Tema 4: Estimación en sistemas lineales con observaciones inciertas

1. Sistemas con observaciones inciertas

2. Estimación óptima de mínimos cuadrados

3. Estimación lineal mínimo cuadrática

En el estudio del problema de estimación en sistemas lineales discretos realizado en el tema anterior, las observaciones siempre tenían información del estado que se deseaba estimar y la perturbación de éstas se debía, únicamente, a un ruido aditivo. En este caso, el filtro de Kalman proporciona el estimador lineal de mínimos cuadrados, que coincide con el estimador óptimo para cualquier función de pérdida admisible bajo condiciones de gaussianidad e independencia mutua de la condición inicial y los ruidos del estado y las observaciones.

Sin embargo, en muchas aplicaciones prácticas como, por ejemplo, en telecomunicaciones o en procesamiento de imágenes, el mecanismo de medidas puede estar afectado por interrupciones aleatorias, de modo que el estado interviene en la ecuación de la observación de manera aleatoria; esto es, existe una probabilidad positiva (probabilidad de falsa alarma) de que las observaciones contengan únicamente ruido. Estas situaciones se modelizan mediante los denominados sistemas con observaciones inciertas, que incluyen en la ecuación de observación un ruido escalar multiplicativo descrito por una sucesión de variables aleatorias de Bernoulli. La presencia de ruido multiplicativo en la ecuación de observación hace que no se satisfagan las condiciones necesarias para la aplicación del filtro de Kalman y, así, se plantea la necesidad de resolver el problema de estimación en este tipo de sistemas. En este tema describiremos dichos sistemas y estudiaremos el problema de estimación mínimo cuadrática del estado en cada instante de tiempo a partir de observaciones inciertas hasta dicho instante.

1 Sistemas con observaciones inciertas

Como ya hemos comentado, en diversos problemas prácticos, la observación en un instante puede ser sólo ruido; estas situaciones se modelizan mediante una ecuación de observación que no sólo está perturbada por un ruido aditivo, sino que incluye también una componente (ruido) multiplicativa que hace que el estado intervenga en la observación de forma aleatoria. Por tanto, la diferencia entre los sistemas que aquí trataremos y los descritos en el tema anterior radica en la ecuación de observación, ya que el problema planteado sólo afecta a ésta.

1.1 Ecuación del estado

En virtud de los comentarios anteriores, la ecuación del estado es la misma considerada en el tema anterior y, por tanto, el estado $x\,(k)$ es un vector aleatorio n-dimensional que verifica

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

donde $\Phi(k+1,k)$ y $\Gamma(k+1,k)$ son matrices determinísticas conocidas de dimensiones apropiadas. $\{w(k);\ k \geq 0\}$ es un proceso ruido blanco r-dimensional (*ruido del estado*) con media cero y $E\left[w(k)w^T(k)\right] = Q(k)$. El estado inicial, x_0, es un vector aleatorio n-dimensional con media cero y matriz de covarianzas P_0.

1.2 Ecuación de observación

En la ecuación de observación $z(k) = H(k)x(k) + v(k)$, considerada en el tema anterior, al ser la matriz de observación $H(k)$ determinística, el estado $x(k)$ está siempre presente en la observación $z(k)$. Ahora, para reflejar la posibilidad de que la observación en un instante k pueda contener únicamente ruido con probabilidad $1 - p(k)$, definimos la ecuación de observación de la siguiente forma:

$$z(k) = \begin{cases} H(k)x(k) + v(k), & \text{con probabilidad } p(k) \\ v(k), & \text{con probabilidad } 1 - p(k). \end{cases}$$

