

FILTRADO ESTADÍSTICO

ONÉSIMO HERNÁNDEZ-LERMA

Departamento de Matemáticas, Centro de Investigación y Estudios Avanzados (CINVESTAV)
y Sección de Graduados de la Escuela Superior de Ingeniería Mecánica y Eléctrica (ESIME)
Instituto Politécnico Nacional, México

Resumen: En este trabajo se presenta una breve introducción a la teoría de filtrado estadístico, con énfasis en el filtro de Kalman para sistemas lineales en tiempo discreto.

STATISTICAL FILTERING

Abstract. A brief introduction to the theory of statistical filtering is presented, with emphasis on the Kalman filter for discrete-time linear systems.

INTRODUCCIÓN

En este trabajo se presentan algunos aspectos introductorios sobre la teoría del filtrado estadístico, con especial énfasis en la teoría de los filtros de Kalman. Con el propósito de mantener la presentación a un nivel razonablemente accesible, se considera únicamente el caso de modelos lineales en tiempo discreto.

La palabra filtro tiene distintos significados: puede ser un material o un dispositivo a través del cual se hace pasar un fluido para eliminar las impurezas o partículas sólidas en suspensión; en un receptor de señales, es un dispositivo que elimina frecuencias fuera de un cierto rango, y también puede ser una pantalla que elimina ciertos rayos del espectro. En términos generales podemos pensar en un filtro como un mecanismo que elimina "impurezas" o información no deseable y que, mediante esta acción de filtrado, nos permite obtener el objeto de nuestro interés (agua limpia, señales no distorsionadas, etcétera).

En este trabajo estamos interesados en el filtrado estadístico: deseamos estimar una señal o el estado de un sistema a partir de un conjunto dado de observa-

ciones o datos que pueden tener errores o perturbaciones aleatorias. En otras palabras, estamos interesados en mecanismos (filtros) que eliminen las "impurezas" en los datos y nos permitan obtener información (estimadores) de la señal deseada. Desde este punto de vista la teoría del filtrado y la teoría de estimación estadística son lo mismo. Sin embargo, en los problemas usuales de estimación, dada una muestra, se desea inferir información sobre características de una población *estática*, que no cambia en el tiempo, mientras que en problemas de filtrado la población tiene un comportamiento *dinámico*, está cambiando con el tiempo, de manera que su estado o las características de interés se expresan como una *señal* $x(t)$, es decir, como una función del tiempo t . Cuando t se mide en periodos discretos (e.g. meses o años), a la señal, digamos $x(t)$, $t = 0, 1, 2, \dots$, se le llama una *serie de tiempo*. En algunas áreas de la ingeniería, a la teoría de filtrado también se le conoce con el nombre de procesamiento de señales.

PROBLEMAS DE LA TEORÍA DE FILTRADO

Los problemas de la teoría de filtrado tienen una estructura básica (representada gráficamente en la fi-

gura 1) que consiste de los siguientes componentes: en primer lugar se tiene un sistema generador de señales $x(t)$ o, alternativamente, $x(t)$ se puede pensar como el estado del sistema en el tiempo t . Después tenemos un mecanismo mediante el cual se hacen observaciones (mediciones, muestreos, censos) de la señal o del estado del sistema.

Ahora bien, tanto el sistema mismo como el proceso de observaciones están sujetos a perturbaciones aleatorias —que se describen genéricamente con el nombre de “ruido” aleatorio— y además, el sistema y las observaciones están relacionados entre sí estadísticamente. Para fijar ideas se puede pensar que entre el sistema y las observaciones $z(t)$ existe una relación funcional de la forma

$$z(t) = f(x(t), v(t)),$$

en donde $\{v(t)\}$ es un proceso estocástico que representa el ruido aleatorio. Finalmente, el objetivo es “filtrar” las observaciones $z(t)$ para obtener un “buen estimador” $\hat{x}(t)$ de $x(t)$.

