

# ESTIMACIÓN ADAPTATIVA EN TIEMPO REAL DE FUNCIONES DE TRANSFERENCIA REVISIÓN DE LAS TÉCNICAS DISPONIBLES Y PRESENTACIÓN DE NUEVOS ALGORITMOS

D. F. GARCÍA\*, D. DE LA FUENTE\*\*

Universidad de Oviedo

*En este artículo se analizan los problemas planteados en la estimación en tiempo real de los parámetros de sistemas variantes con el tiempo, con el objeto de definir los requisitos que debe verificar un estimador de este tipo. Seguidamente, se realiza un análisis crítico de las técnicas de estimación adaptativa más usuales, incidiendo en los aspectos fundamentales: la seguridad de funcionamiento del método, su capacidad de adaptación, la facilidad de uso del mismo y su coste computacional. Por último, en base a los análisis anteriores se desarrollan dos nuevos algoritmos que verifican todos los requisitos anteriores.*

**Real-Time Adaptive Estimation of Transfer Functions Review of Available Techniques and Presentation of New Solutions**

**Keywords:** Estimadores en tiempo real; Sistemas variantes en el tiempo; Seguimiento de parámetros variantes; Factores de olvido; Identificación.

---

\*Dpt. de Ingeniería Eléctrica, Electrónica de Computadores y Sistemas.

\*\*Dpt. de Admon. de Empresas y Contabilidad.

-Article rebut el setembre de 1988.

## 1. INTRODUCCIÓN

En la actualidad existe una gran cantidad de algoritmos de estimación de parámetros, que son adecuados para trabajar con sistemas variantes en el tiempo. Su utilización off-line no presenta problemas; no obstante, para su utilización on-line en un entorno industrial (en tiempo real), los algoritmos deben cumplir varios requisitos:

- Seguridad de funcionamiento
- Sencillez de uso
- Capacidad de adaptación
- Eficiencia computacional

Una gran parte de los métodos disponibles actualmente, no garantizan el control de los elementos de la matriz de covarianzas del estimador y por tanto, su utilización en un sistema industrial es totalmente desaconsejable.

Otros métodos requieren un delicado ajuste de uno o más parámetros, resultando muy difícil la correcta realización de este tipo de correcciones.

La escasa investigación sobre los requisitos necesarios para la aplicabilidad de este tipo de estimadores, ha sido la principal motivación para la realización de este trabajo. Conceptualmente, las aportaciones del trabajo se pueden agrupar en tres grandes bloques: en primer lugar, se realiza *un detallado análisis de los problemas que plantea la estimación en tiempo real de procesos*, cuyos parámetros varían en el tiempo. La segunda aportación, consiste en la *realización de un análisis comparativo y crítico de los métodos de estimación on-line disponibles en la actualidad*, determinando su aplicabilidad a procesos industriales. Por último, basándose en los estudios y análisis de los dos puntos anteriores se han desarrollado dos *nuevos métodos que presentan notables ventajas sobre los precedentes*.

Físicamente, el artículo está estructurado en cinco secciones: en primer lugar se realiza una introducción al problema, en la segunda sección se analizan los problemas que se plantean en la estimación adaptativa de parámetros. El siguiente apartado está dedicado al estudio de los métodos actualmente disponibles, desarrollándose nuevos métodos en la cuarta y por último se resumen las conclusiones de este trabajo.

## 2. PROBLEMAS BÁSICOS QUE PLANTEA LA ESTIMACIÓN DE PARÁMETROS EN TIEMPO REAL

El primer aspecto a considerar en un problema de estimación adaptativa, es el tipo de variaciones que pueden experimentar los parámetros del proceso. Estas variaciones, se pueden clasificar en dos grandes grupos:

- Variaciones *abruptas*
- Variaciones *moderadas* y/o derivas suaves

En el primer caso, el problema clave es detectar el cambio y tomar las acciones apropiadas para controlarlo; en el otro, el estimador debe estar en alerta permanente. Es posible desarrollar estimadores que traten cada uno de los casos anteriores de forma específica; no obstante, el objetivo del artículo es el análisis y desarrollo de estimadores para el segundo problema principalmente, sin perjuicio de que se añadan a estos estimadores mecanismos de detección y aceleración de la convergencia, para tratar de modo específico las variaciones abruptas de parámetros.

La *eliminación de información* del estimador es necesaria, para seguir correctamente los parámetros de un proceso variante en el tiempo, pero plantea otros problemas graves si no se realiza adecuadamente.

El problema a resolver se resume en la pregunta: ¿cómo desechar la información?. Para responder a esta pregunta, es preciso considerar al conjunto de parámetros como un espacio vectorial, cuya dimensión coincide con el número de parámetros a estimar. La información que entra al estimador en cada muestreo no está uniformemente distribuida en el tiempo, ni en el espacio de parámetros. Por lo tanto, se debe desechar la información teniendo en cuenta que es preciso:

- eliminar más información en unas componentes del espacio de parámetros que en otras, y
- calculando la cantidad exacta de información a eliminar

según el volumen de información que entra al conjunto de parámetros, cada vez que se actualizan los datos.

Del análisis anterior, se deduce que para adaptarse a variaciones en los parámetros, es preciso olvidar información, dando entonces más peso a los datos más recientes. Obviamente, si la velocidad con que se elimina la información aumenta, el estimador puede seguir con mayor rapidez las variaciones de los parámetros.

Este análisis es válido presuponiendo que todos los datos empleados en la estimación son correctos, sólo que unos contienen más información y otros menos

debido a las diferencias en la amplitud de las señales utilizadas por el estimador. No obstante si durante un cierto período de tiempo se “alimenta al estimador” con datos erróneos o atípicos, pero ricos en información, las estimaciones no convergerán a los valores correctos. Si se sospecha que pueden existir datos atípicos, es preciso incorporar un detector de los mismos para proceder a su eliminación o sustitución antes de la estimación.

Cuando no se producen perturbaciones externas que exciten al sistema, las señales de entrada y salida no contienen información útil para el estimador y las filas de la matriz de información se hacen linealmente dependientes a medida que transcurre el tiempo. Si en cada paso se “olvida” una gran cantidad de información, el estimador desecha rápidamente la información antigua sobre el proceso, que era la correcta y recibe información de los nuevos datos que son poco fiables. Como contrapartida, se tiene una gran capacidad para seguir variaciones rápidas de la dinámica del proceso. Si la excitación es elevada, los datos de entrada-salida contienen mucha información y entonces, se puede “olvidar” con rapidez, siendo la capacidad de adaptación elevada. En ausencia de excitación no se debe olvidar nada y la capacidad de adaptación es nula.

En resumen, la capacidad que posee un estimador para adaptarse a un proceso variante es, básicamente, función de:

- La forma en que el estimador “olvida” la información para adaptarse a las variaciones del sistema.
- La excitación a la que se encuentra sometido el propio sistema.

