

Tema 3: Estimación en sistemas lineales discretos

1. Modelo de espacio de estados
2. Predicción óptima
3. Filtrado óptimo
4. Suavizamiento óptimo

Hemos estudiado que en el caso gaussiano el estimador óptimo, $\hat{x}(k/j)$, bajo cualquier función de pérdida admisible, es el estimador lineal de menor error cuadrático medio y tiene la siguiente expresión

$$\hat{x}(k/j) = E[x(k)] + E[x(k)Z_j^T] \left(E[(Z_j - E[Z_j])(Z_j - E[Z_j])^T] \right)^{-1} (Z_j - E[Z_j])$$

donde Z_j denota el vector $(j+1)m$ -dimensional $Z_j = (z^T(0), \dots, z^T(j))^T$.

Desde el punto de vista teórico, el problema de estimación está resuelto; sin embargo, la utilidad práctica de este estimador es bastante limitada ya que como puede verse en su expresión, es necesario invertir una matriz de dimensión $(j+1)m \times (j+1)m$, siendo $j+1$ el número de medidas y m el número de componentes del vector de medidas. Si j varía, y se desea realizar estimaciones “on-line”, esto es, procesar la información conforme ésta se va haciendo disponible, la aplicación de la expresión anterior para generar el estimador óptimo se vuelve impracticable.

Desde el punto de vista de las aplicaciones, lo deseable sería disponer de algoritmos eficientes y prácticos que procesen secuencialmente las observaciones y que proporcionen estimaciones en cada instante a partir de las previas (es decir, algoritmos recursivos). En este tema desarrollaremos tales algoritmos para los problemas de predicción, filtrado y suavizamiento.

1 Modelo de espacio de estados

Consideremos un sistema lineal en tiempo discreto, cuya modelización matemática se realiza mediante las siguientes ecuaciones.

1.1 Ecuación del estado

Es una ecuación vectorial en diferencias, también denominada ecuación de transición del estado

$$x(k+1) = \Phi(k+1, k)x(k) + \Gamma(k+1, k)w(k), \quad k \geq 0; \quad x(0) = x_0$$

donde

- $\{x(k); k \geq 0\}$ es un proceso estocástico n -dimensional que representa el *estado* del sistema.
- $\{w(k); k \geq 0\}$ es un proceso estocástico r -dimensional; el vector $w(k)$ se denomina *vector de perturbación aleatoria del estado* o *ruido del estado*.
- $\Phi(k+1, k)$ y $\Gamma(k+1, k)$ son matrices determinísticas conocidas, de dimensiones apropiadas, denominadas *matriz de transición del estado* y *matriz de transición o de ponderación de la perturbación aleatoria*, respectivamente.

1.2 Ecuación de observación

Se supone que el vector estado no es directamente accesible, sino que sólo disponemos de una sucesión de observaciones, relacionadas con el estado y perturbadas por un ruido. La ecuación de observación es una ecuación vectorial,

$$z(k) = H(k)x(k) + v(k), \quad k \geq 0$$

donde

- $\{z(k); k \geq 0\}$ es un proceso estocástico m -dimensional; $z(k)$ representa el *vector de medidas u observaciones* en el instante k .
- $\{v(k); k \geq 0\}$ es un proceso estocástico m -dimensional; el vector $v(k)$ se denomina *vector de perturbación aleatoria de la observación* o *ruido de la observación*.
- $H(k)$ es una matriz determinística conocida, de dimensión $m \times n$, llamada *matriz de medidas o de observaciones*.

1.3 Hipótesis

Sobre los ruidos y el estado inicial se imponen las siguientes hipótesis:

1. El estado inicial, x_0 , es un vector aleatorio n -dimensional gaussiano con media cero y matriz de covarianzas $E[x_0x_0^T] = P_0$.
2. El proceso $\{w(k); k \geq 0\}$ es una sucesión ruido blanco gaussiana, centrada, con matrices de covarianzas $E[w(k)w^T(k)] = Q(k)$, $k \geq 0$.
3. El proceso $\{v(k); k \geq 0\}$ es un proceso ruido blanco gaussiano, centrado y con covarianzas $E[v(k)v^T(k)] = R(k)$, $k \geq 0$, siendo $R(k)$ una matriz definida positiva.
4. El estado inicial, x_0 , y los ruidos aditivos, $\{w(k); k \geq 0\}$ y $\{v(k); k \geq 0\}$, son mutuamente independientes.