En la práctica, el concepto abstracto de sistema o señal puede tomar formas muy diversas. La descripción abstracta del párrafo anterior puede referirse a señales sísmicas (Findley, 1978), geofísicas (Robin-

son y Treitel, 1980), de satélite, radar o navegación (Anderson y Moore, 1979; Bucy y Joseph, 1968; Chin, 1979; Gelb, 1974), sistemas económicos (Aoki, 1976; Chow, 1975; Mehra, 1979), pesqueros (Balchen, 1979) y otra gran variedad de aplicaciones diversas que van desde la demografía hasta prácticamente cualquier área de la ingeniería.

En todo caso, los problemas de la teoría de filtrado se pueden clasificar en tres tipos. Supongamos que se desea determinar un estimador, $\hat{x}(t)$, del estado del sistema $x(t)$ en el tiempo t , dado que se tienen las observaciones hasta un tiempo T , digamos $z(s)$, $s \leq T$. (Véase la figura 2.) Si $t > T$, se dice entonces que el problema de estimar $x(t)$ es un problema de *predicción* o *extrapolación*; si $t < T$, es un problema de *interpolación* o *suavizamiento*; si $t = T$, es un problema de *filtrado*.

Así pues, en un sentido estricto, el problema de filtrado consiste en estimar $x(T)$ dadas las observaciones hasta el tiempo T . Sin embargo, los problemas anteriores se conocen con el nombre genérico de problemas de “filtrado” porque, como veremos adelante, en muchos casos de interés el problema fundamental es el de filtrado, en el sentido de que los predictores $\hat{x}(t)$, con $t > T$, o interpoladores $\hat{x}(t)$, con $t < T$, se pueden determinar en base al “filtro” $\hat{x}(T)$.

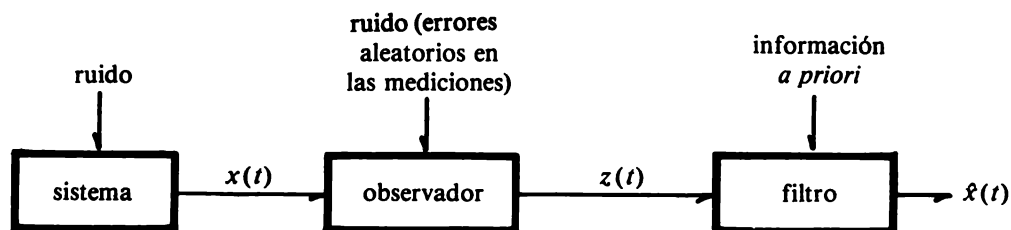


Fig. 1. Componentes principales de un problema de filtrado: sistema, mecanismo de observaciones, filtro (estimador).

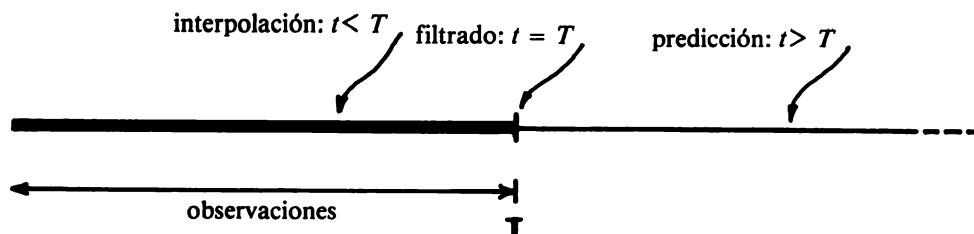


Fig. 2. Problema: estimar $x(t)$ dadas las observaciones hasta el tiempo T .

