Filtro de Kalman

scardenol

14 de abril de 2022

1. Sistemas modelados en forma de espacio de estado [1]

1.1. Sistema lineal

La representación (modelo) más general en forma de espacio de estado de un sistema lineal con p entradas, q salidas y n variables de estado es la siguiente

$$\dot{\mathbf{x}}(t) = \mathbf{A}(t)\mathbf{x}(t) + \mathbf{B}(t)\mathbf{u}(t)$$
$$\mathbf{y}(t) = \mathbf{C}(t)\mathbf{x}(t) + \mathbf{D}(t)\mathbf{u}(t)$$

donde

- $\mathbf{x}(\cdot)$ se le llama el "vector de estado", $\mathbf{x}(t) \in \mathbb{R}^n$.
- $\mathbf{y}(\cdot)$ se le llama el "vector de salida", $\mathbf{y}(t) \in \mathbb{R}^q$.
- $\mathbf{u}(\cdot)$ se le llama el "vector de entradas (o de control)", $\mathbf{u}(t) \in \mathbb{R}^p$.
- $\mathbf{A}(\cdot)$ es la "matriz de estado (o del sistema)", $\dim[\mathbf{A}(\cdot)] = n \times n$.
- $\mathbf{B}(\cdot)$ es la "matriz de entradas", $\dim[\mathbf{B}(\cdot)] = n \times p$.
- $\mathbf{C}(\cdot)$ es la "matriz de salidas", $\dim[\mathbf{C}(\cdot)] = q \times n$.
- $\mathbf{D}(\cdot)$ es la "matriz de transmisión directa" conocida en inglés como *feedtrhough* o *feedforward*, es la que permite que la entrada del sistema afecte directamente la salida (si el modelo del sistema no tiene transmisión directa la matriz se convierte en la matriz cero), $\dim[\mathbf{D}(\cdot)] = q \times p$.
- $\dot{\mathbf{x}(t)} := \frac{d}{dt}\mathbf{x}(t)$.

1.1.1. Sistema lineal invariante en el tiempo

Observe que en la representación general todas las matrices $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{D}$ varían (dependen) del tiempo, sin embargo, es común trabajar con sistemas con matrices invariantes (que no dependen) del tiempo y representarlo con el modelo

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t)$$
$$\mathbf{y}(t) = \mathbf{C}\mathbf{x}(t) + \mathbf{D}\mathbf{u}(t)$$

Por otra parte, la variable que representa el tiempo t puede ser continua (con $t \in \mathbb{R}$) o discreta (con $t \in \mathbb{Z}$). En el caso discreto usualmente se usa la variable k y no t para denotar el tiempo.

1.1.2. Sistema lineal en tiempo discreto

La representación general de un sistema lineal puede escribirse en tiempo discreto k como

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{A}(k)\mathbf{x}(k) + \mathbf{B}(k)\mathbf{u}(k)\\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}(k)\mathbf{u}(k)\end{aligned}$$

- Recuerde que la discretización de $\dot{\mathbf{x}}(t)$ es $\mathbf{x}(k+1)$.

1.1.3. Sistema lineal en tiempo discreto e invariante en el tiempo

El sistema lineal con matrices invariantes en el tiempo puede escribirse en tiempo discreto k (explícita) como

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k)\\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \mathbf{D}\mathbf{u}(k)\end{aligned}$$

Nota:

- Recuerde que la discretización de $\dot{\mathbf{x}}(t)$ es $\mathbf{x}(k+1)$.

1.2. Controlabilidad

La condición de controlabilidad implica que es posible, mediante entradas admisibles, dirigir los estados desde cualquier valor inicial a cualquier valor final dentro de un intervalo de tiempo. **Un modelo de espacio de estado de un sistema lineal invariante en el tiempo es controlable si y solo si**

$$\text{rank}\begin{bmatrix}B & AB & A^2B & ... & A^{n-1}B\end{bmatrix} = n$$

1.3. Observabilidad

La observabilidad es la medida de que tan bien se pueden hacer inferencias de los estados estados internos del sistema al conocer sus salidas externas. **Un modelo de espacio de estado de un sistema lineal invariante en el tiempo es observable si y solo si**

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

Importante:

La observabilidad y controlabilidad de un sistema son "duales matemáticos", es decir, mientras que la controlabilidad implica que haya una entrada disponible que lleve cualquier estado inicial a cualquier estado final deseado, la observabilidad implica que el conocimiento de una trayectoria de salida provee información suficiente para predecir el estado inicial del sistema.