Si denotamos mediante $\gamma(k)$ una variable aleatoria de Bernoulli con

$$P\left(\gamma(k) = 1\right) = p(k),$$

la ecuación de observación se puede reescribir como

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

donde $H(k)$ es una matriz determinística conocida de dimensión apropiada, y $\{v(k);\ k \geq 0\}$ es un proceso ruido blanco m-dimensional (*ruido de las observaciones*) con media cero y $E\left[v(k)v^T(k)\right] = R(k)$, siendo $R(k)$ una matriz definida positiva. El ruido multiplicativo $\{\gamma(k);\ k \geq 0\}$, que modeliza la incertidumbre de las observaciones, es una sucesion de variables de Bernoulli con $P\left(\gamma(k) = 1\right) = p(k),\ \forall k \geq 0$; la probabilidad de que la observación $z(k)$ contenga únicamente ruido es $1 - p(k)$ y se denomina *probabilidad de falsa alarma.*

2 Estimación óptima de mínimos cuadrados

Consideremos el sistema lineal en tiempo discreto con observaciones inciertas descrito en la sección anterior:

$$\begin{aligned} x(k+1) &= \Phi(k+1,k)x(k) + \Gamma(k+1,k)w(k), \qquad k \geq 0; \qquad x(0) = x_0, \\ z(k) &= \gamma(k)H(k)x(k) + v(k), \qquad k \geq 0. \end{aligned}$$

Como ya hemos estudiado, el estimador óptimo de mínimos cuadrados de $x(k)$ basado en las observaciones hasta el instante k, el filtro óptimo, es

$$\widehat{x}(k/k) = E\left[x(k)/z(0), \ldots, z(k)\right],$$

y denotando mediante $f(x(k)/Z_k)$ a la función de densidad condicionada del vector $x(k)$ dadas las observaciones $Z_k = (z^T(0), \ldots, z^T(k))^T$, este estimador viene dado por

$$\widehat{x}(k/k) = E\left[x(k)/Z_k\right] = \int x(k) f(x(k)/Z_k) dx(k).$$

Como las observaciones $z(0), \ldots, z(k)$ son función de $\gamma(0), \ldots, \gamma(k)$, tenemos que

$$f(x(k)/Z_k) = \sum_{\Upsilon_k \in S^{k+1}} f(x(k)/Z_k, \Upsilon_k) P\left(\Upsilon_k/Z_k\right),$$

donde $S = \{0, 1\}$ representa el espacio de estados de las variables $\gamma(k)$, S^{k+1} es el producto cartesiano $S \times \overset{(k+1)}{\cdots} \times S$ ($card\ S^{k+1} = 2^{k+1}$) y $\Upsilon_k = (\gamma(0), \ldots, \gamma(k))$ es un elemento de S^{k+1}.

Así, el estimador óptimo viene dado por

$$\widehat{x}(k/k) = \int x(k) \sum_{\Upsilon_k \in S^{k+1}} f(x(k)/Z_k, \Upsilon_k) P\left(\Upsilon_k/Z_k\right) dx(k),$$

e intercambiando la sumatoria y el signo integral, podemos escribir

$$\widehat{x}(k/k) = \sum_{\Upsilon_k \in S^{k+1}} \widehat{x}^{\Upsilon}(k/k) P\left(\Upsilon_k/Z_k\right), \tag{1}$$

donde

$$\widehat{x}^{\Upsilon}(k/k) = E\left[x(k)/Z_k, \Upsilon_k\right] = \int x(k) f(x(k)/Z_k, \Upsilon_k) dx(k)$$

es el estimador de $x(k)$ condicionado a la observaciones Z_k y a un valor concreto $\Upsilon_k \in S^{k+1}$.

En 1971, Jaffer y Gupta establecieron un algoritmo recursivo para la obtención de $\widehat{x}(k/k)$ dado en (1), suponiendo que la sucesión ruido multiplicativo, $\{\gamma(k);\ k \geq 0\}$, es un proceso de Markov con probabilidades de transición y distribución inicial conocidas. Además, se supuso que el estado inicial, x_0, y los ruidos blancos, $\{w(k);\ k \geq 0\}$ y $\{v(k); k \geq 0\}$, eran gaussianos y centrados, y que x_0, $\{w(k);\ k \geq 0\}$, $\{v(k);\ k \geq 0\}$ y $\{\gamma(k);\ k \geq 0\}$ eran mutuamente independientes. En estas condiciones, los estimadores $\widehat{x}^{\Upsilon}(k/k)$, $\Upsilon_k \in S^{k+1}$ tienen una estructura recursiva similar al filtro de Kalman y las probabilidades $P\left(\Upsilon_k/Z_k\right)$, $\Upsilon_k \in S^{k+1}$ pueden calcularse también de forma recursiva.