CRITERIOS DE ESTIMACIÓN

Como parte integral de un problema de filtrado debemos definir un criterio de estimación y seleccionar un filtro que sea óptimo de acuerdo con ese criterio. Lo usual es basar este criterio en algún aspecto de la densidad de probabilidad *a posteriori* $p(x(t)|z(0), z(1), \dots, z(T))$, la cual esencialmente nos da toda la información que podemos obtener sobre $x(t)$ dadas las observaciones $z(0), z(1), \dots, z(T)$. Si denotamos por $Z(T)$ el conjunto de observaciones hasta el tiempo T , entonces podemos escribir

$$\tilde{p}(x(t)|z(0), z(1), \dots, z(T)) = p(x(t)|Z(T)),$$

y el estimador *máximo a posteriori* es aquel que maximiza esta densidad. Por otra parte, si definimos el *error cuadrado medio* (e.c.m.) entre $x(t)$ y cualquier estimador $y = g(Z(T))$ como $E(x(t) - y)^2$, es fácil ver que el estimador en la *medida condicional* $\hat{x}(t|T)$ definido como

$$\hat{x}(t|T) := E(x(t)|Z(T)) = E(x(t)|z(0), \dots, z(T)),$$

es el único estimador que minimiza el e.c.m., es decir,

$$E(x(t) - \hat{x}(t|T))^2 \leq E(x(t) - y)^2$$

para todo estimador $y = g(Z(T))$. El estimador en la media condicional es más conocido usualmente como el *estimador de varianza mínima* debido a que el e.c.m. es

$$E(x(t) - \hat{x}(t|T))^2 = \text{var}(\hat{x}(t|T)),$$

pues el *error de estimación* $\tilde{x}(t) := x(t) - \hat{x}(t|T)$ tiene media cero. Otro hecho importante sobre $\hat{x}(t|T)$ es que es igual al estimador máximo a posteriori cuando la distribución a posteriori es simétrica y unimodal, que es precisamente el caso de los sistemas

lineales con ruido gaussiano que veremos a continuación.

SISTEMAS LINEALES

Tal como se mencionó antes, como parte de la estructura básica de un problema de filtrado debemos especificar los modelos dinámicos del sistema de interés y del proceso de observaciones. En el caso general (no lineal) en tiempo discreto, estos modelos serán ecuaciones en diferencias (o ecuaciones diferenciales en el caso de tiempo continuo de la forma

$$\begin{aligned} x(t+1) &= h(x(t), w(t)) \\ z(t) &= f(x(t), v(t)), \end{aligned}$$

en donde h y f son funciones conocidas y $w(t)$ y $v(t)$ son procesos estocásticos que representan el ruido en el sistema y en las observaciones respectivamente. Aquí deseamos considerar el caso particular en el que tanto h como f son funciones lineales, digamos

$$x(t+1) = A(t)x(t) + w(t) \quad (1)$$

$$\begin{aligned} t &= 0, 1, 2, \dots, \\ z(t) &= H(t)x(t) + v(t) \quad (2) \end{aligned}$$

con $w(t)$ y $v(t)$ *ruidos blancos*. Esto último significa que $w(t)$ y $v(t)$, $t = 0, 1, 2, \dots$, son sucesiones de variables (o vectores) aleatorias independientes, con media cero y matrices de covarianza conocidas:

$$\text{cov}(w(t), w(s)) = Q(t) \delta_{ts}$$

$$\text{cov}(v(t), v(s)) = R(t) \delta_{ts},$$

en donde δ_{ts} es el símbolo de Kronecker (igual a 1 si $t = s$, y 0 si $t \neq s$). En (1) y (2), si $x(t)$ es un vector de dimensión n y $z(t)$ es un vector de dimensión m , entonces $w(t)$ y $v(t)$ son procesos vectoriales de dimensiones n y m respectivamente; $A(t)$ es una matriz n por n , $H(t)$ es m por n , $Q(t)$ es n por n , y $R(t)$ es m por m .

