

ALGUNAS POSIBILIDADES Y LIMITACIONES DE LAS TECNICAS DE  
FILTRAJE NO LINEAL

P. BRUNET J. PAGES

El filtraje no lineal es una técnica con grandes posibilidades; previsiblemente podrá utilizarse para problemas tales como el de previsión, estudio de series cronológicas y estimación de parámetros.

Cabe distinguir entre las técnicas de filtraje óptimo y las de filtraje subóptimo. Las primeras, basadas en el seguimiento de la densidad de probabilidad condicionada por las observaciones, conducen a algoritmos que quedan limitados por el tiempo que su implementación exige. Una línea de trabajo consiste pues en estudiar métodos que permitan disminuir el tiempo de cálculo. Las técnicas de filtraje subóptimo, más ampliamente utilizadas, conducen a algoritmos de implementación más simple. Sin embargo, la convergencia no queda asegurada como sucede en el caso óptimo. El estudio de la misma constituye otra importante línea de trabajo.

En este trabajo se presentan resultados obtenidos en las direcciones citadas. De ellos se deduce que la extensión del filtraje a problemas nuevos debe realizarse con cierta prudencia.

### 1. EL PROBLEMA DEL FILTRAJE

El problema del filtraje consiste en obtener una estimación de un proceso estocástico a partir del conocimiento de otro proceso denominado de observación relacionado con el primero. Se suponen conocidos modelos matemáticos de evolución del proceso a estimar así como de la relación entre éste y el de observación.

#### 1.1 Modelos discretos

La evolución del proceso a estimar viene dada por una ecuación del tipo

$$x_{k+1} = f(x_k, u_k)$$

en que  $u_k$  es una variable aleatoria y la relación entre  $x_k$  y  $x_{k+1}$  es tal que define la probabilidad de transición

$$P(x_{k+1}/x_k)$$

La ecuación de observación viene dada a su vez por

$$z_k = h(x_k, v_k)$$

con  $v_k$  variable aleatoria y la función  $h$  tal que defina la probabilidad

$$P(z_k/x_k)$$

El problema del filtraje en este caso consiste en obtener una estimación  $\hat{x}_k$  de  $x_k$  conociendo únicamente el proceso de observación  $\{z_l, 1 \leq l \leq k\}$ . Otro problema directamente relacionado con éste es el denominado de predicción, donde conociendo  $\{z_l, 1 \leq l \leq k\}$  se intenta estimar  $x_{k+h}$ ,  $h > 0$ .

#### 1.2 Modelos continuos

En este caso, la evolución del proceso a estimar viene dada por una ecuación diferencial con perturbaciones. Estas perturbaciones reciben el nombre genérico de ruidos, y en el caso continuo únicamente existe una teoría matemática del filtraje cuando éstos son gaussianos y actúan de forma aditiva:

$$dx = f(x, t) dt + G(t) d\beta$$

- P. Brunet del Laboratori de Càlcul de la ETSEIB. Diagonal, 647. Barcelona 28.
- J. Pagés de la Càtedra de Matemàtiques II de la ETSEIB. Diagonal, 647. Barcelona 28.
- Article rebut l'Agost de 1978.

$\beta_t$  es un movimiento Browniano, proceso que cumple:

$$E[d\beta] = 0 \quad E[d\beta d\beta^T] = Qdt$$

A su vez, la ecuación de observación es del tipo

$$dz = h(x, t) dt + d\eta$$

con  $\eta$  otro movimiento Browniano tal que

$$E[d\eta] = 0 \quad E[d\eta d\eta^T] = Rdt$$

Con el proceso  $x_t$  continuo, puede considerarse el caso en que las observaciones sean discretas,

$$z_k = h(x(t_k), t_k) + u_k$$

La generación de la estimación  $\hat{x}_t$  del proceso  $x_t$  es en este caso un problema de predicción entre cada par de sucesivas observaciones.