Las matrices $\Phi(k+1, k)$, $\Gamma(k+1, k)$, $Q(k)$, $H(k)$ y $R(k)$ se denominan *matrices del sistema*. Si estas matrices no dependen de k , se dice que el sistema es *invariante en el tiempo*.

Las ecuaciones del estado y de la observación, junto con las hipótesis sobre los ruidos y el estado inicial, se conocen como *modelo de espacio de estados*.

Nota. A veces no interesa, o incluso no es posible, estimar el estado del sistema propiamente dicho, sino una función del mismo. En general, se denomina *señal* a una función $y(k)$ que contiene información relevante sobre el estado del sistema. A menudo, la señal es una función lineal del estado, es decir, una función de la forma $y(k) = A(k)x(k)$, siendo $A(k)$ una matriz determinística conocida. En tal caso, los estimadores mínimo cuadráticos del estado y de la señal verifican la siguiente relación

$$\hat{y}(k/j) = A(k)\hat{x}(k/j)$$

y, por tanto, podemos centrarnos en el estudio del problema de estimación del estado.

2 Predicción óptima

Vamos a establecer el algoritmo de predicción óptima para el sistema definido en la sección anterior; es decir, el algoritmo para obtener los estimadores $\hat{x}(k/j)$ con $j = 0, 1, \dots, k-1$, $k = 0, 1, \dots$. Estudiaremos también las propiedades del proceso error, $\{\tilde{x}(k/j); k > j\}$, y el comportamiento de su matriz de covarianzas, $P(k/j) = E[\tilde{x}(k/j)\tilde{x}^T(k/j)]$.

Tema 3: Estimación en sistemas lineales discretos

Las hipótesis de gaussianidad e independencia mutua (hipótesis 1-4) exigidas al sistema garantizan que el filtro óptimo de $x(j)$ para $j = 0, 1, \dots$ es

$$\hat{x}(j/j) = E[x(j)/z(0), \dots, z(j)].$$

Es claro que $\hat{x}(j/j)$ tiene media cero y, ya que es combinación lineal de los vectores gaussianos $z(0), \dots, z(j)$, es también un vector aleatorio gaussiano. Análogamente, el error del filtro, $\tilde{x}(j/j) = x(j) - \hat{x}(j/j)$, es un vector aleatorio gaussiano centrado.

En el siguiente teorema se establece el resultado fundamental sobre predicción óptima.

Teorema 1. *Si el filtro óptimo, $\hat{x}(j/j)$, y la matriz de covarianzas, $P(j/j)$, del correspondiente error de filtrado, $\tilde{x}(j/j) = x(j) - \hat{x}(j/j)$, son conocidos para algún $j = 0, 1, \dots$, entonces, para todo $k > j$, se tiene que*

(i) *El predictor óptimo, $\hat{x}(k/j)$, para toda función de pérdida admisible, es*

$$\hat{x}(k/j) = \Phi(k, j)\hat{x}(j/j).$$

(ii) *El proceso $\{\tilde{x}(k/j); k = j + 1, j + 2, \dots\}$ definido por*

$$\tilde{x}(k/j) = x(k) - \hat{x}(k/j)$$

es gaussiano, tiene media cero y su matriz de covarianzas verifica la siguiente relación

$$\begin{aligned} P(k/j) &= \Phi(k, j)P(j/j)\Phi^T(k, j) \\ &\quad + \sum_{i=j+1}^k \Phi(k, i)\Gamma(i, i-1)Q(i-1)\Gamma^T(i, i-1)\Phi^T(k, i). \end{aligned}$$

Notas:

- La demostración del Teorema 1 se basa en la siguiente representación del vector estado (ver Tema 1):

$$x(k) = \Phi(k, j)x(j) + \sum_{i=j+1}^k \Phi(k, i)\Gamma(i, i-1)w(i-1)$$

donde Φ y Γ son las matrices de transición del estado y de la perturbación, respectivamente, que se suponen conocidas por la descripción del modelo.

- Teniendo en cuenta las propiedades de la matriz de transición Φ (Tema 1) se obtiene una expresión alternativa para la matriz de covarianzas del error de predicción:

$$P(k/j) = \Phi(k, k-1)P(k-1/j)\Phi^T(k, k-1) + \Gamma(k, k-1)Q(k-1)\Gamma^T(k, k-1)$$

para $k = j + 1, j + 2, \dots$

- El predictor óptimo es único en virtud de la unicidad de la esperanza condicionada.
- El Teorema 1 es de uso limitado para realizar predicciones, pues es necesario conocer $\hat{x}(j/j)$ y $P(j/j)$ para algún $j \geq 0$.