El sistema (1)-(2), representado gráficamente en la figura 3, es más general de lo que aparenta a primera

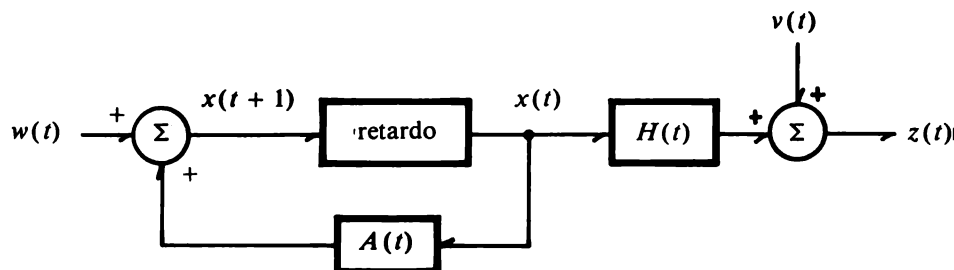


Fig. 3. Representación gráfica del sistema lineal (1), (2).

vista e incluye muchos casos de interés práctico. Por ejemplo, en (1) se puede añadir una variable de entrada o variable exógena $u(t)$, escribiendo (1) en la forma

$$x(t+1) = A(t)x(t) + B(t)u(t) + w(t) \quad (3)$$

y dejando las observaciones como aparecen en (2). Si tomamos $v(t) \equiv 0$ en (2) y suponemos que el sistema es observable directamente, es decir $z(t) = x(t)$ (la matriz $H(t) = \text{identidad}$), entonces (3)-(2) se reducen a

$$z(t+1) = A(t)z(t) + B(t)u(t) + w(t)$$

Esta ecuación representa un modelo muy común en diversas aplicaciones y se le conoce como un *proceso autorregresivo de primer orden* con entrada o variable exógena $u(t)$ y errores aleatorios $w(t)$. Se obtiene una interpretación diferente si (suponiendo $H(t) = \text{identidad}$ y $w(t) = 0$) sustituimos $x(t) = z(t) - v(t)$ en (3), lo cual da como resultado

$$z(t+1) = A(t)z(t) + B(t)u(t) + [v(t+1) - A(t)v(t)].$$

Este es un modelo (vectorial) conocido como un *promedio móvil autorregresivo* de primer orden. El mismo tipo de modelo se obtiene cuando $w(t) \neq 0$. La ecuación (2) en sí misma (ignorando (1) y (3)) representa un modelo muy común consistente en una *señal más ruido*. En las referencias se pueden ver otros casos especiales del modelo lineal (1)-(2) o (3)-(2).

FILTROS DE KALMAN

Teniendo ya el modelo dinámico (1)-(2) del sistema y las observaciones, procedemos al cálculo del filtro (o estimador) óptimo. Empezaremos por considerar el *predicador en un paso*:

$$\hat{x}(t|t-1) = E(x(t)|Z(t-1)) = E(x(t)|z(0), \dots, z(t-1)).$$

El filtro de Kalman (Kalman, 1960) es un esquema iterativo que nos permite calcular $\hat{x}(t|t-1)$ en forma rápida y eficiente, y fácil de implementar en una computadora digital. El desarrollo de los filtros de Kalman tiene una larga historia que se remonta a Gauss, en el siglo XVIII, y que incluye a matemáticos y estadísticos tan venerados como Fisher, Wiener y

Kolmogorov, entre otros. Un interesante artículo sobre los orígenes, evolución y desarrollos recientes de los filtros de Kalman es el de Sorenson (1970).

Para describir el filtro de Kalman debemos imponer algunas hipótesis adicionales al modelo lineal (1)-(2). Supondremos que $w(t)$ y $v(t)$, además de ser ruidos blancos, son procesos *gaussianos*, no correlacionados, o sea,

$$\text{cov}(w(t), v(s)) = 0 \text{ para todo } t \text{ y } s,$$

y supondremos también que el estado inicial, $x(0)$, del sistema es un vector gaussiano con media y covarianza conocidas:

$$Ex(0) = \hat{x}(0), \text{cov}(x(0), x(0)) = P(0),$$

y que $x(0)$ es independiente de los ruidos $w(t)$ y $v(t)$.