### 1.3 El filtraje no lineal óptimo discreto

Si se adopta como criterio la minimización de

$$E[(\lambda^T(x-v))^2]$$

con  $\lambda$  un vector arbitrario y  $v$  un funcional de las observaciones, se obtiene como solución óptima la esperanza matemática condicionada, es decir

$$v_{opt} = \hat{x} = E[x/\text{observaciones}]$$

En el caso discreto puede calcularse  $\hat{x}_k$  de forma recurrente a partir de la densidad de probabilidad condicionada. Denominado

$$z_k = \{z_k, h_{\leq k}\}$$

la recurrencia se establece entre  $P(x_k/z_k)$  y  $P(x_{k+1}/z_{k+1})$ , obteniéndose como paso intermedio  $P(x_{k+1}/z_k)$ . En efecto, a partir  $P(x_k/z_k)$  y teniendo en cuenta la ecuación de Chapman-Kolmogorov para el proceso  $x_k$

$$P(x_{k+1}/z_k) = \int P(x_{k+1}/x_k z_k) P(x_k/z_k) dx_k =$$

$$= \int P(x_{k+1}/x_k) P(x_k/z_k) dx_k$$

$P(x_k/z_k)$  es la base de la recurrencia y  $P(x_{k+1}/x_k)$  puede obtenerse a partir de la ecuación de evolución del proceso  $x_k$ .

En el siguiente paso, mediante el teorema de Bayes se integra la información suministrada por la observación siguiente  $z_{k+1}$

$$P(x_{k+1}/z_{k+1}) = P(x_{k+1}/z_k) z_k =$$

$$= \frac{P(z_{k+1}/x_{k+1} z_k) P(x_{k+1}/z_k)}{\int P(z_{k+1}/x_{k+1} z_k) P(x_{k+1}/z_k) dx_{k+1}} =$$

$$= \frac{P(z_{k+1}/x_{k+1}) P(x_{k+1}/z_k)}{\int P(z_{k+1}/x_{k+1}) P(x_{k+1}/z_k) dx_{k+1}}$$

La implementación de este algoritmo presenta el problema del tiempo exigido para el cálculo de la integral que aparece en el primer paso. En efecto si las densidades de probabilidad  $P(x_k/z_k)$  y  $P(x_{k+1}/z_k)$ , se representan mediante sus valores en  $N$  puntos, el tiempo de cálculo es proporcional a  $N^2$ .

Por ello es interesante investigar métodos que permitan reducir sustancialmente dicho tiempo de cálculo. Con esta óptica se ha estudiado la posibilidad de calcular dicha integral en paralelo, y la de utilizar la transformación de Fourier,  $1/$ .

## 2. POSIBLES IMPLEMENTACIONES DEL FILTRO OPTIMO DISCRETO

### 2.1 Cálculo en paralelo de $P(x_{k+1}/z_k)$

En este caso el procedimiento utilizado queda esquematizado en la figura 1.

Se ha utilizado el calculador híbrido para obtener por este procedimiento el filtro óptimo del problema conocido como "problema del sensor cúbico". Las ecuaciones correspondientes son

$$x_{k+1} = ax_k + u_k \quad E(u_k^2) = q$$

$$z_k = x_k^3 + v_k \quad E(v_k^2) = r$$

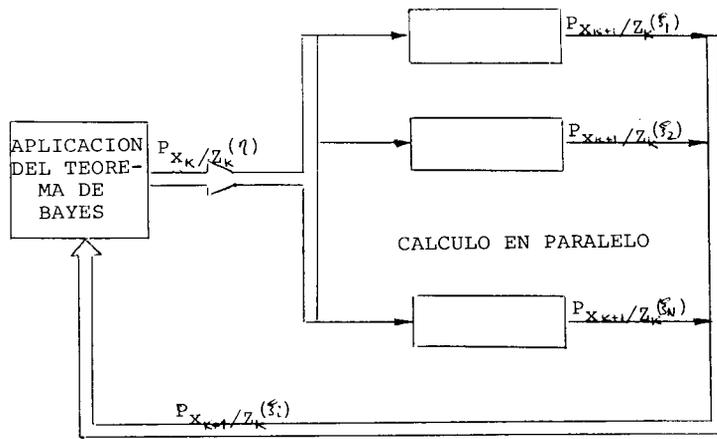


Fig. 1  
Filtraje discreto  
Cálculo en paralelo de la predicción

El carácter no lineal del problema es debido a la ecuación de la observación. La ecuación de evolución del proceso  $x$  es en cambio lineal debido a lo cual la ecuación de Chapman-Kolmogorov es de la forma