Para resolver estos problemas, se sigue el siguiente enfoque: desarrollar mecanismos que adecúen el ratio de “olvido” (capacidad de adaptación) a la excitación. El objetivo de estos mecanismos, es evitar la “explosión” del estimador y permitir el seguimiento de parámetros cuya velocidad de variación es lenta o media. Para tratar variaciones abruptas (muy rápidas) y de gran magnitud se pueden incorporar en los estimadores mecanismos adicionales de detección e intervención.

El análisis de los diversos mecanismos de olvido, se ha realizado empleando *estimadores mínimo-cuadráticos*. Cuando se utiliza una ponderación exponencial constante de los datos mediante la inclusión de un *factor de olvido exponencial fijo*,  $\beta$ , el algoritmo es el siguiente:

$$(1) \quad e(t) = y(t) - z(t)^T \theta(t-1)$$

$$(2) \quad K(t) = \frac{P(t-1)z(t)}{\beta + z(t)^T P(t-1)z(t)}$$

$$(3) \quad \theta(t) = \theta(t-1) + K(t)e(t)$$

$$(4) \quad P(t) = [I - K(t)z(t)^T] \frac{P(t-1)}{\beta}$$

donde  $y(t)$  es la medición de la respuesta del sistema en el instante actual,  $z(t)$  es el vector de observaciones, constituido por la entrada al sistema en el instante actual  $u(t)$ , e instantes anteriores  $u(t-1), u(t-2), \dots$  y por las respuestas del sistema en instantes anteriores,  $y(t-1), y(t-2), \dots$   $\theta(t)$  denota al vector de parámetros estimados en el instante actual.  $P(t)$  es la matriz de covarianzas del estimador y al vector  $K(t)$  se le denomina Ganancia de Kalman en la literatura especializada, ya que actúa como una ganancia por la cual se multiplica el error de la estimación y cuyo resultado se utiliza para actualizar el vector de parámetros, como se puede comprobar en la ecuación (3).

La inclusión de este factor  $\beta$ , es la forma más sencilla de cambiar la función de coste para incluir una ponderación exponencial de los residuos a posteriori.

El empleo de un factor  $\beta < 1$ , evita que la matriz  $P(t)$  converja a cero, pero si el proceso no está convenientemente excitado,  $P(t)$  crece exponencialmente. Si el estimador trabaja con datos diferenciados, el vector de observaciones,  $z(t)$ , tiende a cero y la ecuación de actualización de  $P(t)$  se convierte en:

$$(5) \quad P(t) = \frac{P(t-1)}{\beta}$$

Este fenómeno de crecimiento sin límite de la matriz de covarianzas se denomina “blow-up”, entonces el estimador se vuelve inestable, ya que  $P(t)$  aparece como una ganancia en el algoritmo y los parámetros evolucionan de una forma bastante errática.

Como ejemplo ilustrativo, se muestran los problemas que plantea la utilización de un factor de olvido exponencial fijo (el mecanismo de olvido más sencillo), ante diversas excitaciones del sistema. Para simular estas condiciones de funcionamiento se excita el siguiente sistema de primer orden:

$$(6) \quad (1 - a_1 z^{-1})y(t) = b_1 z^{-1}u(t) + v(t) \quad [a_1 = 0.9; b_1 = 0.1]$$

con una señal  $u(t)$  de tipo PRBS (“Pseudo Random Binary Signal”, Señal Binaria Seudoaleatoria), durante ciertos períodos, siendo nula entre ellos. Además, el sistema está excitado por un ruido blanco  $v(t)$ . El operador  $z^{-1}$  es el operador retardo unitario [ $z^{-1}y(t) = y(t-1)$ ]. Se ha elegido un factor de olvido fijo de 0.9 para mostrar claramente el efecto del “blow-up”.

Los resultados se muestran en la figura 1, en la que se puede comprobar que los dos elementos de la diagonal de la matriz de covarianzas,  $p_1$  y  $p_2$ , evolucionan

de forma exponencial cuando falta la excitación. No obstante, el elemento  $p_2$  se llega a estabilizar en valores muy altos, a causa de la ligera excitación aportada por  $v(t)$ . Debido al efecto de este ruido en combinación con los elevados valores de la matriz de covarianzas, el parámetro  $a_1$  presenta los clásicos síntomas de “burst” (evolución errática de los parámetros estimados), sin embargo, el parámetro  $b_1$  permanece estacionario. Las tres gráficas comparten el mismo eje de abscisas en el que se representa el número de muestra, o tiempo discreto. Esta simulación, ilustra claramente los dos problemas a resolver: el “blow-up” y la excitación no uniforme de todos los elementos del vector de parámetros.

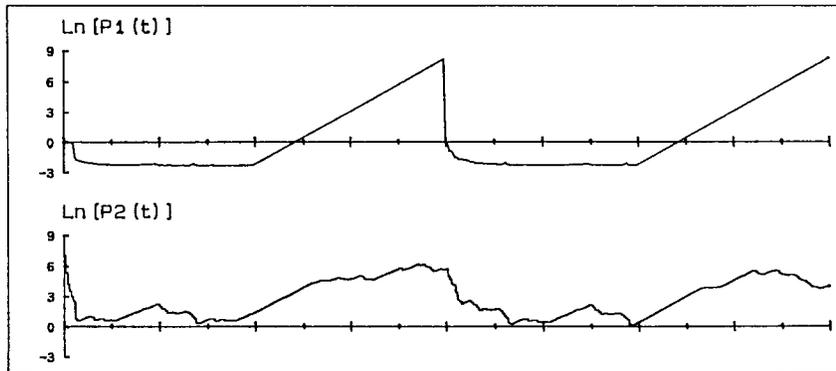
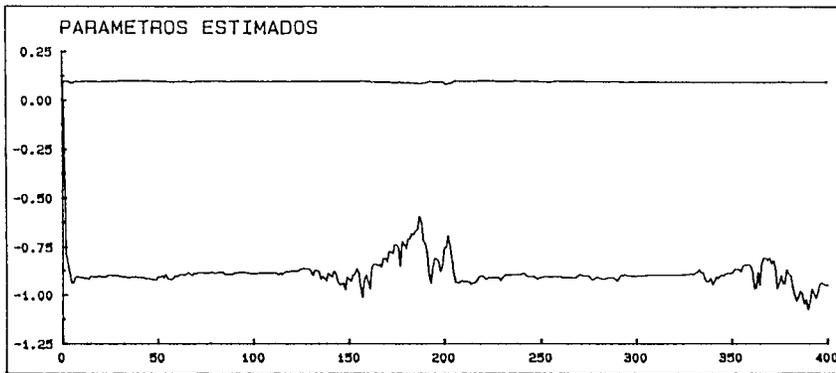
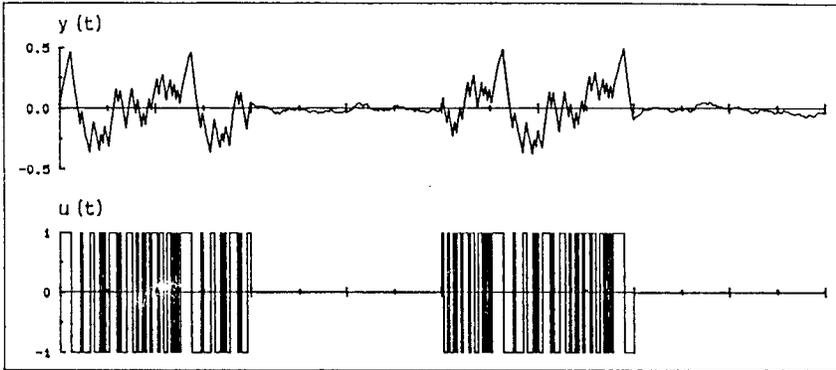


Figura 1. Problemas que plantea la utilización de un factor de olvido escalar fijo.