Para simplificar la notación escribamos el predictor como

$$\hat{x}(t) = \hat{x}(t|t-1) = E(x(t)|Z(t-1)).$$

Entonces, si la matriz de covariancia $R(t) = \text{cov}(v(t), v(t))$ es positiva-definida, el predictor $\hat{x}(t)$, $t = 1, 2, \dots$, satisface la ecuación

$$\begin{aligned} \hat{x}(t+1) &= A(t)\hat{x}(t) + K(t)(z(t) - H(t)\hat{x}(t)), \\ \hat{x}(0) &= Ex(0), \end{aligned} \quad (4)$$

en donde el coeficiente $K(t)$, llamado la *matriz de ganancia* del filtro, está dada por

$$K(t) = A(t)P(t)H(t)^* [H(t)P(t)H(t)^* + R(t)]^{-1}; \quad (5)$$

(el asterisco * denota matriz "transpuesta").

$P(t) = \text{cov}(\tilde{x}(t), \tilde{x}(t))$ es la matriz de covariancia del error de estimación $\tilde{x}(t) = x(t) - \hat{x}(t)$, y satisface la ecuación (matricial) de Riccati:

$$\begin{aligned} P(t+1) &= A(t)P(t)A(t)^* + Q(t) - \\ &\quad - A(t)P(t)H(t)^* [H(t)P(t)H(t)^* + R(t)]^{-1} H(t)P(t)A(t)^*, \end{aligned}$$

con condición inicial

$$P(0) = \text{cov}(\tilde{x}(0), \tilde{x}(0)) = \text{cov}(x(0), x(0)).$$

El algoritmo del filtro de Kalman consta de las ecuaciones (4), (5) y (6). El algoritmo se inicia con $\hat{x}(0) = E x(0)$ y $P(0) = \text{cov}(x(0), x(0))$, y usando las ecuaciones (4)-(6) se puede obtener $\hat{x}(t)$ para cualquier tiempo t : sustituyendo $P(0)$ en (5), se calcula $K(0)$, y sustituyendo después $K(0)$ y $\hat{x}(0)$ en (4), se obtiene $\hat{x}(1)$. Después, sustituyendo $P(0)$ en (6) se obtiene $P(1)$ y sustituyendo este valor en (5) se obtiene $F(1)$. Finalmente, sustituyendo $F(1)$ y el valor ya calculado $\hat{x}(1)$ en (4), se obtiene $\hat{x}(2)$. Continuando en forma recursiva de esta manera se puede obtener $\hat{x}(t)$ para cualquier $t > 0$. Nótese que las matrices $P(t)$ y $K(t)$ se calculan "fuera de línea". En la figura 4 se representa el filtro, escribiendo la ecuación (4) en la forma equivalente

$$\hat{x}(t+1) = [A(t) - K(t)H(t)] \hat{x}(t) + K(t)z(t) \quad (4')$$

Algunas propiedades del filtro de Kalman

a) Nótese la forma tan especial que tiene la ecuación (4) para el estimador $\hat{x}(t+1) = E(x(t+1)|Z(t))$, en donde

$$Z(t) = \{z(0), z(1), \dots, z(t)\}$$

son las observaciones hasta el tiempo t . De (4), vemos que $\hat{x}(t+1)$ consta de una transformación lineal, $A(t) \hat{x}(t)$, del estimador anterior $\hat{x}(t) = E(x(t)|Z(t-1))$, más un término actualizante

$$K(t)(z(t) - H(t)\hat{x}(t)) \equiv K(t)\tilde{z}(t)$$

que consiste de la nueva información en $z(t)$ no contenida en las observaciones hasta el tiempo $t-1$. Debido a esta propiedad el proceso definido por

$$\begin{aligned} \tilde{z}(t) &:= z(t) - H(t)\hat{x}(t) \\ &= z(t) - H(t)E(x(t)|Z(t-1)) \end{aligned}$$

se llama el *proceso de innovaciones*. De hecho, este proceso tiene la propiedad de ser un ruido *blanco gaussiano* (es decir, es un proceso gaussiano, con media cero y covariancia $E(\tilde{z}(t)\tilde{z}(s)) = 0$ para todo $t \neq s$), y lo más importante de todo es que contiene tanta información como el proceso de observaciones original, en el sentido de que