$$P_{x_{k+1}/z_k}(\xi) = \int_{-\infty}^{\infty} P_{x_{k+1}/x_k}(\xi/\eta) P_{x_k/z_k}(\eta) d\eta = \int_{-\infty}^{\infty} P(u_k)(\xi - a\eta) P_{x_k/z_k}(\eta) d\eta =$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi q}} e^{-\frac{1}{2q}(\xi - a\eta)^2} P_{x_k/z_k}(\eta) d\eta$$

ecuación parecida a una convolución.

En el computador híbrido, la parte analógica realiza el cálculo en paralelo de

$$P_{x_{k+1}/z_k}(\xi_i)$$

para distintos puntos  $\xi_i$ ; a base de integrar el producto de

$$P_{x_k/z_k}(\eta)$$

por

$$P(u_k)(\xi_i - a\eta)$$

La función

$$P_{x_k/z_k}(\eta)$$

la recibe del digital mientras que las funciones

$$P(u_k)(\xi_i - a\eta)$$

que resultan para los distintos valores  $\xi_i$ , se generan mediante circuitos analógicos; - estas funciones son en este caso gaussianas que difieren únicamente en la media.

El computador digital realiza además del control del analógico el producto y la normalización que aparece en la fórmula de Bayes.

## 2.2 Utilización de la Transformación de Fourier

En el problema del sensor cúbico y en general siempre que la ecuación del proceso observado sea lineal, la ecuación de Chapman-Kolmogorov resulta del tipo integral de convolución. Por ello puede pensarse en la utilización de la transformación de Fourier que reduce el cálculo de dicha integral al de una transformada, un producto y una transformación inversa.

El cálculo de la transformación de Fourier directa e inversa se realiza con ventajas - mediante el algoritmo de Transformación rápida.

El tiempo de cálculo resulta ser proporcional a  $N \log N$ . Puede compararse tiempos de cálculo obtenidos en las gráficas de las figuras 2 y 3.

## 3. FILTRAJE CONTINUO

Así como en el caso discreto la dificultad

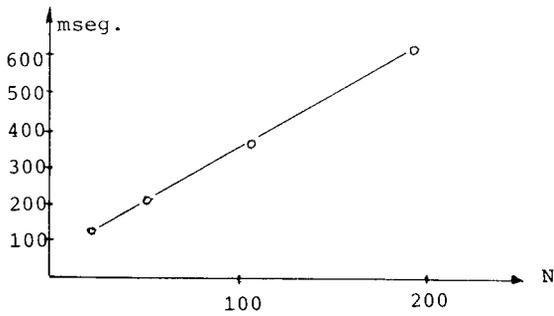


Fig. 2  
Cálculo híbrido en paralelo.  
Tiempo por iteración

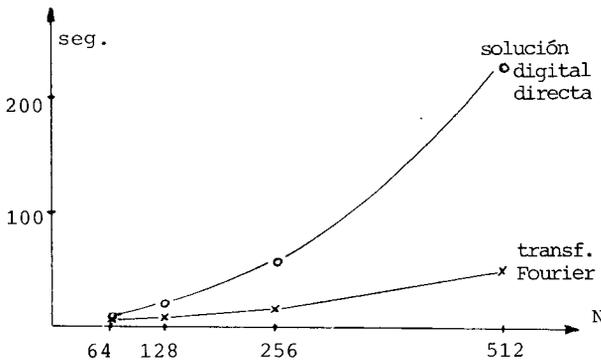


Fig. 3  
Transformación de Fourier comparada con la  
solución digital directa

para la implementación del filtro óptimo reside en el tiempo de cálculo, no ocurre lo mismo en el caso continuo. Aquí, la obtención de la evolución temporal de la densidad de probabilidad condicionada conduce a un sistema de infinitas ecuaciones diferenciales para sus momentos, en que interviene el proceso de observación.