2.1 Predicción en una etapa

Un caso particular dentro del ámbito de la predicción, y que merece especial interés, es la llamada *predicción óptima adelantada una etapa* o, simplemente, *predicción en una etapa* ($j = k - 1$). En este caso, el Teorema 1 proporciona el siguiente resultado que, como veremos, será muy útil en el desarrollo de las ecuaciones del filtro óptimo.

Corolario. *Si el filtro óptimo, $\hat{x}(k/k)$, y la matriz de covarianzas del error de filtrado, $P(k/k)$, se conocen para algún $k \geq 0$, entonces*

(i) *El predictor óptimo en una etapa, bajo cualquier función de pérdida admisible, es*

$$\hat{x}(k/k-1) = \Phi(k, k-1)\hat{x}(k-1/k-1), \quad k > 0.$$

Para $k = 0$, se tiene

$$\hat{x}(0/-1) = E[x_0 / \text{sin medidas}] = E[x_0] = 0.$$

(ii) *El proceso estocástico $\{\tilde{x}(k/k-1); k \geq 0\}$, definido por*

$$\tilde{x}(k/k-1) = x(k) - \hat{x}(k/k-1)$$

es gaussiano, tiene media cero y su matriz de covarianzas viene dada por la expresión

$$\begin{aligned} P(k/k-1) &= \Phi(k, k-1)P(k-1/k-1)\Phi^T(k, k-1) \\ &\quad + \Gamma(k, k-1)Q(k-1)\Gamma^T(k, k-1). \end{aligned}$$

Para $k = 0$, se tiene

$$P(0/-1) = E[\tilde{x}(0/-1)\tilde{x}^T(0/-1)] = E[x_0x_0^T] = P_0$$

donde la matriz P_0 se supone conocida por la descripción del sistema.

Nota. Al igual que el Teorema 1, este corolario es de uso limitado para realizar predicciones, ya que es necesario conocer $\hat{x}(k/k)$ y $P(k/k)$, para algún $k \geq 0$.

En la siguiente sección, abordaremos el problema de filtrado óptimo. Veremos que predicción en una etapa y filtrado están relacionados, en el sentido de que puede determinarse el filtro, dado el predictor en una etapa, y viceversa.

3 Filtrado óptimo

Como ya hemos indicado, desde el punto de vista de las aplicaciones, lo deseable sería disponer de algoritmos eficientes y prácticos que procesen secuencialmente las observaciones y proporcionen estimaciones en cada instante en base a las estimaciones en el instante anterior. El más significativo de estos algoritmos es el célebre *Filtro de Kalman*. Se trata de un algoritmo recursivo para el cálculo del estimador óptimo del vector estado de un sistema dinámico discreto en un instante determinado, basado en la información disponible hasta ese instante. Este filtro permite realizar estimaciones “on-line”, es decir, actualizar en cada iteración el valor del estimador conforme se va haciendo disponible la información, sin necesidad de repetir todos los cálculos. Además, proporciona una medida de la bondad de la estimación, a través de la matriz de covarianzas del error de estimación.

3.1 Filtro de Kalman

En 1960, Kalman obtuvo el algoritmo recursivo para el problema de estimación mínimos cuadráticos haciendo uso de la técnica de proyecciones ortogonales. Inicialmente, Kalman consideró un sistema en el que la ecuación de la observación no estaba perturbada por ruido; no obstante, aquí vamos a presentar el algoritmo de filtrado para los sistemas descritos en la Sección 1. Señalemos que, si se suprime la hipótesis de normalidad del estado inicial y de los ruidos, no hay garantía de que el filtro proporcione la esperanza condicionada del vector estado dadas las observaciones, si bien, como se ha probado en el Tema 2, podemos asegurar que el filtro es el óptimo, en el sentido de mínimos cuadrados, dentro de la clase de todos los estimadores lineales.

Para desarrollar el algoritmo de filtrado supondremos que, en cada instante k , disponemos de observaciones hasta dicho instante, $\{z(0), \dots, z(k)\}$.