$$\begin{aligned} \hat{x}(t) &= E(x(t)|z(0), \dots, z(t-1)) \\ &= E(x(t)|\tilde{z}(0), \dots, \tilde{z}(t-1)). \end{aligned}$$

b) Tanto la demostración de las ecuaciones (4)-(6) del filtro como las propiedades del proceso de innovaciones mencionadas en el párrafo anterior se basan fuertemente en el hecho de que bajo las hipótesis impuestas al modelo (1)-(2), conocido como modelo de Gauss-Markov, los procesos $\{x(t)\}$ y $\{z(t)\}$ son conjuntamente gaussianos y también $\{\hat{x}(t)\}$ es gaussiano. Estas afirmaciones son fáciles de probar. Por ejemplo, iterando la ecuación (1) se tiene que

$$\begin{aligned} x(t+1) &= a(0)x(0) + a(1)W(0) + \dots \\ &\quad + a(t)W(t-1) + W(t), \end{aligned}$$

en donde

$$a(i) := A(t)A(t-1)\dots A(i),$$

$$i = 0, 1, \dots, t,$$

de manera que $x(t+1)$ es una combinación lineal de vectores gaussianos independientes y, por lo tanto, $x(t+1)$ es un vector gaussiano. De aquí se sigue fácilmente que $\{x(t)\}$ es un proceso gaussiano. En forma análoga se obtiene la gaussianidad, valga la expresión, de $\{z(t)\}$ y $\{x(t)\}$. Las demostraciones detalladas se pueden ver en cualquiera de los textos sobre filtrado citados en la bibliografía.

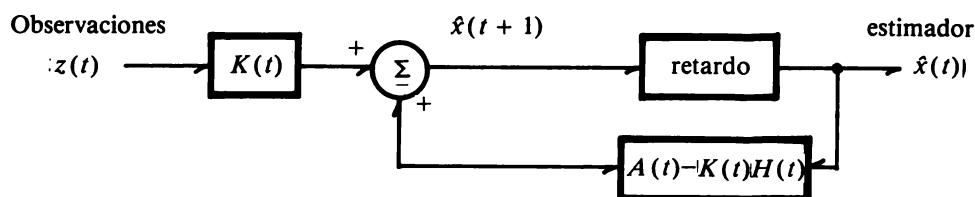


Fig. 4. Representación gráfica de la ecuación del filtro (4) = (4').

c) El filtro de Kalman (4)-(6) nos da el predictor en un paso, digamos de t a $t + 1$:

$$\hat{x}(t + 1) \equiv \hat{x}(t + 1|t) = E(x(t + 1)|Z(t)).$$

El predictor en k pasos ($k = 1, 2, \dots$) es el estimador de $x(t + k)$ dadas las observaciones hasta el tiempo t , es decir,

$$\hat{x}(t + k|t) = E(x(t + k)|Z(t)),$$

y se puede calcular en términos de $\hat{x}(t + 1) = \hat{x}(t + 1|t)$. De hecho, de (1) y del hecho de que los ruidos son blancos se puede obtener que

$$\hat{x}(t + k|t) = A(t + k - 1)A(t + k - 2) \dots$$

$$A(t + 1)\hat{x}(t + 1|t);$$

en particular, si $A(t) = A$ constante,

$$\hat{x}(t + k|t) = A^{k-1}\hat{x}(t + 1|t).$$

d) También los “filtros” e interpoladores se pueden calcular en base al predictor en un paso. Por ejemplo, si

$$\hat{x}(t|t) = E(x(t)|Z(t))$$

es el “filtro” (es decir, el estimador de $x(t)$ dadas las observaciones hasta el tiempo t) se tiene que