1.4. Sistema no lineal

La forma más general de un modelo en espacio de estado se escribe como las 2 funciones

$$\begin{aligned}\dot{\mathbf{x}}(t) &= \mathbf{f}(t, x(t), u(t))\\ \mathbf{y}(t) &= \mathbf{g}(t, x(t), u(t))\end{aligned}$$

2. Filtro de Kalman [2]

2.1. Definición y punto de partida

Es un observador del estado, lo que quiere decir que su funcionamiento dependerá de la **observabilidad** del estado. Considere como punto de partida el siguiente sistema (discretizado e invariante en el tiempo) en forma de espacio de estado

$$\begin{aligned}\mathbf{x}(k+1) &= \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) + \mathbf{V}E(k)\\ \mathbf{y}(k) &= \mathbf{C}\mathbf{x}(k) + \nu(k)\end{aligned}$$

donde $E(k)$ y $\nu(k)$ son ruidos blancos con media 0 y covarianza 0. Note que:

- El vector de estado $\mathbf{x}(k+1)$ tiene un nuevo término $\mathbf{V}E(k)$ que corresponde a una filtración $\mathbf{V}$ del ruido blanco $E(k)$ o "ruido coloreado".
- El vector de salida $\mathbf{y}$ no tiene el término $\mathbf{D}\mathbf{u}(k)$, lo que quiere decir que el sistema no tiene matriz de transmisión directa (*feedthrough*) $\mathbf{D}$.
- El vector de salida $\mathbf{y}(k)$ ahora depende del ruido blanco $\nu(k)$.

Adicionalmente, se denotan M y N a las matrices diagonales y constantes definidas como

$$M := \mathbb{E}[E(k)E^T(k)], \quad N := \mathbb{E}[n(k)n^T(k)]$$

2.2. Objetivo y desarrollo

Nos interesa encontrar una predicción $\hat{\mathbf{x}}(k+1)$ de los estados con el mínimo error (predicción óptima). Sea $\mathcal{E}$ una función de costo definida como

$$\mathcal{E} := \mathbb{E}\Big[\|\hat{\mathbf{x}}(k+1) - \mathbf{x}(k+1)\|_2^2\Big]$$

Desarrollando la expresión se obtiene

$$\begin{aligned}\mathcal{E} &= \mathbb{E}\Bigg[\Big(\hat{\mathbf{x}}(k+1) - \mathbf{x}(k+1)\Big)^T\Big(\hat{\mathbf{x}}(k+1) - \mathbf{x}(k+1)\Big)\Bigg]\\ &= \mathbb{E}\Bigg[\operatorname{tr}\Bigg[\Big(\hat{\mathbf{x}}(k+1) - \mathbf{x}(k+1)\Big)\Big(\hat{\mathbf{x}}(k+1) - \mathbf{x}(k+1)\Big)^T\Bigg]\Bigg]\\ &= \operatorname{tr}\Bigg[\mathbb{E}\Bigg[\Big(\hat{\mathbf{x}}(k+1) - \mathbf{x}(k+1)\Big)\Big(\hat{\mathbf{x}}(k+1) - \mathbf{x}(k+1)\Big)^T\Bigg]\Bigg]\\ &= \operatorname{tr}\Big[\mathbf{P}(k+1)\Big]\end{aligned}$$

Donde $\mathbf{P}(k+1)$ denota la matriz de covarianzas del estado. Ahora, sabemos que la verdadera dinámica del sistema está dada por

$$\mathbf{x}(k+1) = \mathbf{A}\mathbf{x}(k) + \mathbf{B}\mathbf{u}(k) + \mathbf{V}E(k)$$

pero como $E(k)$ es desconocido resulta natural estimarlo como su valor esperado $\mathbb{E}[E(k)] = 0$. Así, podemos proponer la estimación

$$\hat{\mathbf{x}}(k+1|k) = \mathbf{A}\hat{\mathbf{x}}(k) + \mathbf{B}\mathbf{u}(k) \tag{1}$$

Nota: la estimación $\hat{\mathbf{x}}(k+1|k)$ significa la estimación del vector $\hat{\mathbf{x}}(k+1)$ en el instante de tiempo (discreto) k.