Sin embargo, aún en las condiciones de Jaffer y Gupta, la expresión (1) indica que el cálculo del filtro óptimo requiere un crecimiento exponencial de memoria pues, en el instante k, se necesita haber almacenado 2^{k+1} estimadores $\widehat{x}^{\Upsilon}(k/k)$ y 2^{k+1} probabilidades $P\left(\Upsilon_k/Z_k\right)$. Esta dificultad computacional, que se deriva de la presencia de las variables $\gamma(k)$ en la ecuación de observación, motiva la búsqueda de estimadores subóptimos en sistemas con observaciones inciertas.

En la siguiente sección establecemos un algoritmo recursivo para el filtro lineal óptimo en sistemas con observaciones inciertas en el caso más simple, cuando las variables de Bernoulli que modelizan la incertidumbre son independientes.

3 Estimación lineal mínimo cuadrática

Consideremos el sistema con observaciones inciertas descrito en la Sección 1,

$$\begin{aligned} x(k+1) &= \Phi(k+1,k)x(k) + \Gamma(k+1,k)w(k), \qquad k \geq 0; \qquad x(0) = x_0, \\ z(k) &= \gamma(k)H(k)x(k) + v(k), \qquad k \geq 0. \end{aligned}$$

Además de las condiciones supuestas en dicha sección sobre los ruidos del sistema y la condición inicial impondremos las siguientes hipótesis:

- El ruido multiplicativo $\{\gamma(k);\ k \geq 0\}$ es una sucesión de variables aleatorias independientes de Bernoulli, con $P\left(\gamma(k) = 1\right) = p(k)$.
- El estado inicial x_0 y los ruidos $\{w(k);\ k \geq 0\}, \{v(k);\ k \geq 0\}, \{\gamma(k);\ k \geq 0\}$ son mutuamente independientes.

En 1969, Nahi dedujo un algoritmo recursivo que resuelve el problema de filtrado lineal de menor error cuadrático medio en este tipo de sistemas. Este algoritmo generaliza al de

Kalman pues, si la probabilidad de falsa alarma, $1-p(k) = P(\gamma(k) = 0)$, es cero, el sistema bajo consideración es un sistema con certidumbre en las observaciones y el algoritmo de Nahi coincide con el de Kalman. Deducimos a continuación dicho algoritmo basándonos en la técnica de proyecciones ortogonales estudiada en el Tema 2.

3.1 Algoritmo recursivo para el problema de filtrado lineal

Según estudiamos en el Tema 2, el filtro de mínimos cuadrados en cualquier instante de tiempo puede obtenerse al partir del predictor mediante la siguiente expresión:

$$\begin{aligned} \widehat{x}(k/k) &= \widehat{x}(k/k-1) + K(k)\widetilde{z}(k/k-1), \qquad k > 0 \\ \widehat{x}(0/0) &= K(0)\widetilde{z}(0/-1), \end{aligned} \tag{2}$$

donde los vectores $\widetilde{z}(k/k-1) = z(k) - \widehat{z}(k/k-1)$, $k > 0$ y $\widetilde{z}(0/-1) = z(0)$ constituyen, como en el Tema 3, las innovaciones en cada instante de tiempo, y $K(k)$ es una matriz a determinar con la condición de minimizar el error cuadrático medio.