$$\hat{x}(t|t) = \hat{x}(t|t-1) + K_f(t)[z(t) - H(t)\hat{x}(t|t-1)],$$

en donde la ganancia

$$K_f(t) = P(t)H(t)^*[H(t)P(t)H(t)^* + R(t)]^{-1};$$

$R(t)$ es la covariancia del ruido $v(t)$ y $P(t)$ es la matriz de covariancia del error $\tilde{x}(t) = x(t) - \hat{x}(t|t-1)$ del predictor dada por (6). Por otra parte, la matriz de covariancia del error de estimación $\tilde{x}_f(t) = x(t) - \hat{x}(t|t)$ en el filtro se puede calcular como

$$P_f(t) = P(t) - K_f(t)H(t)P(t).$$

Para no abundar en tecnicismos y complicar más la presentación, pediremos al lector interesado que consulte la bibliografía sobre más propiedades del filtro de Kalman.

Antes de concluir deseamos hacer algunas observaciones sobre varias de las alternativas posibles que se presentan en el problema de filtrado cuando no se satisfacen las hipótesis impuestas sobre el modelo (1)-(2).

FILTRADO ADAPTIVO

Los filtros de Kalman requieren que la dinámica de la señal (o sistema) y de las observaciones sea lineal y que las perturbaciones aleatorias sean ruidos blancos gaussianos con parámetros conocidos. En la práctica, estos requisitos rara vez se satisfacen. Los problemas con que el estadístico se encuentra frecuentemente son tan oscuros y complejos que es casi imposible dar una descripción precisa del modelo de la señal y/o de los parámetros del ruido. Ante esta situación, una posible alternativa es usar técnicas de series de tiempo (Findley, 1978) para estimar covariancias y a partir de ellas tratar de obtener un modelo de la señal. Otra alternativa es usar algoritmos de filtrado adaptivo, los cuales permiten introducir mediante procedimientos “en línea”, cambios en los parámetros del sistema generador de la señal. En otras palabras, un filtro *adaptivo* (figura 5) puede visualizarse como filtro de Kalman que contiene además un

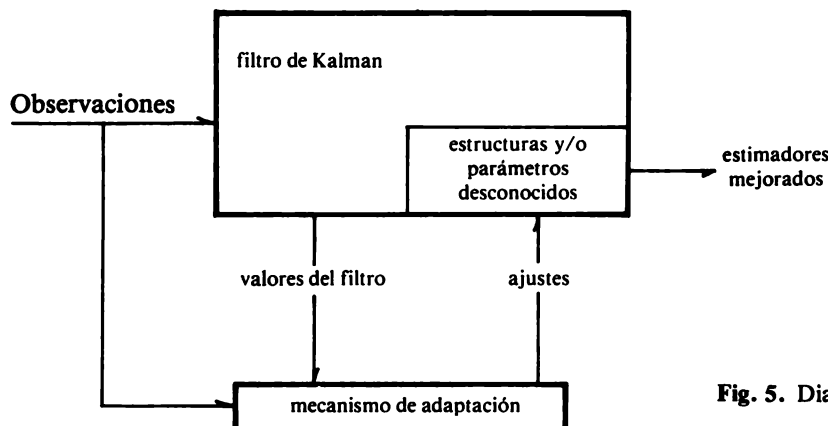


Fig. 5. Diagrama de un filtro adaptivo.

mecanismo de adaptación o "aprendizaje" de tal manera que el sistema integrado puede alcanzar su funcionamiento óptimo aun cuando el medio en el que opera el sistema no es completamente conocido. El mecanismo de adaptación tiene el propósito de tomar las decisiones necesarias para reajustar, adaptar o modificar el comportamiento del filtro conforme se tienen más observaciones. Por supuesto, el mecanismo de adaptación apropiado dependerá de la situación específica que se presente; por ejemplo, puede haber dudas sobre la estructura del modelo del sistema (su dimensión; si es lineal o no lineal, etc.) y/o sobre parámetros del ruido o del modelo mismo (digamos, parámetros generados en base a un número insuficiente de datos o datos de calidad dudosa). Algunas referencias en donde se puede consultar sobre el tema son Anderson y Moore (1979), Chin (1979).