### 3. ESTUDIO COMPARATIVO Y CRÍTICO DE LOS MÉTODOS MÁS COMUNES PARA LA ESTIMACIÓN ADAPTATIVA PARÁMETROS VARIANTES

A continuación, se estudian diversos métodos para modificar los estimadores recursivos de modo que puedan cumplir los dos objetivos antagónicos, de seguir fielmente parámetros variantes y evitar el “blow-up” en ausencia de excitación. Estos métodos se han agrupado del siguiente modo:

- Factores de olvido exponenciales-escalares
- Factores de olvido exponenciales-vectoriales
- Modificación directa de la matriz de covarianzas
- Utilización de ventanas de datos finitas
- Activación y desactivación del estimador
- Métodos de olvido exponenciales-direccionales
- Métodos de olvido exponenciales-direccionales-variables
- Otras técnicas

Seguidamente, se analiza cada uno de los grupos propuestos.

#### 3.1 Factores de Olvido Exponenciales Escalares

Para evitar los problemas mencionados anteriormente, se puede realizar una *ponderación exponencial de los datos*, con un factor de olvido variable. Todos estos algoritmos utilizan el método de los mínimos cuadrados, o lo que es lo mismo, las ecuaciones (1-4), pero añaden una ecuación extra para adaptar el factor de olvido  $\beta$ . A continuación se presentan varios algoritmos basados en esta idea.

*Fortescue, Kershenbaum e Ydstie* (1981), han propuesto un factor de olvido variable, cuyo objetivo es mantener constante la suma de los cuadrados de los errores de estimación a posteriori. Esta magnitud puede expresarse recursivamente en función del error de estimación a priori, del modo siguiente:

$$(7) \quad \sum(t) = \beta(t) \sum(t-1) + [1 - z(t)^T K(t)] e(t)^2$$

El objetivo propuesto, se introduce en (7) mediante la siguiente igualdad:

$$(8) \quad \Sigma(t) = \Sigma(t-1) = \dots = \Sigma_0$$

De este modo, la cantidad de información a olvidar en cada paso, se elige en función de la información aportada por los nuevos datos, asegurándose que la estimación siempre está basada en la misma cantidad de información. Realizando las operaciones convenientes, se puede establecer la siguiente expresión para el factor de olvido exponencial escalar:

$$(9) \quad \beta(t) = 1 - (1/\Sigma_0) \frac{e(t)^2}{1 + z(t)^T P(t-1) z(t)}$$

Para elegir el valor de  $\Sigma_0$ , se le descompone del siguiente modo:

$$(10) \quad \Sigma_0 = \sigma^2 N_0$$

donde  $\sigma^2$ , representa la varianza esperada del ruido de medida, basándose en el conocimiento real del proceso.  $N_0$  se elige, estableciendo un compromiso entre la sensibilidad del sistema y la velocidad de adaptación deseada.

*Cordero y Mayne* (1981), demostraron que el algoritmo de Fortescue sólo garantiza la eliminación del "blow-up" en el caso determinista y en ausencia de ruido de medida. Ejemplos típicos de la incapacidad de este algoritmo para evitar el "blow-up", han sido presentados por Saelid y Foss (1983), entre otros.

*Astrom* (1980), ha propuesto otro método para ajustar automáticamente el factor de olvido, mediante la siguiente expresión:

$$(11) \quad \beta(t) = 1 - \alpha(e^2/\underline{e}^2)$$

donde  $\underline{e}^2$  representa el valor medio de  $e^2$  sobre un determinado período y  $e$ , es el error de estimación a priori. El parámetro  $\alpha$  es una constante de pequeño valor, por ejemplo 1/1000. Se puede demostrar que este algoritmo es muy similar al de Fortescue.

*Wellstead y Sanoff* (1981), han propuesto el siguiente algoritmo para ajustar el factor de olvido:

$$(12) \quad \beta(t) = \beta_1(t) \cdot \beta_2(t)$$

$$(13) \quad \beta_1(t) = \alpha + (1 - \alpha) [1 - \text{EXP}(-t/\tau)] \quad 0.9 \leq \alpha \leq 1.0$$

$$(14) \quad \beta_2(t) = 1 - \frac{e(t-1)^2}{\tau s(t)}$$

$$(15) \quad s(t) = \frac{(\tau - 1)}{\tau} s(t-1) + \frac{e(t-1)^2}{\tau}$$

donde  $\tau$  es la memoria deseada para el estimador. El algoritmo es muy similar al propuesto por Astrom (1980) y Fortescue et al. (1981).

*Tuffs* (1985), propone otra medida de la información que debe mantener constante el estimador: la norma de  $P(t)$ , ponderada con el vector de observaciones actual. Esto es,  $z(t)^T P(t-1)z(t)$ . De esta forma, el factor de olvido resultante es:

$$(16) \quad \sigma(t) = \frac{z(t)^T P(t-1)z(t)}{1 + z(t)^T P(t-1)z(t)}$$

$$(17) \quad \beta(t) = \text{MAX}(1 - \sigma(t), \beta_{\min})$$

donde  $\sigma(t)$ , es una medida de información propuesta por Clarke y Gawthrop (1979), para analizar la convergencia de los reguladores autoajustables y  $\beta_{\min}$  el valor mínimo preestablecido que puede tomar el factor de olvido.

Los algoritmos de Fortescue *et al.* (1981), Astrom (1980) y Wellstead *et al.* (1981) se basan en la misma medida de información: el cuadrado del error de estimación, sometido a diversas ponderaciones, todas ellas similares. Estos algoritmos tratan de mantener esta medida a un nivel constante seleccionado por el usuario. Hemos comprobado que esta medida de la información es inadecuada, ya que:

- Su comportamiento es imprevisible en el caso de parametrizaciones incorrectas.
- Es muy sensible ante cambios de las perturbaciones estocásticas que actúan sobre la planta.
- Es muy sensible ante el ruido de medida.

Con ninguno de estos algoritmos se elimina con seguridad el problema del "blow-up" y plantean otro problema adicional, que es, el difícil ajuste del parámetro de ponderación que requieren. El algoritmo propuesto por Tuffs (1985), aunque se comporta mejor que los anteriores y es muy sencillo de configurar, tampoco elimina totalmente el problema del "blow-up", en gran parte debido a que la matriz de covarianzas puede incluir elementos con excitaciones muy diferentes.