Únicamente en el caso lineal

$$dx = F x dt + G d\beta \quad E[d\beta d\beta^T] = Q dt$$

$$dz = H x dt + d\eta \quad E[d\eta d\eta^T] = R dt$$

en que todos los procesos que aparecen son gaussianos, el sistema anterior se reduce a las ecuaciones para la evolución de los dos primeros momentos de la densidad de probabilidad condicionada. Llamando  $P$  a la matriz de covariaciones de la estimación

$$P = E[(x-\hat{x})(x-\hat{x})^T] / z_T, \quad \tau \leq t$$

se tiene

$$d\hat{x} = F \hat{x} dt + P H^T R^{-1} (dz - H \hat{x} dt)$$

$$\frac{dP}{dt} = FP + PF^T + GQG^T - P H^T R^{-1} H P$$

ecuaciones que corresponden al denominador - filtro de Kalman.

Las dificultades señaladas han conducido a la utilización en el caso lineal de filtros subóptimos, en los que, partiendo de diversas aproximaciones, se llega a una ecuación o sistema de ecuaciones diferenciales para la estimación  $\hat{x}$ , que ya es resoluble. Entre los más usados podemos citar /3/,

- El Extended Kalman, que se obtiene desarrollando las funciones no lineales  $f(x,t)$  y  $h(x,t)$  en serie de Taylor alrededor de la estimación  $\hat{x}$ :

$$f(x,t) = f(\hat{x},t) + \left. \frac{\partial f}{\partial x} \right|_{\hat{x}} (x-\hat{x}) + \dots$$

$$h(x,t) = h(\hat{x},t) + \left. \frac{\partial h}{\partial x} \right|_{\hat{x}} (x-\hat{x}) + \dots$$

y aplicando el filtro de Kalman al sistema lineal resultante, cuando se desprecian los términos de orden superior al primero. Las ecuaciones que se obtienen en este caso son:

$$d\hat{x} = f(\hat{x},t) dt + P \left[ \frac{\partial h}{\partial x} \right]_{\hat{x}}^T R^{-1} (dz - h(\hat{x},t) dt)$$

$$\frac{dP}{dt} = \left[ \frac{\partial f}{\partial x} \right]_{\hat{x}} P + P \left[ \frac{\partial f}{\partial x} \right]_{\hat{x}}^T + \left[ GQG^T \right]_{\hat{x}} - P \left[ \frac{\partial h}{\partial x} \right]_{\hat{x}}^T R^{-1} \left[ \frac{\partial h}{\partial x} \right]_{\hat{x}} P$$

- El filtro de segundo orden, resulta desarrollando en serie de Taylor hasta el segundo orden las funciones  $f(x,t)$  y  $h(x,t)$ , y anulando los momentos de orden superior al segundo que aparecerían en las ecuaciones diferenciales para los dos primeros.
- Filtro gaussiano, se obtiene utilizando las mismas aproximaciones que en caso anterior, salvo los momentos de cuarto orden; en este caso se suponen función de los de segundo como lo serían en un caso gaussiano.

Aunque estos filtros conducen a algoritmos en los que el tiempo no es un factor condicionante, no está asegurada su convergencia. Resulta por ello importante el análisis del comportamiento de dichos filtros para cada

posible aplicación.