Por las hipótesis de gaussianidad e independencia mutua (hipótesis 1-4), podemos afirmar que el filtro óptimo de $x(k)$ basado en las observaciones hasta el instante k es

$$\hat{x}(k/k) = E[x(k)/z(0), \dots, z(k)],$$

y ya que este estimador coincide con el lineal de menor error cuadrático medio, a partir de la ecuación de Wiener-Hopf (Tema 2) se obtiene que

$$\hat{x}(k/k) = \hat{x}(k/k - 1) + K(k)[z(k) - \hat{z}(k/k - 1)],$$

donde $\hat{z}(k/k - 1)$ es el estimador lineal de menor error cuadrático medio de $z(k)$ basado en $z(0), \dots, z(k - 1)$. Aplicando de nuevo la ecuación de Wiener-Hopf, es inmediato ver que

$$\hat{z}(k/k - 1) = H(k)\hat{x}(k/k - 1), \quad k > 0.$$

Para $k = 0$ se tiene

$$\hat{z}(0/k - 1) = E[z(0)/\sin \text{medidas}] = E[z(0)] = 0.$$

Demostración de la expresión del filtro óptimo

Expresamos $\hat{x}(k/k)$ y $\hat{z}(k/k - 1)$ como función lineal de las observaciones correspondientes:

$$\hat{x}(k/k) = \sum_{i=0}^k K(k,i)z(i), \quad \hat{z}(k/k - 1) = \sum_{i=0}^{k-1} F(k,i)z(i).$$

Aplicando el LPO obtenemos las ecuaciones de Wiener-Hopf:

$$E[x(k)z^T(\sigma)] = \sum_{i=0}^k K(k,i)E[z(i)z^T(\sigma)], \quad \sigma = 0, \dots, k$$

$$E[z(k)z^T(\sigma)] = \sum_{i=0}^{k-1} F(k,i)E[z(i)z^T(\sigma)], \quad \sigma = 0, \dots, k-1,$$

y, por lo tanto, para $\sigma = 0, \dots, k-1$, se tiene:

$$\begin{aligned} E[x(k)z^T(\sigma)] &= \sum_{i=0}^{k-1} K(k,i)E[z(i)z^T(\sigma)] + K(k,k)E[z(k)z^T(\sigma)] \\ &= \sum_{i=0}^{k-1} K(k,i)E[z(i)z^T(\sigma)] + \sum_{i=0}^{k-1} K(k,k)F(k,i)E[z(i)z^T(\sigma)] \\ &= \sum_{i=0}^{k-1} (K(k,i) + K(k,k)F(k,i))E[z(i)z^T(\sigma)], \quad \sigma = 0, \dots, k-1. \end{aligned}$$

Entonces, $\sum_{i=0}^{k-1} (K(k,i) + K(k,k)F(k,i))z(i)$ es una función lineal de $z(0), \dots, z(k-1)$ que satisface la ecuación de Wiener-Hopf correspondiente a $\hat{x}(k/k - 1)$ y, por tanto:

$$\begin{aligned} \hat{x}(k/k - 1) &= \sum_{i=0}^{k-1} (K(k,i) + K(k,k)F(k,i))z(i) \\ &= \sum_{i=0}^{k-1} K(k,i)z(i) + K(k,k) \sum_{i=0}^{k-1} F(k,i)z(i) \\ &= \hat{x}(k/k) - K(k,k)z(k) + K(k,k)\hat{z}(k/k - 1). \end{aligned}$$

Despejando $\hat{x}(k/k)$, tenemos la expresión del filtro.

Con estos preliminares, estamos en condiciones de enunciar el resultado fundamental de la teoría de filtrado óptimo para sistemas lineales gaussianos, el célebre *Filtro de Kalman*.

Teorema 2 (Filtro de Kalman). *Consideremos el sistema definido en la Sección 1.*

1. *El filtro óptimo, $\hat{x}(k/k)$, viene dado por la relación*

$$\begin{aligned}\hat{x}(k/k) &= \Phi(k, k-1)\hat{x}(k-1/k-1) \\ &\quad + K(k)[z(k) - H(k)\Phi(k, k-1)\hat{x}(k-1/k-1)], \quad k > 0 \\ \hat{x}(0/0) &= K(0)z(0).\end{aligned}$$

2. *$K(k)$ es una matriz $n \times m$, denominada matriz de ganancia del filtro y especificada por las siguientes expresiones*

$$K(k) = P(k/k-1)H^T(k)[H(k)P(k/k-1)H^T(k) + R(k)]^{-1}, \quad k \geq 0 \quad (1)$$

$$\begin{aligned}P(k/k-1) &= \Phi(k, k-1)P(k-1/k-1)\Phi^T(k, k-1) \\ &\quad + \Gamma(k, k-1)Q(k-1)\Gamma^T(k, k-1), \quad k > 0 \\ P(0/-1) &= P_0\end{aligned} \quad (2)$$

siendo

$$P(k/k) = [I - K(k)H(k)]P(k/k-1), \quad k \geq 0. \quad (3)$$

3. *El proceso estocástico $\{\tilde{z}(k/k); k \geq 0\}$ definido por*

$$\tilde{z}(k/k) = z(k) - \hat{x}(k/k)$$

es gaussiano, con media cero y matriz de covarianzas dada por (3).