Irving (1979), ha propuesto mantener la traza de la matriz  $P$  constante, a un determinado valor, por ejemplo  $\text{Tr}_0$ . El algoritmo es el siguiente:

$$\begin{aligned}
 &P_1(t) = P(t-1) - K(t)z(t)^T P(t-1) \\
 &\text{SI } \text{Tr}[P(t)] \geq \text{Tr}_0 \text{ ENTONCES} \\
 &\beta(t) = 1.0; \quad P(t) = P_1(t) \\
 (18) \quad &\text{SINO} \\
 &\beta(t) = \frac{\text{Tr}[P-1(t)]}{\text{Tr}_0}; \quad P(t) = \frac{P_1(t)}{\beta(t)} \\
 &\text{FIN - SI}
 \end{aligned}$$

Este método elimina el problema del “blow-up” de la matriz de covarianzas, pero surgen problemas cuando  $z(t)$  se hace constante, o sea, cuando el sistema no está persistentemente excitado. Entonces, la matriz  $P$  se hace singular y para evitarlo, es necesario emplear una técnica de regularización. Irving (1979), propone el siguiente método para modificar la fórmula de actualización de la matriz de covarianzas:

$$(19) \quad P(t) = P_2(t)/\beta(t)$$

$$(20) \quad P_2(t) = P_1(t) + \sigma \{ \text{Diag} P_1(t) + \text{Tr} P_1(t) \cdot I \}$$

con esta regularización, se evita que los valores propios de la matriz  $P(t)$  se reduzcan demasiado. Desgraciadamente, este algoritmo no resuelve los problemas originados por las excitaciones no uniformes, en el espacio de parámetros.

Después de programar y utilizar todos estos algoritmos podemos decir como conclusión, que todos los métodos basados en factores de olvido escalares, son incapaces de resolver los problemas causados por una excitación no uniforme en el espacio de parámetros. Por lo tanto, es preferible emplear otras técnicas que describiremos a continuación.

### 3.2 Factores de Olvido Exponenciales Vectoriales

Empleando un factor de olvido escalar, los elementos de  $P(t)$  asociados con datos de baja excitación se incrementarán y los asociados con datos de excitación elevada, disminuirán. Por tanto, un único factor de olvido no es la solución idónea para manejar conjuntamente todos los datos. *Saelid y Foss* (1983) han propuesto el empleo de un factor de olvido variable vectorial, o lo que es lo mismo, el empleo de *un factor de olvido escalar para cada elemento del vector de observaciones*.

Conceptualmente, la extensión al caso vectorial de un factor de olvido exponencial escalar es sencilla. El problema subyacente, es definir el criterio de ajuste automático de los factores de olvido; para ello Saelid y Foss (1983) han propuesto como objetivo a conseguir, el mantener a un nivel constante los elementos de la diagonal de  $P(t)$ , asegurando por tanto que la traza de  $P(t)$ , la cual es una buena medida de la información contenida en el estimador, permanezca constante. Si se emplea la matriz de covarianzas en forma factorizada ( $UDU^T$ , por ejemplo, donde  $U$  es triangular y  $D$  diagonal), es difícil utilizar la traza como la medida de información a mantener constante, siendo en este caso más sencillo mantener el determinante (diagonal de  $P$  en la forma  $UDU^T$ ), que también es una excelente medida de la información contenida en el estimador. A continuación se describe este método, donde la ecuación de actualización de la matriz  $P$  es:

$$(21) \quad P(t) = P(t-1) - \frac{P(t-1)z(t)z(t)^T P(t-1)}{1 + z(t)^T P(t-1)z(t)}$$

Teniendo en cuenta la expresión del vector de corrección,  $K(t)$ , definida por la ecuación (2), y la fórmula (21), los elementos de la diagonal de  $P$  se pueden actualizar con la siguiente fórmula:

$$(22) \quad P_{ii}(t) = P_{ii}(t-1) - [1 + z(t)^T P(t-1)z(t)] [K_i(t)]^2$$

El objetivo es conseguir que  $P_{ii}(t) = P_{ii}(t-1)$ . Para ello, se introduce un factor de olvido para cada parámetro,  $\beta_i(t)$ . Estos factores de olvido  $\beta_i(t)$ , no tienen ninguna relación directa con el factor de olvido escalar  $\beta$ . Bajo estas condiciones, la fórmula de actualización de la matriz de covarianzas es:

$$(23) \quad P(t) = L(t)^{-1} [I - K(t)z(t)^T] P(t)L(t)^{-1}$$

donde  $L(t) = \text{diag} \{\beta_1(t) \cdots \beta_n(t)\}$  y  $n = \dim\theta$ .

Empleando esta definición del factor de olvido, la actualización de la diagonal de la matriz de covarianzas es la siguiente:

$$(24) \quad P_{ii}(t) = \left\{ P_{ii}(t-1) - [1 + z(t)^T P(t-1)z(t)] [K_i(t)]^2 \right\} / \beta_i(t)$$

Suponiendo que  $P_{ii}(t) = P_{ii}(t-1)$ , se despeja  $\beta_i(t)$  en (24), obteniéndose:

$$(25) \quad \beta_i(t) = 1 - [1 + z(t)^T P(t) z(t)] [K_i(t)]^2 / P_{ii}(t)$$

Usando la fórmula (25) para actualizar  $P(t)$ , los elementos de la diagonal permanecerán constantes para todo  $t$ , mientras que los elementos fuera de la diagonal cambiarán de acuerdo con las variaciones de  $z(t)$ .

No obstante, la aplicación de la fórmula (25) puede resultar peligrosa, ya que si la excitación es muy elevada, los factores de olvido resultantes, con objeto de mantener la diagonal de  $P(t)$  constante, pueden ser muy reducidos (próximos a cero) volviéndose el estimador muy sensible ante el ruido.

Tuffs (1985), utilizó este método imponiendo un límite inferior a los factores de olvido; de este modo, se evita que el estimador olvide la información que tiene almacenada a una velocidad demasiado elevada. No obstante, esta técnica plantea problemas; cuando la excitación es muy elevada y no se permite que los factores de olvido tomen valores inferiores a  $\beta_{\min}$ , la traza disminuye. Una vez pasada la etapa de excitación, el algoritmo tiende a mantener el nuevo valor de la traza y después de sucesivos períodos de elevada excitación, el estimador llega a “congelarse”, o sea, la matriz de covarianzas es llevada a un valor muy bajo del que no puede recuperarse, perdiendo entonces el estimador toda su capacidad para seguir las variaciones de parámetros. Tuffs (1985), indicó la dificultad de encontrar mecanismos de recuperación sin “blow-up”. Un ejemplo de los problemas planteados por este método se muestra más adelante.

### 3.3 Modificación Directa de la Matriz de Covarianzas

El método, consiste en *adicionar una matriz  $Q(t)$ , a la de covarianzas*, en cada paso de actualización de ésta:

$$(26) \quad P(t) = [I - K(t)z(t)^T] P(t-1) + Q(t)$$

Con una adecuada elección de la secuencia de matrices  $Q(t)$ , se consigue que la matriz de covarianzas no converja a la matriz nula. No obstante, es posible que los elementos de  $P(t)$  crezcan demasiado y por lo tanto, es conveniente supervisar la traza de  $P(t)$ , asignando temporalmente  $Q(t)$  a cero, si la traza supera un valor máximo predeterminado.

