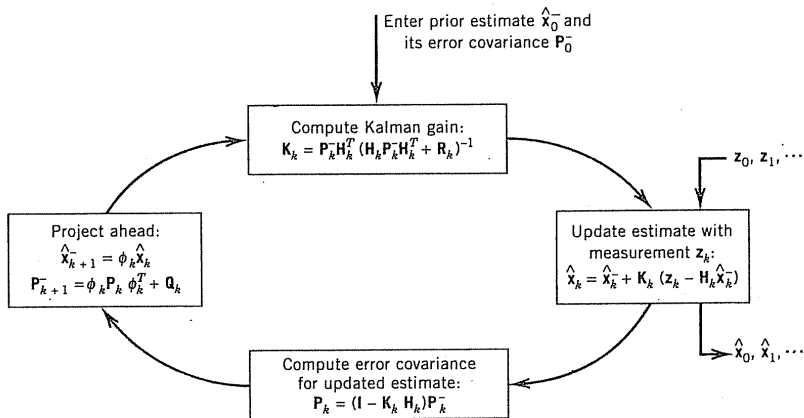
Filtro de Kalman: Resumen

- $\mathbf{x}_{k+1} = \Phi_k \mathbf{x}_k + \mathbf{w}_k$, $y_k = \mathbf{C}_k \mathbf{x}_k + \mathbf{v}_k$.
- \mathbf{x}_0^- , \mathbf{P}_0^- .

Para $k = 0, 1, 2, \dots$, calcular:

- 1 Ganancia de Kalman: $\mathbf{K}_k = \mathbf{P}_k^- \mathbf{C}_k^T (\mathbf{C}_k \mathbf{P}_k^- \mathbf{C}_k^T + \mathbf{R}_k)^{-1}$
- 2 Actualización del estimador: $\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_k^-)$
- 3 Covarianza del estimador: $\mathbf{P}_k = (\mathbf{I} - \mathbf{K}_k \mathbf{C}_k) \mathbf{P}_k^-$
- 4 Proyección a $k + 1$: $\hat{\mathbf{x}}_{k+1}^- = \Phi_k \hat{\mathbf{x}}_k$,
 $\mathbf{P}_{k+1}^- = \Phi_k \mathbf{P}_k \Phi_k^T + \mathbf{Q}_k$

Filtro de Kalman: Resumen (cont.)



En la figura, \mathbf{z}_k es y_k y \mathbf{H}_k es \mathbf{C}_k

Estabilidad del filtro de Kalman en régimen estacionario

Si el filtro de Kalman converge a un estado estacionario, se comporta como cualquier otro filtro digital y se puede estudiar su estabilidad mediante la transformada \mathcal{Z} .

Suponemos que alcanzamos un estado de ganancia constante:

$$\begin{aligned}\hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_k^- + \mathbf{K}_k(\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_k^-) \\ &= (\Phi_{k-1} - \mathbf{K}_k \mathbf{C}_k \Phi_{k-1}) \hat{\mathbf{x}}_{k-1} + \mathbf{K}_k \mathbf{y}_k \quad (\text{aquí usamos que } \hat{\mathbf{x}}_k^- = \Phi_{k-1} \hat{\mathbf{x}}_{k-1})\end{aligned}$$

Luego,

$$\begin{aligned}\hat{\mathbf{X}}_k(z) &= (\Phi_{k-1} - \mathbf{K}_k \mathbf{C}_k \Phi_{k-1}) z^{-1} \hat{\mathbf{X}}_k(z) + \mathbf{K}_k \hat{\mathbf{Y}}_k(z) \\ \Rightarrow \underbrace{[z\mathbf{I} - (\Phi_{k-1} - \mathbf{K}_k \mathbf{C}_k \Phi_{k-1})]}_{\text{describe los modos naturales}} \hat{\mathbf{X}}_k(z) &= z\mathbf{K}_k \hat{\mathbf{Y}}_k(z).\end{aligned}$$

describe los modos naturales

El polinomio característico es $\det[\cdot]$ y las raíces nos dan la información de estabilidad. El filtro es estable si y sólo si todas sus raíces caen dentro del círculo unidad.

- En muchas situaciones los procesos que consideramos están excitados por entradas determinísticas así como aleatorias.
- En estos casos, las ecuaciones del proceso son

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{\Phi}(k)\mathbf{x}(k) + \mathbf{w}(k) + \mathbf{\Gamma}(k)\mathbf{u}(k) \\ \mathbf{y}(k) &= \mathbf{C}(k)\mathbf{x}(k) + \mathbf{v}(k).\end{aligned}$$

- Veremos como esta representación en modelo de estados puede ser reformulada para hacer uso de los resultados de estimación óptima anteriores. La linealidad nos permite basarnos en el [principio de superposición](#).