4. ESTABILIDAD DE LOS FILTROS SUBOPTIMOS EN LA ESTIMACION DE PARAMETROS

En un sistema cuya evolución depende de varios parámetros que supondremos constituyen el vector a,

$$dx = f(x,a,t) dt + G(x,a,t) d\beta$$

es posible estimar sus valores mediante las técnicas de filtraje y a partir del proceso de observación z,

$$dz = h(x,t) dt + d\eta$$

Para ello, basta considerar el vector

$$y = (x^T, a^T)^T$$

y definir el nuevo sistema:

$$dy = f^*(y,t) dt + G^*(y,t) d\beta$$

donde  $f^*(y,t)=f(x,a,t)$  y  $G^*(y,t)=G(x,a,t)$  - para las primeras componentes de y, mientras que  $f^*(y,t)=G^*(y,t)=0$  para las restantes. - Usualmente, esta transformación conduce a - un sistema no lineal.

Sea por ejemplo el sistema lineal y escalar

$$dx = a x dt + d\beta \quad E[d\beta^2] = qdt$$

Definiendo  $y_1=x$ ,  $y_2=a$  el sistema puede escribirse en la forma

$$dy_1 = y_1 y_2 dt + d\beta$$

$$dy_2 = 0$$

en la que el carácter no lineal aparece en el producto  $y_1 y_2$ . Resulta pues un caso en - que la utilización de algoritmos subóptimos parece a priori adecuada.

La deducción de sus ecuaciones para este caso concreto suponiendo el proceso de observación definido por

$$dz = y_1 dt + d\eta \quad E[d\eta^2] = rdt$$

conduce a

$$d\hat{y}_1 = (c P_{12} + \hat{y}_1 \hat{y}_2) dt + \frac{1}{r} P_{11} (dz - \hat{y}_1 dt)$$

$$d\hat{y}_2 = \frac{P_{12}}{r} (dz - \hat{y}_1 dt)$$

$$\frac{dP_{11}}{dt} = 2\hat{y}_1 P_{12} + 2\hat{y}_2 P_{11} + q - \frac{1}{r} P_{11}^2$$

$$\frac{dP_{12}}{dt} = \hat{y}_2 P_{12} + \hat{y}_1 P_{22} - \frac{1}{r} P_{11} P_{12}$$

$$\frac{dP_{22}}{dt} = - \frac{1}{r} P_{12}^2$$

ecuaciones que corresponden al filtro de Extended Kalman para  $c=0$ , y a los gaussianos y de segundo orden para  $c=1$ .

Para estos algoritmos es especialmente significativo el estudio de su convergencia al verdadero valor del parámetro, por ser éste constante. Sustituyendo en el sistema anterior dz por su valor, e incluyendo a su vez la ecuación del sistema obtendremos /2/,

$$d\hat{y}_1 = (c P_{12} + \hat{y}_1 \hat{y}_2 + \frac{1}{r} y_1 P_{11} - \frac{1}{r} \hat{y}_1 P_{11}) \cdot dt + \frac{1}{r} P_{11} d\eta$$

$$d\hat{y}_2 = \frac{1}{r} P_{12} (y_1 - \hat{y}_1) dt + \frac{1}{r} P_{12} d\eta$$

$$dP_{11} = (2 \hat{y}_1 P_{12} + 2\hat{y}_2 P_{11} + q - \frac{1}{r} P_{11}^2) dt$$

$$dP_{12} = (\hat{y}_2 P_{12} + \hat{y}_1 P_{22} - \frac{1}{r} P_{11} P_{12}) dt$$

$$dP_{22} = - \frac{1}{r} P_{12}^2 dt$$

$$dy_1 = a y_1 dt + d\beta$$

De la observación de este sistema de ecuaciones diferenciales perturbado por ruido, se desprende que existen una infinidad de - estados de equilibrio para las variables --  $\hat{y}_2, P_{11}, P_{12}, P_{22}$  definidos por:

$$P_{12}=P_{22}=0, \quad \hat{y}_2, \quad P_{11} = r\hat{y}_2 \pm \sqrt{r^2 \hat{y}_2^2 + qr}$$

independientemente de los valores de  $y_1, \hat{y}_1$ .