3.2 Proceso innovación

Las ecuaciones del filtro de Kalman se basan en los predictores en una etapa de las observaciones, $\{\tilde{z}(k/k-1); k \geq 0\}$. El proceso error asociado, $\{\tilde{z}(k/k-1); k \geq 0\}$, definido por

$$\tilde{z}(k/k-1) = z(k) - \hat{x}(k/k-1), \quad k \geq 0,$$

se denomina *proceso innovación* o *residual*; para cada $k \geq 0$, $\tilde{z}(k/k-1)$ es la diferencia entre la medida observada en el instante k y la estimación de dicha medida basada en las observaciones anteriores hasta el instante $k-1$. La innovación representa así la nueva información, no contenida en $z(0), \dots, z(k-1)$, que proporciona la medida $z(k)$ para

estimar el estado $x(k)$. El siguiente resultado establece una importante propiedad de la sucesión de innovaciones.

Teorema 3. *La sucesión de innovaciones $\{\tilde{z}(k/k - 1); k \geq 0\}$ es un proceso blanco, gaussiano, centrado y con matriz de covarianzas*

$$E [\tilde{z}(k/k - 1)\tilde{z}^T(k/k - 1)] = H(k)P(k/k - 1)H^T(k) + R(k).$$

Señalemos, por último, que el proceso innovación se puede emplear para reformular el modelo de espacio de estados. Esta nueva formulación, expresada en términos de predicciones, se obtiene combinando las ecuaciones del predictor y el filtro, lo que proporciona una ecuación de transición para el predictor:

$$\hat{x}(k + 1/k) = \Phi(k + 1, k)\hat{x}(k/k - 1) + G(k)\tilde{z}(k/k - 1)$$

donde $G(k) = \Phi(k + 1, k)K(k)$.

La ecuación de medidas puede expresarse como:

$$z(k) = H(k)\hat{x}(k/k - 1) + \tilde{z}(k/k - 1),$$

y esta nueva formulación recibe el nombre de *forma innovadora* del modelo de espacio de estados definido en la Sección 1.

3.3 Aspectos computacionales

Una de las principales características del filtro de Kalman es su carácter recursivo. Supongamos que conocemos $\hat{x}(k/k)$, para algún k y que queremos determinar $\hat{x}(k + 1/k + 1)$. El ciclo computacional a seguir es el siguiente:

1. Multiplicar el estimador $\hat{x}(k/k)$ por la matriz de transición, $\Phi(k + 1, k)$. El resultado es el predictor $\hat{x}(k + 1/k)$.
2. Multiplicar $\hat{x}(k + 1/k)$ por $H(k + 1)$, obteniendo $\tilde{z}(k + 1/k)$, que, restado de la observación actual, $z(k + 1)$, proporciona la innovación $\tilde{z}(k + 1/k)$.
3. Multiplicar la innovación por la matriz de ganancia, $K(k + 1)$, y sumar el resultado a $\hat{x}(k + 1/k)$, obteniendo $\hat{x}(k + 1/k + 1)$.
4. Almacenar $\hat{x}(k + 1/k + 1)$ y repetir el ciclo cuando dispongamos de la nueva observación, $z(k + 2)$.

Observemos que el filtro puede interpretarse como un procedimiento de corrección del predictor en una etapa. En otras palabras, lo que se hace es añadir el término de corrección $K(k+1)\tilde{z}(k+1/k)$ al predictor $\hat{x}(k+1/k)$, para determinar el filtro.

A veces, interesa predecir el estado en un instante $k+m$ ($m \geq 0$), a partir de las observaciones hasta el instante k . En tales circunstancias, el Teorema 1 garantiza que

$$\hat{x}(k+m/k) = \Phi(k+m, k)\hat{x}(k/k).$$

Seguidamente, consideramos el cálculo de la matriz de ganancias, $K(k+1)$, y las matrices de covarianzas, $P(k+1/k)$ y $P(k+1/k+1)$. Un adecuado ciclo computacional es el siguiente:

1. Dadas $P(k/k)$, $Q(k)$, $\Phi(k+1, k)$, $\Gamma(k+1, k)$, calculamos $P(k+1/k)$, mediante la ecuación (2).
2. $P(k+1/k)$, $H(k+1)$, y $R(k+1)$ se sustituyen en la ecuación (1), obteniendo la matriz de ganancias, $K(k+1)$.
3. $P(k+1/k)$, $K(k+1)$ y $H(k+1)$ se sustituyen en la ecuación (3) para determinar $P(k+1/k+1)$, que se almacenará para repetir el ciclo.