Filtro de Kalman con entradas determinísticas (2)

Dinámica del desvío con respecto al proceso determinístico

- Definimos el siguiente **sistema determinístico**

$$\mathbf{x}^d(k+1) = \mathbf{\Phi}(k)\mathbf{x}^d(k) + \mathbf{\Gamma}(k)\mathbf{u}(k)$$

$$\mathbf{y}^d(k) = \mathbf{C}(k)\mathbf{x}^d(k).$$

- Luego estudiamos el desvío del **sistema estocástico** con respecto al determinístico:

$$\mathbf{x}(k+1) - \mathbf{x}^d(k+1) = \mathbf{\Phi}(k) \left(\mathbf{x}(k) - \mathbf{x}^d(k) \right) + \mathbf{w}(k)$$

$$\mathbf{y}(k) - \mathbf{y}^d(k) = \mathbf{C}(k) \left(\mathbf{x}(k) - \mathbf{x}^d(k) \right) + \mathbf{v}(k).$$

Definiendo $\Delta\mathbf{x}(k) = \mathbf{x}(k) - \mathbf{x}^d(k)$ y $\Delta\mathbf{y}(k) = \mathbf{y}(k) - \mathbf{y}^d(k)$, tenemos:

$$\Delta\mathbf{x}(k+1) = \mathbf{\Phi}(k)\Delta\mathbf{x}(k) + \mathbf{w}(k)$$

$$\Delta\mathbf{y}(k) = \mathbf{C}(k)\Delta\mathbf{x}(k) + \mathbf{v}(k).$$

De esta forma **la estimación óptima del desvío se rige por las ecuaciones de Kalman ya vistas.**

Filtro de Kalman con entradas determinísticas (3)

Ecuación de predicción

- Para recuperar la ecuación de predicción volvemos a la definición de $\Delta \mathbf{x}(k)$ y expandimos la ecuación de predicción del desvío:

$$\begin{aligned}\hat{\mathbf{x}}_{k+1}^- - \mathbf{x}_{k+1}^d &= \Phi_k \left(\hat{\mathbf{x}}_k - \mathbf{x}_k^d \right) \\ \Leftrightarrow \hat{\mathbf{x}}_{k+1}^- &= \mathbf{x}_{k+1}^d + \Phi_k \hat{\mathbf{x}}_k - \Phi_k \mathbf{x}_k^d \\ \Leftrightarrow \hat{\mathbf{x}}_{k+1}^- &= \Phi(k) \mathbf{x}^d(k) + \Gamma(k) \mathbf{u}(k) + \Phi_k \hat{\mathbf{x}}_k - \Phi_k \mathbf{x}_k^d \\ \Leftrightarrow \hat{\mathbf{x}}_{k+1}^- &= \Gamma(k) \mathbf{u}(k) + \Phi_k \hat{\mathbf{x}}_k.\end{aligned}$$

- **Obs.:** La matriz de covarianza de $\Delta \mathbf{x}_k^- = \hat{\mathbf{x}}_k^- - \mathbf{x}_k^d$ es la misma que la de \mathbf{x}_k^- ya que $\{\mathbf{x}_k^d\}$ es una señal determinística.

Filtro de Kalman con entradas determinísticas (4)

Ecuación de corrección

- Partimos de la ecuación de actualización para el desvío y operamos:

$$\begin{aligned}\Delta \hat{\mathbf{x}}_k &= \Delta \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\Delta \mathbf{y}_k - \mathbf{C}_k \Delta \hat{\mathbf{x}}_k^-) \\ \Leftrightarrow \hat{\mathbf{x}}_k - \mathbf{x}_k^d &= \hat{\mathbf{x}}_k^- - \mathbf{x}_k^d + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{y}_k^d - \mathbf{C}_k (\hat{\mathbf{x}}_k^- - \mathbf{x}_k^d) \right) \\ \Leftrightarrow \hat{\mathbf{x}}_k &= \hat{\mathbf{x}}_k^- + \mathbf{K}_k \left(\mathbf{y}_k - \mathbf{y}_k^d - \mathbf{C}_k (\hat{\mathbf{x}}_k^- - \mathbf{x}_k^d) \right).\end{aligned}$$

Como $\mathbf{y}_k^d = \mathbf{C}_k \mathbf{x}_k^d$, tenemos

$$\hat{\mathbf{x}}_k = \hat{\mathbf{x}}_k^- + \mathbf{K}_k (\mathbf{y}_k - \mathbf{C}_k \hat{\mathbf{x}}_k^-).$$

- El resto de las ecuaciones para la ganancia de Kalman y la actualización de la covarianza no cambian, ya que no dependen directamente de la estimación del estado.
- Por consiguiente, en suma, el único cambio al introducir entradas deterministas es la inclusión de estas entradas en la ecuación de predicción de estado (intuitivamente tiene sentido).