CONCLUSIONES

En los párrafos anteriores se ha intentado dar un panorama general sobre la teoría de filtrado estadístico, haciendo énfasis en la teoría de los filtros de Kalman. Según algunos autores (e.g. Balakrishnan, 1978), la teoría de Kalman es el avance específico de la teoría de sistemas que ha tenido mayor impacto en las aplicaciones. Citar algunas de estas aplicaciones es prácticamente imposible en un trabajo como el presente porque requeriría introducirnos en una serie de detalles técnicos y especializados del área en cuestión, lo cual no es nuestro propósito. Sin embargo, una breve ojeada a la bibliografía da de inmediato una idea de la gran variedad de aplicaciones que tienen los filtros de Kalman. Por último, también se hizo una breve mención del filtrado adaptivo, particularmente útil cuando existen dudas sobre la validez de algunas de las hipótesis en las que se basa la construcción de los filtros de Kalman.

Con el fin de mantener la presentación a un nivel razonablemente elemental, no hicimos mención alguna de filtrado para modelos no lineales y/o en tiempo continuo (ecuaciones diferenciales, en lugar de ecuaciones en diferencias como (1) y (2)). Estos temas

se pueden ver, por ejemplo en Bucy y Joseph (1968), Jazwinsky (1970), o Hazewinkel y Willems (1981). El problema conjunto de filtrado y control adaptable de ciertos sistemas markovianos se puede ver en Hernández-Lerma y Marcus (1986).

Agradecimientos. Este trabajo se realizó con apoyo parcial del Consejo Nacional de Ciencia y Tecnología (CoNaCyT) bajo el proyecto PCCBBNA-020630.

BIBLIOGRAFIA

- Anderson, B.D.O. y J.B. Moore, 1979. *Optimal Filtering*, Prentice-Hall.
- Aoki, M., 1976. *Optimal Control and System Theory in Dynamic Economic Analysis*, North Holland.
- Balakrishnan, A.V., 1978. "The Kalman filter". *The Mathematical Intelligencer*, vol. 1, pp. 90-92.
- Balchen, J.G., 1979. "Modeling, prediction, and control of fish behavior". *Control and Dynamic Systems*, Vol. 15, Academic Press, pp. 99-146.
- Bickel, P.J. and K.A. Doksum, 1977. *Mathematical Statistics*, Holden-Day.
- Bucy, R.S. and P.D. Joseph, 1968. *Filtering for Stochastic Processes with Applications to Guidance*, Wiley.
- Chin, L., 1979. "Advances in adaptive filtering". *Control and Dynamic Systems*, vol. 15, Academic Press, pp. 277-356.
- Chow, G., 1975. *Analysis and Control of Dynamic Economic Systems*, Wiley.
- Findley, D.F., editor, 1978. *Applied Time Series Analysis*, Academic Press.
- Gelb, A., 1974. *Applied Optimal Estimation*, The MIT Press.
- Hazewinkel, M. and J.C. Willems, editores, 1981. *Stochastic Systems: The Mathematics of Filtering and Identification and Applications*, D. Reidel.
- Jazwinsky, A.H., 1970. *Stochastic Processes and Filtering Theory*, Academic Press.
- Kalman, R.E., 1960. "A new approach to linear filtering and prediction problems", *Journal of Basic Engineering*, vol. 1, pp. 597-622.
- Mehra, R.K. 1979. "Kalman filters and their applications to forecasting", *TIMS Studies in the Management Sciences*, vol. 12, pp. 75-94.
- Melsa, J.L. y D.L. Cohn, 1978. *Decision and Estimation Theory*, McGraw-Hill.
- Robinson, E.A. y S. Treitel, 1980. *Geophysical Signal Analysis*, Prentice-Hall.
- Sorenson, H.W., 1970. "Least-squares estimation: from Gauss to Kalman", *IEEE Spectrum*, julio 1970, pp. 63-68.