Por otra parte, considerando las ecuaciones de Wiener-Hopf correspondientes a $\widehat{x}(k/k-1)$ y $\widehat{x}(k-1/k-1)$,

$$E\left[x(k)z^T(i)\right] = E\left[\widehat{x}\left(k/k-1\right)z^T(i)\right], \quad \forall i = 0, \ldots, k-1,$$

$$E\left[x(k-1)z^T(i)\right] = E\left[\widehat{x}\left(k-1/k-1\right)z^T(i)\right], \quad \forall i = 0, \ldots, k-1,$$

y teniendo en cuenta que $E\left[x(k)z^T(i)\right] = \Phi(k,k-1)E\left[x(k-1)z^T(i)\right]$, es inmediato probar que

$$\begin{aligned} \widehat{x}(k/k-1) &= \Phi(k,k-1)\widehat{x}(k-1/k-1), \qquad k > 0 \\ \widehat{x}(0/-1) &= 0. \end{aligned} \tag{3}$$

Así, una vez determinadas las innovaciones $\widetilde{z}(k/k-1)$ y las matrices de ganancia $K(k)$, las ecuaciones (2) y (3) proporcionan un algoritmo recursivo para el cálculo del filtro.

Proceso innovación

Con objeto de determinar $\widehat{z}(k/k-1)$, consideramos la ecuación de Wiener-Hopf,

$$E\left[z(k)z^T(i)\right] = E\left[\widehat{z}\left(k/k-1\right)z^T(i)\right], \quad \forall i = 0, \ldots, k-1,$$

y teniendo en cuenta que

$$E\left[z(k)z^T(i)\right] = p(k)H(k)E\left[x(k)z^T(i)\right] = p(k)H(k)E\left[\widehat{x}\left(k/k-1\right)z^T(i)\right] \;\; i = 0, \ldots, k-1,$$

se deduce que $\widehat{z}(k/k-1) = p(k)H(k)\widehat{x}\left(k/k-1\right)$. Así, el proceso innovación para el problema que estamos estudiando viene dado por

$$\begin{aligned} \widetilde{z}(k/k-1) &= z(k) - p(k)H(k)\widehat{x}(k/k-1), \qquad k > 0, \\ \widetilde{z}(0/-1) &= z(0). \end{aligned}$$

Este proceso tiene media cero y su matriz de covarianzas, $\Pi(k) = E[\widetilde{z}(k/k-1)\widetilde{z}^T(k/k-1)]$, está dada por

$$\begin{aligned} \Pi(k) &= p(k)(1-p(k))H(k)D(k)H^T(k) + p^2(k)H(k)P\left(k/k-1\right)H^T(k) + R(k) \\ \Pi(0) &= p(0)H(0)P_0H^T(0) + R(0), \end{aligned}$$

donde $D(k) = E[x(k)x^T(k)]$ y $P(k/k-1) = E\left[\widetilde{x}\left(k/k-1\right)\widetilde{x}^T\left(k/k-1\right)\right]$ es la matriz de covarianzas del error de predicción, $\widetilde{x}\left(k/k-1\right) = x(k) - \widehat{x}\left(k/k-1\right)$. Teniendo en cuenta la ecuación del estado y la expresión (3) para $\widehat{x}\left(k/k-1\right)$ estas matrices vienen dadas por

$$\begin{aligned} D(k) &= \Phi(k,k-1)D(k-1)\Phi^T(k,k-1) + \Gamma(k,k-1)Q(k-1)\Gamma^T(k,k-1) \\ D(0) &= P_0, \end{aligned}$$

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

Cálculo de la matriz de ganancia

La matriz de ganancia en la expresión (2) del filtro se calcula minimizando el error cuadrático medio, $E\left[\widetilde{x}^T\left(k/k-1\right)\widetilde{x}\left(k/k-1\right)\right]$ con $\widetilde{x}(k/k) = x(k) - \widehat{x}(k/k)$, o, equivalentemente, minimizando su matriz de covarianzas. Claramente, por la ecuación del estado y la del filtro, el error de estimación satisface