Esta estimación conduce a la matriz de varianzas y covarianzas intermedia (o antes de la corrección) denotada por $\bar{\mathbf{P}}(k+1)$ y expresada como

$$\bar{\mathbf{P}}(k+1) = \mathbf{A}\mathbf{P}(k)\mathbf{A}^T + \mathbf{V}\mathbf{M}\mathbf{V}^T \tag{2}$$

Ahora, si se tiene una medición $\mathbf{y}(k+1)$ es posible realizar una corrección a la estimación $\hat{\mathbf{x}}(k+1|k)$ en el siguiente instante de tiempo $(k+1)$

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + \mathcal{K}(k+1)\Big(\mathbf{y}(k+1) - \mathbf{C}\hat{\mathbf{x}}(k+1|k)\Big) \tag{3}$$

Donde $\mathcal{K}$ representa un factor de corrección o peso, que representa la confianza que se tiene en las mediciones. Para determinar cuál es la mejor elección de $\mathcal{K}$ se deduce primero la expresión de la matriz de varianzas y covarianzas $\mathbf{P}(k+1)$ como

$$\mathbf{P}(k+1) = \Big(\mathbf{I} - \mathcal{K}(\mathbf{k+1})\mathbf{C}\Big)\bar{\mathbf{P}}(k+1)\Big(\mathbf{I} - \mathcal{K}(k+1)\mathbf{C}\Big)^T + \mathcal{K}(k+1)\mathbf{N}\mathcal{K}^T(k+1)$$

lo cual puede simplificarse a

$$\mathbf{P}(k+1) = \Big(\mathbf{I} - \mathcal{K}(k+1)\mathbf{C}\Big)\bar{\mathbf{P}}(k+1) \tag{4}$$

donde $\mathbf{I}$ es la matriz identidad de dimensiones adecuadas. Luego, la elección óptima de $\mathcal{K}$ está dada por la solución de la ecuación

$$\frac{\partial}{\partial \mathcal{K}}\mathcal{E} = 0$$

cuya solución es

$$\mathcal{K}(k+1) = \bar{\mathbf{P}}(k+1)\mathbf{C}^T\Big(\mathbf{C}\bar{\mathbf{P}}(k+1)\mathbf{C}^T + \mathbf{N}\Big)^{-1} \tag{5}$$

2.3. Algoritmo

El filtro de Kalman se compone entonces de 2 pasos

1. **Predicción** (*time update*)

$$\begin{aligned}\hat{\mathbf{x}}(k+1|k) &= \mathbf{A}\hat{\mathbf{x}}(k) + \mathbf{B}\mathbf{u}(k)\\ \bar{\mathbf{P}}(k+1) &= \mathbf{A}\mathbf{P}(k)\mathbf{A}^T + \mathbf{V}\mathbf{M}\mathbf{V}^T\end{aligned}$$

2. **Estimación** (*measurement update*)

$$\begin{aligned}\mathcal{K}(k+1) &= \bar{\mathbf{P}}(k+1)\mathbf{C}^T\Big(\mathbf{C}\bar{\mathbf{P}}(k+1)\mathbf{C}^T + \mathbf{N}\Big)^{-1}\\ \hat{\mathbf{x}}(k+1|k+1) &= \hat{\mathbf{x}}(k+1|k) + \mathcal{K}(k+1)\Big(\mathbf{y}(k+1) - \mathbf{C}\hat{\mathbf{x}}(k+1|k)\Big)\\ \mathbf{P}(k+1) &= \Big(\mathbf{I} - \mathcal{K}(k+1)\mathbf{C}\Big)\bar{\mathbf{P}}(k+1)\end{aligned}$$

2.3.1. Incógnitas

- $\mathbf{M}, \mathbf{N}, \mathbf{V}$.
- Condiciones iniciales.

Nota:

- Usualmente $\mathbf{V} = \mathbf{I}$.
- Se dan condiciones iniciales de los estados $\mathbf{x}(0)$.
- Se asume $\bar{\mathbf{P}}(0) = \mathbf{M}$.
- $\mathbf{M}, \mathbf{N}$ se determinan por estimaciones fuera de línea o **heurística**.

2.3.2. Conocidos

- $\mathbf{A}, \mathbf{B}, \mathbf{C}, \mathbf{u}, \mathbf{y}$

2.4. Selección de M y N

- Cuando $N \to \infty$ ocurre que: hay mucha incertidumbre (poca confianza) en las mediciones (salidas), el algoritmo solo debería estimar (no corregir), causando que $\mathcal{K}(k+1) = 0$.
- Cuando $N \to 0$ ocurre que: hay mucha confianza en las mediciones y $\mathcal{K} \to \mathbf{C}^{-1} \Rightarrow \hat{\mathbf{x}}(k+1|k+1) \to \mathbf{y}(k+1)$, lo que quiere decir que se desecha la predicción.

2.5. Convergencia

En general $\mathbf{P}$ y $\mathcal{K}$ tienden a ser constantes (cuando se descubre suficiente del proceso), entonces para evaluar la convergencia se grafica $\bar{\mathbf{P}} = \mathrm{tr}[\mathbf{P}]$ y $\|\mathcal{K}\|$.
Nota: recuerde que como $\mathbf{P}$ es la matriz de varianzas y covarianzas, en su diagonal se encuentra el error de predicción de cada estado.