Notemos, por último, que el cálculo de la matriz de covarianzas del error, $P(k/k)$, es independiente de la obtención del filtro, $\hat{x}(k/k)$. Esto permite obtener una medida de la bondad de la estimación sin necesidad de calcular explícitamente el estimador.

3.4 Programa en MATLAB: Filtro de Kalman

En este apartado realizamos un programa en MATLAB que, en cada iteración, simula el estado que se desea estimar, así como la observación correspondiente, y proporciona, tanto las estimaciones de predicción y filtrado, como las matrices de covarianzas de los errores asociados a los estimadores. Para ilustrar gráficamente los resultados, el programa también llevará a cabo la representación gráfica simultánea del estado, las observaciones y dichas estimaciones, así como las varianzas de los errores de predicción y filtrado. Concretamente, el programa se ha realizado para el siguiente modelo AR(1) escalar:

Ecuación del estado

$$x(k+1) = \Phi x(k) + w(k), \quad k \geq 0; \quad x(0) = x_0$$

donde x_0 , es una variable gaussiana con media cero y varianza P_0 y $\{w(k); k \geq 0\}$ es una sucesión ruido blanco gaussiana, centrada con varianzas $Q(k)$, $\forall k \geq 0$.

En el programa presentado se supone $\Phi = 0.95$, $P_0 = 1$ y $Q(k) = 0.1$, $\forall k \geq 0$.

- *Generación del estado*

$\Phi = 0.95;$

$P0 = 1; \quad x0 = sqrt(P0) * randn(1, 1);$

$Q0 = 0.1; \quad w0 = sqrt(Q0) * randn(1, 1);$

$x(1) = \Phi * x0 + w0;$

for $k = 1 : \text{número de iteraciones}$

$Q(k) = 0.1; \quad w(k) = sqrt(Q(k)) * randn(1, 1);$

$x(k + 1) = \Phi * x(k) + w(k);$

end

Ecuación de observación

$$z(k) = H(k)x(k) + v(k), \quad k \geq 0$$

donde $\{v(k); k \geq 0\}$ es un proceso ruido blanco gaussiano, centrado y con varianzas $R(k)$, $\forall k \geq 0$. Supondremos $H(k) = 1$, $\forall k \geq 0$ y $R(k) = 0.5$, $\forall k \geq 0$.

- *Generación de las observaciones*

$H0 = 1;$

$R0 = 0.5; \quad v0 = sqrt(R0) * randn(1, 1);$

$z0 = H0 * x0 + v0;$

for $k = 1 : \text{número de iteraciones}$

$H(k) = 1;$

$R(k) = 0.5; \quad v(k) = sqrt(R(k)) * randn(1, 1);$

$z(k) = H(k) * x(k) + v(k);$

end

Algoritmo de filtrado de Kalman

- *Notación:* fil=filtro, Cfil=Covarianza del filtro,
pred=predictor, Cpred=Covarianza del predictor

- *Condiciones iniciales:*

$$C_{pred0} = P_0;$$

$$K_0 = C_{pred0} * H_0 * [H_0^2 * P_0 + R_0]^{-1};$$

$$fil0 = K_0 * z_0;$$

$$C_{fil0} = [1 - K_0 * H_0] * C_{pred0};$$

$$pred(1) = \Phi * fil0;$$

$$C_{pred}(1) = \Phi^2 * C_{fil0} + Q_0;$$

- for $k = 1$: número de iteraciones

$$K(k) = C_{pred}(k) * H(k) * [H^2(k) * C_{pred}(k) + R(k)]^{-1};$$

$$fil(k) = pred(k) + K(k) * [z(k) - H(k) * pred(k)];$$

$$C_{fil}(k) = [1 - K(k) * H(k)] * C_{pred}(k);$$

$$pred(k + 1) = \Phi * fil(k);$$

$$C_{pred}(k + 1) = \Phi^2 * C_{fil}(k) + Q(k);$$

end

- *Gráfica del estado, la observación, el predictor y el filtro:*

for $k = 1$: número de iteraciones

$$time(k) = k;$$

$$px(k) = x(k);$$

$$pz(k) = z(k);$$

$$ppred(k) = pred(k);$$

$$pfil(k) = fil(k);$$

end

$$graf1 = plot(time, px, 'k', time, pz, 'g', time, ppred, 'r', time, pfil, 'b');$$