$$\begin{aligned} \widetilde{x}(k/k) &= \widetilde{x}(k/k-1) - K(k)\widetilde{z}(k/k-1) \\ \widetilde{x}(0/0) &= x_0 - K(0)z(0), \end{aligned}$$

y, por tanto, su matriz de covarianzas, $P(k/k) = E\left[\widetilde{x}(k/k)\widetilde{x}^T(k/k)\right]$, viene dada por

$$\begin{aligned} P(k/k) &= P(k/k-1) + K(k)\Pi(k)K^T(k) \\ &\quad -E\left[\widetilde{x}(k/k-1)\widetilde{z}^T(k/k-1)\right]K^T(k) \\ &\quad -K(k)E\left[\widetilde{z}(k/k-1)\widetilde{x}^T(k/k-1)\right]. \end{aligned}$$

Teniendo en cuenta el Lema de Proyecciones Ortogonales y las hipótesis de independencia impuestas al sistema, obtenemos

$$E\left[\widetilde{x}(k/k-1)\widetilde{z}^T(k/k-1)\right] = E\left[\widetilde{x}(k/k-1)z^T(k)\right] = p(k)P(k/k-1)H^T(k),$$

con lo que la matriz de covarianzas del error de filtrado queda

$$\begin{aligned} P(k/k) &= P(k/k-1) + K(k)\Pi(k)K^T(k) \\ &\quad -p(k)P(k/k-1)H^T(k)K^T(k) - p(k)K(k)H(k)P(k/k-1). \end{aligned}$$

Con objeto de encontrar la ganancia $K(k)$ que hace mínima esta expresión, notamos $Z(k) = p(k)P(k/k-1)H^T(k)\Pi^{-1}(k)$, y tenemos la siguiente expresión alternativa:

$$P(k/k) = [K(k) - Z(k)]\,\Pi(k)\,[K(k) - Z(k)]^T - Z(k)\Pi(k)Z^T(k) + P(k/k-1),$$

con la que queda claro que debemos tomar $K(k) = Z(k)$, y la matriz de covarianzas correspondiente es

$$P(k/k) = P(k/k-1) - p(k)K(k)H(k)P(k/k-1).$$

Los resultados obtenidos se resumen en el siguiente algoritmo para los problemas de predicción en una etapa y filtrado.

El filtro lineal de mínimos cuadrados viene dado por la siguiente fórmula recursiva

$$\begin{aligned} &\widehat{x}(k/k) = \widehat{x}(k/k-1) + K(k)\left[z(k) - p(k)H(k)\widehat{x}(k/k-1)\right], \qquad k > 0 \\ &\widehat{x}(0/0) = K(0)z(0) \end{aligned}$$

donde el predictor, $\widehat{x}(k/k-1)$ *está dado por*

$$\begin{aligned} &\widehat{x}(k/k-1) = \Phi(k,k-1)\widehat{x}(k-1/k-1), \qquad k > 0 \\ &\widehat{x}(0/-1) = 0, \end{aligned}$$

y la ganancia, $K(k),\ k \geq 0$*, viene especificada por las siguientes relaciones:*

$$\begin{aligned} &K(k) = p(k)P(k/k-1)H^T(k)\Pi^{-1}(k), \\ &P(k/k-1) = \Phi(k,k-1)P(k/k-1)\Phi^T(k,k-1) + \Gamma(k,k-1)Q(k-1)\Gamma^T(k,k-1) \\ &P(0/-1) = P_0 \\ &P(k/k) = \left[I - p(k)K(k)H(k)\right]P(k/k-1) \\ &\Pi(k) = p(k)(1-p(k))H(k)D(k)H^T(k) + p^2(k)H(k)P(k/k-1)H^T(k) + R(k) \\ &\Pi(0) = p(0)H(0)P_0H^T(0) + R(0) \\ &D(k) = \Phi(k,k-1)D(k-1)\Phi^T(k,k-1) + \Gamma(k,k-1)Q(k-1)\Gamma^T(k,k-1) \\ &D(0) = P_0. \end{aligned}$$