Este método, es análogo al empleado con el filtro de Kalman para estimar un vector de estado (27), perturbado aleatoriamente:

$$(27) \quad x(t) = x(t-1) + v(t)$$

$$(28) \quad E\{v(t) \cdot v(t)^T\} = Q$$

donde  $v(t)$ , es un vector de variables aleatorias, con una matriz de covarianzas conocida  $Q$ . Bajo estas circunstancias, la actualización óptima de la matriz de covarianzas del filtro viene dada al sustituir en la fórmula (26),  $Q(t)$  por  $Q$ . En definitiva, con este método el problema fundamental, consiste en determinar el criterio de construcción de la secuencia  $Q(t)$ .

*Saelid, Egeland y Foss* (1985), proponen mantener la traza de la matriz de covarianzas constante. Si se observa la fórmula (29) de actualización de esta matriz:

$$(29) \quad P(t) = [I - K(t)z(t)^T] P(t-1) + Q(t)$$

se deduce inmediatamente, que eligiendo:

$$(30) \quad Q(t) = \text{diag}\{q_{11}(t) \cdots q_{nn}(t)\} \quad n = \dim\theta \quad y$$

$$(31) \quad q_{ii}(t) = [K(t)z(t)^T P(t-1)]_{ii}$$

se obtiene que  $P_{ii}(t) = P_{ii}(t-1)$ . Como se puede comprobar, el algoritmo es muy simple, ya que empleando este enfoque, la actualización de los elementos fuera de la diagonal de  $P$  viene dada por:

$$(32) \quad P_{ij}(t) = [(I - K(t)z(t)^T) P(t-1)]_{ij} \quad i \neq j$$

mientras que con el método del factor de olvido vectorial (*Saelid y Foss, 1983*), se tiene:

$$(33) \quad P_{ij}(t) = \frac{[(I - K(t)z(t)^T) P(t-1)]_{ij}}{[\beta_i(t)\beta_j(t)]^{1/2}}$$

Como se puede ver, la única diferencia entre ambos métodos está en el factor  $[\beta_i(t)\beta_j(t)]^{1/2}$ . Por tanto, con el factor de olvido vectorial los elementos de  $P$  fuera de la diagonal son "inflados", mientras que con la condición de  $Q(t)$ , estos elementos siempre disminuyen, pues son actualizados sin olvido. A continuación, se analiza este algoritmo. Suponiendo que se desea mantener la diagonal de la matriz  $P(t)$  a un valor  $P^*$ , la matriz  $Q(t)$  a añadir a  $P(t)$ , oscila entre los siguientes valores:

$$(34) \quad 0 \leq Q(t) \leq P^*$$

Si se permite que la matriz  $Q(t)$  tome cualquier valor dentro del rango anterior, no tiene sentido aplicar este algoritmo, ya que es mucho más cómodo actualizar la matriz de covarianzas, con un factor de olvido escalar igual a la unidad y seguidamente fijar los elementos de la diagonal de la matriz  $P(t)$ , al valor  $P^*$ .

Cuando la excitación es elevada,  $Q(t)$  crece para mantener la traza y como consecuencia los parámetros estimados tienden a seguir al ruido.

*Goodwin, Hill y Palaniswami* (1985), han propuesto el siguiente algoritmo para actualizar la matriz de covarianzas:

$$(35) \quad P'(t) = P(t-1) - \frac{P(t-1)z(t)z(t)^T P(t-1)}{1 + z(t)^T P(t-1)z(t)}$$

$$(36) \quad \Gamma = \text{Traza}[P'(t)]$$

$$(37) \quad \text{SI } \Gamma > C_0 \text{ ENTONCES } P(t) = P'(t) + \frac{C_1 - \Gamma}{n}$$

$$(38) \quad \text{SI } \Gamma < C_0 \text{ ENTONCES } P(t) = \frac{C_0}{\Gamma} P'(t) + \frac{C_1 - C_0}{n} I$$

donde  $C_1 > C_0 > 0$ .

$n$  = número de parámetros.

El algoritmo, da una ganancia de estimación constante (traza de  $P(t)$  constante) y presenta unas características intermedias entre los algoritmos de tipo gradiente y los mínimo-cuadráticos. La constante  $C_0$ , es muy fácil de elegir, pero no la  $C_1$ , de la que no se tiene ninguna indicación sobre la forma de ajustarla.

### 3.4 Utilización de Ventanas de Datos Finitas

Ésta, es otra de las posibles técnicas a emplear para eliminar datos antiguos con objeto de mantener al estimador "vivo". La idea, consiste en *descartar los datos previos a un determinado instante, ponderando por igual a todos los datos posteriores a ese instante*. Gráficamente, esto se puede interpretar como una función, o ventana, de ponderación de forma rectangular que se desplaza con los datos. La longitud de esta ventana puede ser fija o variable. A continuación, se analizan algoritmos basados en esta técnica.

En los estimadores que utilizan explícitamente ventanas rectangulares finitas, tales como la mostrada en la figura 2, el problema radica en seleccionar la

longitud del período de observación (ventana), para establecer un compromiso idóneo entre la sensibilidad del estimador ante ruidos y la capacidad de éste para seguir parámetros variantes.

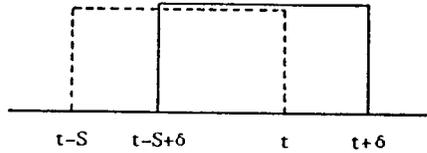


Figura 2. Ponderación rectangular de los datos.

El funcionamiento es el siguiente: después de completar cada solución basada en  $S$  muestras, se toman  $\delta$  nuevas muestras y a continuación se eliminan las  $\delta$  últimas muestras de la ventana. Por tanto, se pueden resolver las ecuaciones de estimación superponiendo (restando) las soluciones correspondientes a las dos ventanas. Para la utilización recursiva de este método en tiempo real, tan sólo tiene interés el caso en el que  $\delta = 1$ . Entonces, el algoritmo recursivo para implementar este mecanismo es el siguiente, *Young (1984)*:

1) Adición del nuevo dato en el  $t$ -ésimo instante

$$(39) \quad e(t) = y(t) - z(t)^T \theta(t-1)$$

$$(40) \quad \theta'(t) = \theta(t-1) + \frac{P(t-1)z(t)}{1 + z(t)^T P(t-1)z(t)} e(t)$$

$$(41) \quad P'(t) = P(t-1) + \frac{P(t-1)z(t)z(t)^T P(t-1)}{1 + z(t)^T P(t-1)z(t)}$$

2) Eliminación del dato recibido en el  $(t-s)$ -ésimo instante

$$(42) \quad e'(t) = y(t-s) - z(t-s)^T \theta'(t)$$

$$(43) \quad \theta(t) = \theta'(t) + \frac{P'(t)z(t-s)}{-1 + z(t-s)^T P'(t)z(t-s)} e'(t)$$

$$(44) \quad P(t) = P'(t) + \frac{P'(t)z(t-s)z(t-s)^T P'(t)}{-1 + z(t-s)^T P'(t)z(t-s)}$$

Como se deduce de las fórmulas (39-44), en el caso más sencillo ( $\delta = 1$ ), como mínimo es preciso duplicar las operaciones realizadas en un estimador que utilice olvido exponencial. Además, es preciso usar dos matrices de covarianzas y recomponer el vector  $z(t-s)$  a partir de las observaciones almacenadas.