2.6. Condición de observabilidad

La observabilidad implica que "a través de las mediciones que tengo puedo reconstruir los estados, en un tiempo finito". Como el filtro de Kalman es un observador del estado, su funcionamiento depende de la observabilidad del estado. El sistema usado en el filtro de Kalman es observable si y solo si

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

3. Filtro de Kalman extendido

En este caso se parte de un sistema no lineal (en tiempo continuo) modelado en espacio de estado como

$$\begin{aligned}
\dot{\mathbf{x}}(t) &= \mathbf{f}(x(t), u(t)) + \mathbf{V}(t)E(t) \\
\mathbf{y}(t) &= \mathbf{g}(x(t)) + \nu(t)
\end{aligned}$$

3.1. Discretización

Se aplica un método de discretización como por ejemplo el método de Euler y se obtiene

$$\begin{aligned}
\mathbf{x}(k+1) &= \mathbf{x}(k) + T_s\mathbf{f}(x(k), u(k)) + \mathbf{V}(t)E(t) \\
\mathbf{y}(k) &= \mathbf{g}(x(k)) + \nu(k)
\end{aligned}$$

donde T_s es el periodo de muestreo y se denotan las funciones

$$f_n(x(k), u(k)) := \mathbf{x}(k) + T_s\mathbf{f}(x(k), u(k)), \quad g_k(x(k)) := \mathbf{g}(x(k))$$

entonces el modelo discretizado resultante es

$$\begin{aligned}
\mathbf{x}(k+1) &= f_k(x(k), u(k)) + \mathbf{V}(t)E(t) \\
\mathbf{y}(k) &= g_k(x(k)) + \nu(k)
\end{aligned}$$

3.2. Linealización

Después de discretizar el modelo, se realiza un proceso de linalización resultando en

$$\begin{aligned}
\mathbf{A}(k) &= \left.\frac{\partial f_k(x(k), u(k))}{\partial \mathbf{x}(k)}\right|_{\mathbf{x}(k)=\hat{\mathbf{x}}(k), \quad u(k)=u(k)} \\
\mathbf{C}(k+1) &= \left.\frac{\partial g_{k+1}(x(k+1))}{\partial \mathbf{x}(k+1)}\right|_{\mathbf{x}(k+1)=\hat{\mathbf{x}}(k+1)}
\end{aligned}$$

3.3. Estimación y predicción

Se actualiza (corrección) con

$$\hat{\mathbf{x}}(k+1|k+1) = \hat{\mathbf{x}}(k+1|k) + \mathcal{K}(k+1)\Big(\mathbf{y}(k+1) - g_{k+1}\big(\hat{\mathbf{x}}(k+1|k)\big)\Big) \tag{6}$$

Se pronostica (predicción) con

$$\hat{\mathbf{x}}(k+1|k) = \hat{\mathbf{x}}(k|k) + f_k\big(\hat{\mathbf{x}}(k|k), \mathbf{u}(k)\big) \tag{7}$$

3.4. Algoritmo

El filtro de Kalman extendido se compone de los mismos pasos que el filtro de Kalman, pero reemplazando las nuevas fórmulas de estimación y predicción

1. **Predicción** (*time update*)

$$\begin{aligned}\hat{\mathbf{x}}(k+1|k) &= \hat{\mathbf{x}}(k|k) + f_k\big(\hat{\mathbf{x}}(k|k), \mathbf{u}(k)\big)\\ \bar{\mathbf{P}}(k+1) &= \mathbf{A}\mathbf{P}(k)\mathbf{A}^T + \mathbf{V}\mathbf{M}\mathbf{V}^T\end{aligned}$$

2. **Estimación** (*measurement update*)

$$\begin{aligned}\mathcal{K}(k+1) &= \bar{\mathbf{P}}(k+1)\mathbf{C}^T\Big(\mathbf{C}\bar{\mathbf{P}}(k+1)\mathbf{C}^T + \mathbf{N}\Big)^{-1}\\ \hat{\mathbf{x}}(k+1|k+1) &= \hat{\mathbf{x}}(k+1|k) + \mathcal{K}(k+1)\Big(\mathbf{y}(k+1) - g_{k+1}\big(\hat{\mathbf{x}}(k+1|k)\big)\Big)\\ \mathbf{P}(k+1) &= \Big(\mathbf{I} - \mathcal{K}(k+1)\mathbf{C}\Big)\bar{\mathbf{P}}(k+1)\end{aligned}$$

Referencias

[1] wikipedia, “State-space representation.” https://en.wikipedia.org/wiki/State-space_representation, 2022.

[2] “Notas de clase.”