- *Gráfica de las varianzas de los errores de predicción y filtrado:*

for $k = 1$: número de iteraciones

$$\begin{aligned} time(k) &= k; \\ pCpred(k) &= Cpred(k); \\ pCfil(k) &= Cfil(k); \end{aligned}$$

end

$$graf2 = plot(time, pCpred, 'r : ', time, pCfil, 'b : ');$$

4 Suavizamiento óptimo

En el problema de filtrado, para estimar el estado en un instante k , se dispone de las observaciones hasta dicho instante, $\{z(0), \dots, z(k)\}$; esto es, la obtención del filtro $\hat{x}(k/k)$ se realiza en el propio instante k . Sin embargo, en numerosas situaciones, puede existir un retraso de N unidades de tiempo entre el instante k y el instante en el que se estima $x(k)$; durante dicho retraso dispondremos de nuevas medidas $\{z(k+1), \dots, z(k+N)\}$, que pueden usarse para obtener el estimador de $x(k)$. Dicho estimador, basado en las observaciones $\{z(0), \dots, z(k+N)\}$, se denomina *estimador de suavizamiento* o *suavizador*.

Para estudiar el problema de suavizamiento óptimo para el sistema definido en la Sección 1, supondremos que en cada instante k se dispone de observaciones hasta un instante j posterior a k , $\{z(0), \dots, z(j)\}$ con $j > k$. De nuevo, las hipótesis de gaussianidad e independencia mutua (hipótesis 1-4) exigidas al sistema garantizan que el suavizador óptimo $\hat{x}(k/j)$, $j > k$, es

$$\hat{x}(k/j) = E [x(k)/z(0), \dots, z(j)].$$

Aunque los dos índices k y j pueden ser variables, en la mayoría de las aplicaciones del suavizamiento no será necesario obtener los estimadores $\hat{x}(k/j)$ para todo k y para todo j . Por tanto, es necesario hacer una clasificación de los estimadores de suavizamiento de acuerdo a las posibles relaciones entre ambos índices; históricamente, se han estudiado tres tipos particulares de problemas de suavizamiento, *suavizamiento punto fijo* (obtención de $\hat{x}(k/j)$ para k fijo y $j = k+1, k+2, \dots$), *suavizamiento intervalo fijo* (obtención de $\hat{x}(k/j)$ para j fijo y cualquier k en el intervalo $0 \leq k \leq j$), y *suavizamiento retraso fijo* (obtención de $\hat{x}(k/k+l)$ para todo k y un retraso l fijo). En los apartados siguientes, establecemos algoritmos recursivos para la obtención de los suavizadores en cada una de las situaciones

mencionadas. Hay que indicar que cada suavizador particular, $\hat{x}(k/j)$, es único; lo que varía es la forma de obtenerlo en cada uno de los algoritmos.

4.1 Algoritmo de suavizamiento punto fijo

Consideremos el sistema definido en la Sección 1, y supongamos que el filtro óptimo y la matriz de covarianzas del error de filtrado son conocidos en cualquier instante de tiempo. El algoritmo de suavizamiento punto fijo proporciona el estimador óptimo del estado $x(k)$ para un instante fijo k basado en el conjunto de observaciones $\{z(0), \dots, z(j)\}$ a partir del estimador de $x(k)$ basado en $\{z(0), \dots, z(j-1)\}$, donde $j = k+1, k+2, \dots$. Es, por tanto, un algoritmo recursivo en j hacia adelante.

1. *El suavizador punto fijo óptimo, $\hat{x}(k/j)$, viene dado por la relación*

$$\hat{x}(k/j) = \hat{x}(k/j-1) + K(k,j) [z(j) - H(j)\Phi(j,j-1)\hat{x}(j-1/j-1)]$$

donde $j = k+1, k+2, \dots$ y la condición inicial es el filtro, $\hat{x}(k/k)$.

2. *$K(k,j)$, la matriz de ganancia del suavizador punto fijo, verifica*

$$K(k,j) = L(k,j-1)\Phi^T(j,j-1)H^T(j) [H(j)P(j/j-1)H^T(j) + R(j)]^{-1}$$

donde $L(k,j)$ es obtenida recursivamente por la siguiente relación

$$\begin{aligned} L(k,j) &= L(k,j-1)\Phi^T(j,j-1) [I - K(j)H(j)]^T, \quad j = k+1, k+2, \dots \\ L(k,k) &= P(k/k). \end{aligned}$$

3. *La matriz de covarianzas del error, $P(k/j)$, verifica*

$$P(k/j) = P(k/j-1) - K(k,j) [H(j)P(j/j-1)H^T(j) + R(j)] K^T(k,j)$$

donde $j = k+1, k+2, \dots$ y la condición inicial es la matriz de covarianzas del error de filtrado $P(k/k)$.