Además, toda evolución conjunta del sistema

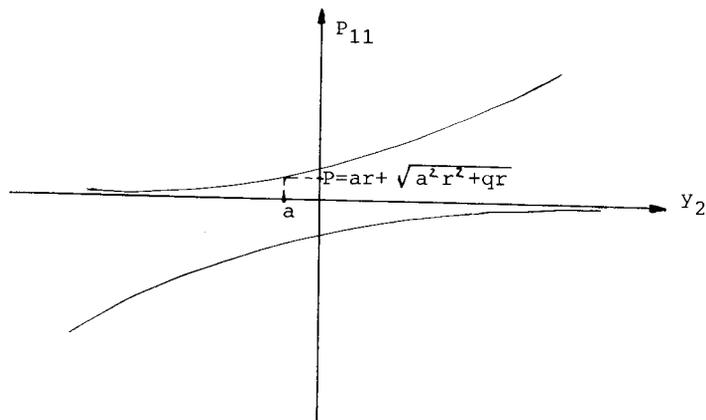


Fig. 4  
Conjunto de estados de equilibrio del filtro

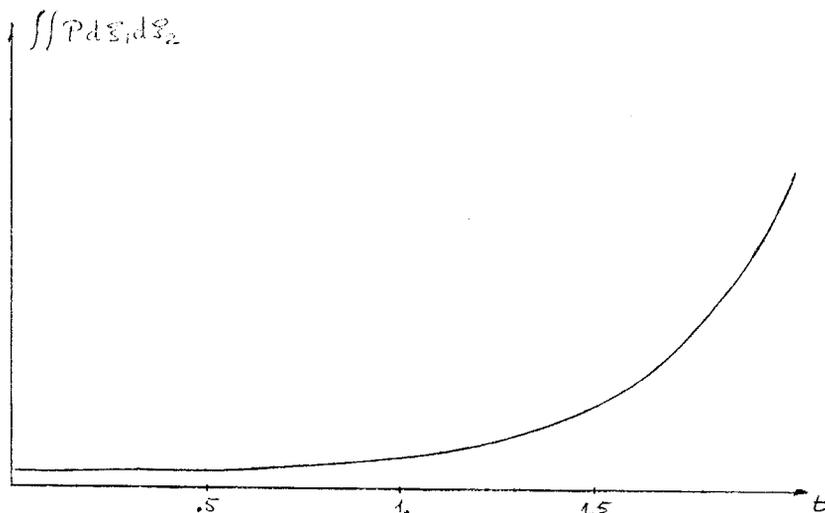


Fig. 5  
Evolución de la probabilidad de que  $P_{12}$  v  $P_{22}$  sean cercanas a cero

y filtro, o bien tiende cuando  $t \rightarrow \infty$  a uno de los anteriores estados, o bien diverge. En la figura 4 se ha representado el conjunto de estados de equilibrio así como el punto al cual tendría que converger el filtro para dar una estimación correcta ( $\hat{y}_2 = a$ ) del parámetro.

La aplicación de la teoría determinista de la estabilidad al sistema anterior suponiendo ruido nulo

$$d\beta = d\eta = 0$$

ha mostrado que el estado de equilibrio  $\hat{y}_2 = a$  es estable pero no asintóticamente estable, lo cual no permite asegurar la estabilidad de estos filtros.

Por otra parte, la aplicación a dicho sistema de la ecuación de Fokker Plank permite es

tudiar la evolución temporal de la densidad de probabilidad de estimación  $\hat{y}_2$  del parámetro, teniendo en cuenta los ruidos  $d\beta$  y  $d\eta$ . En efecto, si definimos la función  $P = P(\xi_1, \xi_2)$  tal que,