Por todo ello, la implementación de estimadores que utilicen ventanas rectangulares, es bastante más compleja que la de los que utilicen ventanas exponenciales, sin que de ello se deriven ventajas adicionales. Por otra parte, para obtener un estimador industrial fiable, es preciso añadir un mecanismo que permita modificar la longitud de la ventana, en función de la información entrante al estimador en cada momento.

Del análisis previo, se deduce que las ventanas rectangulares no son adecuadas (frente a las exponenciales), para ser incorporadas en un estimador adaptativo que deba funcionar en tiempo real.

Los estimadores basados en la reinicialización periódica de  $P(t)$ , pueden analizarse como un caso particular de la utilización de ventanas. Simplemente consiste en emplear un método de estimación secuencial con memoria infinita, cuya matriz de covarianzas es reinicializada periódicamente a un valor  $K \cdot I$ .

El problema fundamental es ¿cuándo se reinicializa la matriz de covarianzas?. Si los parámetros varían a saltos, se puede detectar el cambio y reinicializar  $P(t)$  a continuación; en cambio, si los parámetros del sistema derivan lentamente, es obligatorio elegir un intervalo para las sucesivas reinicializaciones.

El error cuadrático medio, debido al efecto del ruido decrece de forma inversamente proporcional a la longitud del intervalo, debido al efecto de promediado, pero el error debido a la deriva de los parámetros se incrementa proporcionalmente a la longitud del intervalo, debido al efecto acumulativo. Esto se puede expresar mediante la siguiente ecuación:

$$(45) \quad e_m^2 = K_1/L + K_2 * L$$

$L$  = Longitud de la ventana o intervalo de reinicialización

$K_1$  es proporcional a la varianza del ruido

$K_2$  es proporcional a la velocidad con que cambia el proceso

El valor óptimo de  $L$ , ( $L^*$ ) es un compromiso entre los dos efectos:

$$(46) \quad L^* = [K_1/K_2]^{1/2}$$

y el error cuadrático medio resultante es:

$$(47) \quad e_m^2 = 2[K_1 * K_2]^{1/2}$$

Otro método para determinar el instante de reinicialización del algoritmo, consiste en supervisar la magnitud de la matriz  $P(t)$ , mediante por ejemplo, su traza; entonces, se reinicializa el estimador si la traza disminuye por debajo de un umbral definido por el usuario.

La convergencia del método de la reinicialización periódica de  $P(t)$ , en conjunción con un controlador autoajutable, ha sido demostrada por Goodwin, Elliot y Teoh (1983) para el caso determinista. Desde un punto de vista teórico, Ydstie *et al.* (1985) indican que con la reinicialización de la matriz de covarianzas puede obtenerse un control deficiente; esto, se debe al hecho de que todas las direcciones de búsqueda del estimador son reescaladas y ortogonalizadas, cuando se reinicializa la matriz de covarianzas a un valor diagonal u otro cualquiera. De esta forma, el reajuste de los parámetros lleva un cierto tiempo; por el contrario, un factor de olvido escalar exponencial, escala por igual toda la matriz de covarianzas.

La reinicialización de  $P(t)$  es correcta, cuando las direcciones de actualización de parámetros dejan de ser válidas; por ejemplo, al producirse un cambio brusco de parámetros. Fuera de este contexto, es preferible aplicar otros métodos.

Como se puede deducir del estudio previo, estos algoritmos no resultan adecuados para ser implementados en un estimador adaptativo industrial, ya que aunque previenen el “blow-up”, hacen oscilar innecesariamente a los parámetros estimados, o su coste computacional es demasiado elevado. Entonces, es conveniente aplicar otros métodos.

### 3.5 Activación y Desactivación del Estimador

La *activación y desactivación* del estimador, es otra de las técnicas que se pueden utilizar para evitar el “blow-up” de la matriz de covarianzas. El problema consiste en decidir cuando se arranca y cuando se para el estimador.

*Schumann, Lachmann e Iserman* (1981), proponen activar y desactivar el estimador, en función de la varianza de las señales de salida del proceso, manteniendo constante el valor del factor de olvido escalar elegido  $\beta$ , siendo  $\beta < 1$ .

*Bergmann* (1982), propone activar y desactivar el estimador, o pasar de  $\beta = 1$  a  $\beta < 1$  (fijo), en función de las varianzas de las señales  $u(t)$  e  $y(t)$ .

*Isermann* (1985), propone la utilización de la matriz de información  $H(t)$  (o

matriz de correlaciones) para conectar y desconectar el estimador. Su estructura en el caso monovariable es la siguiente:

(48)

$$\left[ \begin{array}{ccc|ccc} \Phi_{yy}(0)\Phi_{yy}(1)\cdots\Phi_{yy}(n-1) & -\Phi_{uy}(d) & \cdots & -\Phi_{uy}(d+m-1) & & \\ \Phi_{yy}(0)\cdots\Phi_{yy}(n-2) & -\Phi_{uy}(d-1) & \cdots & -\Phi_{uy}(d+n-2) & & \\ \vdots & \vdots & & \vdots & & \\ \Phi_{yy}(0) & -\Phi_{uy}(d-n+1) & \cdots & -\Phi_{uy}(d) & & \\ \hline & \Phi_{uu}(0) & \cdots & \Phi_{uu}(n-1) & & \\ & & \cdots & \Phi_{uu}(n-2) & & \\ & & & \vdots & & \\ & & & \Phi_{uu}(0) & & \end{array} \right]$$

En notación compacta, esta matriz se puede denotar por la siguiente expresión:

(49) 
$$H = \begin{bmatrix} H_{11} & H_{12} \\ H_{21} & H_{22} \end{bmatrix}$$

El algoritmo que propone Isermann (1985) es el siguiente:

SI  $\text{Det}[H_{22}] \gg 0$  ENTONCES Estimar  
 SI  $\text{Det}[H_{22}] \approx 0$  ENTONCES No Estimar

Para implementar el condicional, será preciso seleccionar un valor umbral del determinante de  $H$ , por encima del cual se procede a estimar. En principio, no existe mucha información a priori para seleccionar este umbral.

Las correlaciones, se pueden calcular recursivamente mediante la fórmula siguiente:

(50) 
$$\Phi_{uy}(t, k) = \Phi_{uy}(t, k-1) + \frac{1}{k+1} [u(k-t)y(k) - \Phi_{uy}(t, k-1)]$$

La misma fórmula, se puede aplicar para el cálculo de  $\Phi_{uu}$  o de  $\Phi_{yy}$ ; la expresión (50) tiene memoria infinita y por tanto no es adecuada para sistemas variantes. En este caso es preciso sustituir el factor  $1/(k+1)$  por otro fijo  $G$ , con lo que la fórmula se reduce a un filtrado de primer orden de la correlación correspondiente:

$$(51) \quad \Phi_{uy}(t, k) = \frac{G}{1 - (1 - G)z^{-1}} u(k - t)y(k)$$

La posición del polo  $(1 - G)$  en el plano  $z$ , establece un compromiso entre la capacidad del filtro para eliminar el efecto del ruido que afecta a las señales y su rapidez para seguir variaciones del proceso, que se reflejen en estas correlaciones.