4.2 Algoritmo de suavizamiento intervalo fijo

Consideremos el sistema definido en la Sección 1, y supongamos que disponemos de observaciones $\{z(0), \dots, z(j)\}$ en un intervalo fijado $[0, j]$, así como del filtro óptimo y la matriz de covarianzas del error de filtrado en cualquier instante de dicho intervalo. El algoritmo de suavizamiento intervalo fijo proporciona el estimador óptimo del estado $x(k)$ para cada k , $0 \leq k \leq j$, basado en el conjunto $\{z(0), \dots, z(j)\}$ de todas observaciones disponibles en el intervalo $[0, j]$. Cada estimador $\hat{x}(k/j)$ es obtenido a partir de $\hat{x}(k+1/j)$ y, por tanto, es un algoritmo recursivo en k , hacia atrás.

1. El suavizador intervalo fijo óptimo, $\hat{x}(k/j)$, viene dado por la relación

$$\hat{x}(k/j) = \hat{x}(k/k) + A(k) [\hat{x}(k+1/j) - \hat{x}(k+1/k)]$$

donde $k = j-1, j-2, \dots, 0$ y la condición inicial es el filtro $\hat{x}(j/j)$.

2. $A(k)$, la matriz de ganancia del suavizador intervalo fijo está dada por

$$A(k) = P(k/k)\Phi^T(k+1,k)P^{-1}(k+1,k), \quad k = j-1, j-2, \dots, 0.$$

3. La matriz de covarianzas del error, $P(k/j)$, verifica

$$P(k/j) = P(k/k) + A(k) [P(k+1/j) - P(k+1/k)] A^T(k), \quad k = j-1, \dots, 0,$$

con condición inicial $P(j/j)$, la matriz de covarianzas del error de filtrado.

4.3 Algoritmo de suavizamiento retraso fijo

Consideremos el sistema definido en la Sección 1, y supongamos que para cada $k = 0, 1, \dots$ existe un retraso fijo de l unidades de tiempo entre el instante k en el que se desea estimar el estado y la última medida disponible. El algoritmo de suavizamiento retraso fijo proporciona el estimador óptimo del estado $x(k)$ basado en el conjunto $\{z(0), \dots, z(k+l)\}$ (l fijo) de observaciones disponibles en el instante k . Es un algoritmo recursivo *hacia adelante* en k , proporcionando $\hat{x}(k+1/k+1+l)$ a partir de $\hat{x}(k/k+l)$.

1. El suavizador retraso fijo óptimo, $\hat{x}(k+1/k+1+l)$, viene dado por la relación

$$\begin{aligned} \hat{x}(k+1/k+1+l) &= \Phi(k+1,k)\hat{x}(k/k+l) \\ &\quad + C(k+1+l)K(k+1+l)\tilde{z}(k+1+l/k+l) \\ &\quad + U(k+1)[\hat{x}(k/k+l) - \hat{x}(k/k)], \quad k = 0, 1, \dots \end{aligned}$$

con condición inicial $\hat{x}(0/l)$, el suavizador punto fijo.

2. $C(k+1+l)$ y $U(k+1)$, las matrices de ganancia del suavizador retraso fijo, están dadas por

$$C(k+1+l) = \prod_{i=k+1}^{k+l} A(i) \quad \text{con} \quad A(i) = P(i/i)\Phi^T(i+1,i)P^{-1}(i+1,i)$$

$$U(k+1) = \Gamma(k+1,k)Q(k)\Gamma^T(k+1,k)\Phi(k+1,k)P^{-1}(k,k)$$

3. La matriz de covarianza del error, $P(k+1/k+1+l)$, verifica

$$\begin{aligned} P(k+1/k+1+l) &= P(k+1/k) - C(k+1+l)K(k+1+l) \\ &\quad \times H(k+1+l)P(k+1+l/k+l)C^T(k+1+l) \\ &\quad - A^{-1}(k)[P(k/k) - P(k/k+l)][A^T(k)]^{-1}, \quad k = 0, 1, \dots \end{aligned}$$

y la condición inicial, $P(0/l)$, se calcula con el algoritmo de suavizamiento punto fijo.