$$\begin{aligned} &P(\xi_1, \xi_2) d\xi_1 d\xi_2 d\xi_3 d\xi_4 = \\ &= \text{Prob} \left\{ \hat{y}_2 \in [\xi_1, \xi_1 + d\xi_1], P_{11} \in [\xi_2, \xi_2 + d\xi_2], \right. \\ &\quad \left. P_{12} \in \left[ -\frac{d\xi_3}{2}, \frac{d\xi_3}{2} \right], P_{22} \in \left[ -\frac{d\xi_4}{2}, \frac{d\xi_4}{2} \right] \right\} \end{aligned}$$

su evolución viene dada por

$$\frac{\partial P}{\partial t} = \left( \frac{3}{r} \xi_2 - 3\xi_1 \right) P - \left( 2\xi_1 \xi_2 + q - \frac{1}{r} \xi_2^2 \right) \frac{\partial P}{\partial \xi_2}$$

el cálculo numérico de  $\iint P d\xi_1 d\xi_2$  muestra que dicha integral crece rápidamente con el tiempo, como se observa en la figura 5, lo que

indica que va creciendo la probabilidad de -- que  $P_{12}$  y  $P_{22}$  tengan valores cercanos a ce-- ro. En otras palabras, la probabilidad de -- que en una simulación,  $P_{12}$  y  $P_{22}$  estén en un entorno del origen, va creciendo rápidamente con el tiempo. Este resultado confirma la -- afirmación inicial de que en los estados de equilibrio,

$$P_{12} = P_{22} = 0$$

Por otra parte puede verse, a partir de la -- ecuación diferencial para  $P$ , que con el tiem-- po la función

$$P(\xi_1, \xi_2)$$

crece indefinidamente en todos los puntos de la rama de hipérbola superior de estados de equilibrio (fig. 4) y decrece en los puntos restantes del plano  $\xi_1 \xi_2$ . Esto significa que crece la probabilidad de ir a cualquiera de los estados de equilibrio con  $P_{11} > 0$ , y no só-- lo al que cumple  $\hat{y}_2 = a$ . Por ello puede con-- cluirse que los algoritmos subóptimos estu-- diados no convergen al verdadero valor del -- parámetro del sistema.

En el caso de la estimación de los paráme-- tros de un sistema lineal de orden  $n$ , la -- aplicación de la ecuación de Fokker Plank -- conduce a las mismas conclusiones de no con-- vergencia de los algoritmos subóptimos estu-- diados.

## 5. CONCLUSION

El filtraje no lineal es una técnica con -- grandes posibilidades. El tipo de modelos -- con los que se trabaja es suficientemente ge-- neral como para tener aplicación en campos -- muy diversos. El enfoque del problema y la -- formulación empleada hacen pensar que podría utilizarse en problemas tales como los de -- previsión, tratamiento de series cronológi-- cas, etc. Hasta la fecha se ha utilizado con éxito en dominios en los que la estructura -- de los modelos dinámicos es bien conocida. -- (Comunicaciones, Astronáutica). Se realizan intentos de utilización en otros dominios en que los modelos dinámicos son mal conocidos (economía, sociología) e incluso se ha pre-- tendido utilizar estas técnicas para descu-- brir la estructura adecuada de dichos mode--

los. Sin embargo, parece claro que existien-- do graves problemas no resueltos en casos -- de formulación relativamente sencilla, como es el de estimación de parámetros, su utili-- zación en cualquier caso nuevo debería ir -- precedida de un análisis detallado de la -- adecuación del filtraje al problema, estu-- diando aspectos teóricos y considerando a -- la vez otros de implementación, como el -- tiempo de cálculo que en ciertos casos pue-- de constituir un límite efectivo para su -- utilización.

## 6. BIBLIOGRAFIA

- /1/ PAGES, J. "Contribución a las técnicas de síntesis de filtros no lineales". Tesis Doctoral, ETSIIB 1975.
- /2/ BRUNET, P. "Contribución al estudio de la estabilidad de filtros subóptimos y de sistemas estocásticos de dimensión -- elevada". Tesis Doctoral, ETSIIB 1976.
- /3/ JAZWINSKY, A.H. "Stochastic Processes -- and Filtering Theory". Academic Press, 1970.