El método resulta muy costoso computacionalmente, obteniéndose resultados similares si se emplea solamente la diagonal de  $H(t)$ , o sea,  $\Phi_{uu}(0)$  y  $\Phi_{yy}(0)$ , o lo que es lo mismo, las varianzas de  $u(t)$  e  $y(t)$ . Además, resulta mucho más sencillo seleccionar umbrales para estas varianzas, que para el determinante de  $H(t)$ .

Los métodos de conexión/desconexión presentan la ventaja de la sencillez de su ajuste. Tan sólo es preciso, seleccionar las varianzas de  $u(t)$  e  $y(t)$  que se tomarán como umbrales, por encima de las cuales se procede a la activación del estimador.

Como ejemplo de la efectividad de esta técnica para prevenir el "blow-up", se utiliza como parte de un controlador autoajustable de tiempo finito, para controlar adaptativamente el siguiente sistema:

$$(52) \quad (1 - 0.9z^{-1})y(t) = 0.1z^{-1}u(t) + v(t)$$

Los cambios de la variable de consigna del lazo de control,  $r(t)$ , se aprecian claramente en la figura 3. Son 0 en todo momento, excepto en el tramo 30-50 del tiempo de simulación, reflejado en el eje de abcisas. Los valores de la desviación típica del ruido,  $v(t)$ , son los siguientes:

$$60 \leq t \leq 80 \quad \text{Desv. Tip. } v(t) = 0.05$$

$$80 \leq t \leq 110 \quad \text{Desv. Tip. } v(t) = 0.5$$

Para el resto de valores de  $t$ , el ruido  $v(t)$  es nulo. El factor de olvido exponencial fijo es de 0.94 y se utiliza la varianza de la señal de control,  $u(t)$  para analizar la excitación a que está sometido el lazo, calculándose ésta con un filtro de tiempo finito de cuarto orden con todos los coeficientes iguales y de ganancia unidad. Este filtrado equivale a aplicar directamente la fórmula de la varianza sobre los cinco últimos datos.

En la figura 3, puede apreciarse la conexión del estimador con las excitaciones del lazo de control, bien sean debidas a cambios de la consigna o a perturbaciones.

Obsérvese, como para el umbral elegido las perturbaciones de desviación típica 0.05 no provocan la activación del estimador.

En principio, el método resuelve el problema del “blow-up” y también trata, hasta cierto punto, la diferencia de excitación de los componentes del vector de observaciones empleando dos umbrales, uno para los elementos asociados a  $u(t)$  y otro para los asociados a  $y(t)$ . No obstante, es preferible el empleo de métodos que aprovechen automáticamente cualquier excitación en las señales utilizadas por el estimador, sea esta grande o pequeña, sin necesidad de traspasar umbrales predefinidos por el usuario.

### 3.6 Métodos de Olvido Exponencial-Direccional

La información que aporta al estimador cada nuevo conjunto de observaciones, normalmente, no se distribuye por igual entre todos los parámetros. Por tanto, se puede decir que en cada paso, la información le entra al estimador en una determinada “dirección”, dentro del espacio de parámetros. Además, la información casi nunca está uniformemente distribuida a lo largo del tiempo. El objetivo de los métodos de olvido exponencial direccional, es precisamente *olvidar en la misma “dirección” en la que entre la información*. Recuérdese que los factores de olvido escalares olvidan la información por igual en todas las direcciones.

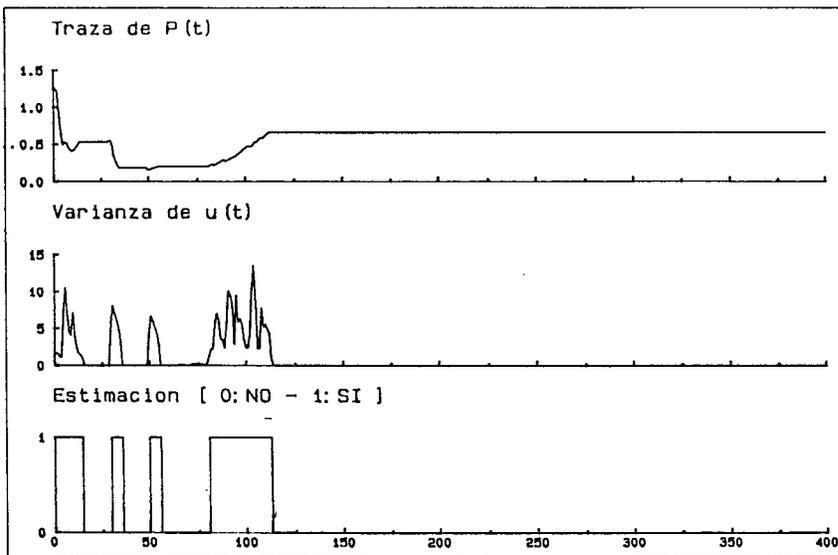
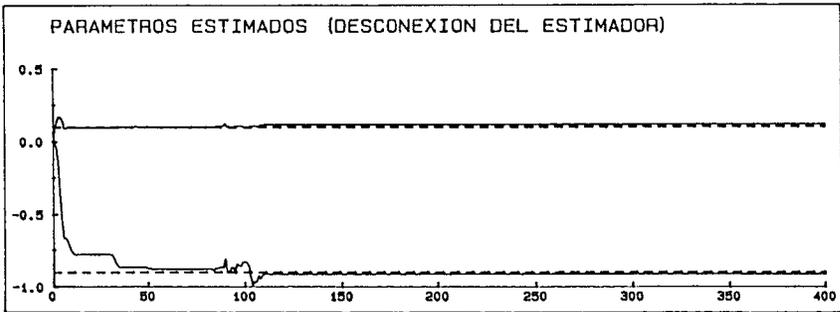
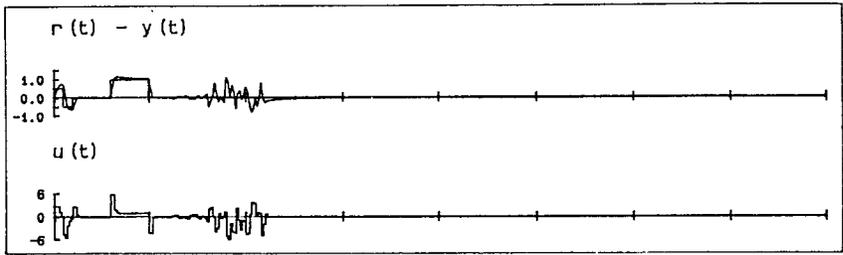


Figura 3. Análisis del método de conexión/desconexión del estimador para prevenir el “blow-up”

Seguidamente, se expone para este tipo de métodos, el más difundido actualmente que es el algoritmo de olvido direccional propuesto por *Kulhavy y Karny* (1984).

La idea básica, consiste en descontar los datos antiguos en la misma dirección en la que entra la nueva información, en el espacio de parámetros. El algoritmo de los mínimos cuadrados se convierte en el siguiente:

$$\begin{aligned}
 (53) \quad & e(t) = y(t) - z(t)^T \theta(t-1) \\
 (54) \quad & K(t) = \frac{P(t-1)z(t)}{1 + z(t)^T P(t-1)z(t)} \\
 (55) \quad & \theta(t) = \theta(t-1) + K(t)e(t) \\
 (56) \quad & \text{SI } \|z(t)\| = 0 \text{ ENTONCES} \\
 (57) \quad & \mu(t) = 1 \\
 & \text{SINO} \\
 (58) \quad & \mu(t) = \beta - (1 - \beta) [z(t)^T P(t-1)z(t)]^{-1} \\
 & \text{FIN SI} \\
 (59) \quad & H(t) = \frac{P(t-1)z(t)}{\mu(t)^{-1} + z(t)^T P(t-1)z(t)} \\
 (60) \quad & P(t) = [I - H(t)z(t)^T] P(t-1)
 \end{aligned}$$

*Hgglund* (1983), demostró que si  $0 < \beta \leq 1$ , entonces  $P(t)$  se mantiene definida positiva para todo  $t$ , suponiendo que  $P(0)$  lo sea. La constante  $\beta$ , se puede interpretar como la cantidad de olvido “exponencial” en la dirección de la información entrante.  $\|z(t)\|$  representa la norma euclídea del vector  $z(t)$  y se define como:

$$(61) \quad \|z(t)\| = (z(t)^T z(t))^{1/2}$$

Empleando la condición anterior, el algoritmo de *Kulhavy y Karny* (1984) no elimina el problema del “blow-up” de la matriz de covarianzas y por tanto, su utilización no es segura en un entorno industrial. O sea, si se toma el cero máquina = 0, ante ausencia de excitación la traza queda limitada en infinito; lógicamente esto no es práctico, mientras que usando la condición:

$$(62) \quad \|z(t)\| < M$$

el límite de  $P(t)$ , queda determinado por los valores de  $\beta$  y  $M$ . Por ejemplo, con  $M = 1E - 7$  se realizó un experimento de estimación empleando el mecanismo

de olvido propuesto por Kulhavy y Karny (1984), bajo las mismas condiciones que el experimento de la figura 3. Los resultados obtenidos de esta experiencia se muestran en la figura 4, en la que se puede observar que en ausencia de excitación la traza sube hasta un valor de 7017.42, con  $\beta = 0.9$ . Repitiendo el experimento, con un valor de  $\beta = 0.95$  la traza se estabiliza en un valor de 13,15. No se aporta la gráfica de la traza de  $P(t)$  para esta segunda experiencia, ya que es totalmente similar a la mostrada en la figura 4, pero con un valor final diferente. Kulhavy y Karny (1984) recomiendan implementar la condición

$$\|z(t)\| = 0 \quad \text{como} \quad [z(t)^T P(t-1)z(t)] < \text{cero} - \text{máquina}$$

pero esto tampoco resuelve el problema del “blow-up”.

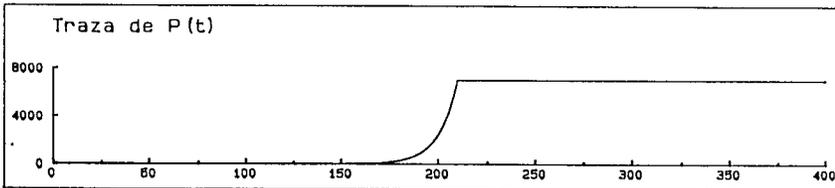
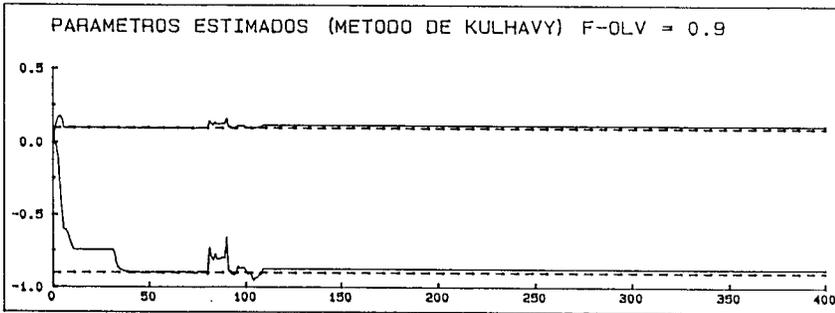
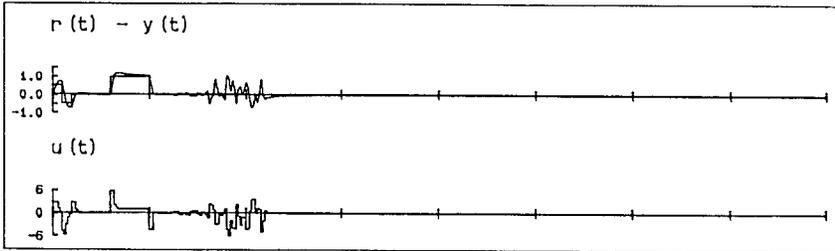


Figura 4. Problemas planteados por el mecanismo de olvido direccional de Kulhavy

### 3.7 Métodos de Olvido Exponencial-Direccional-Variable

Estos métodos, son refinamientos de los algoritmos de olvido direccional presentados anteriormente para controlar el ratio de olvido. Con ellos se pretende conseguir una respuesta más instantánea de los estimadores, ante variaciones de los parámetros de la planta.

Hgglund(1983), ha propuesto mantener constante cierta cantidad de información deseada, si los parámetros son constantes. Como medida de información, emplea las varianzas de los parámetros; los elementos de la diagonal de  $P(t)$  pueden ser interpretados, de forma aproximada, como las varianzas de los parámetros correspondientes. Por tanto, el algoritmo trata de llevar a la matriz  $P(t)$  a un valor deseado  $P_0(t)$ . Es aconsejable elegir  $P_0(t) = A \cdot I$ , así el método requiere una estimación del nivel de ruido al que está sometido el sistema, para que la matriz  $P(t)$  pueda interpretarse como las covarianzas de los parámetros estimados. A continuación, se muestran las ecuaciones de este estimador:

$$(63) \quad w(t) = \frac{1 - z^{-1}}{1 - \tau_1 z^{-1}} \theta(t - 1) \quad \theta \leq \tau_1 < 1$$

$$(64) \quad s(t) = \text{SIGN} [\text{INT} [\theta(t - 1)^r] * w(t)]$$

$$(65) \quad r(t) = \frac{1 - \tau_2}{1 - \tau_2 z^{-1}} s(t) \quad 0 \leq \tau_2 < 1$$

$$(66) \quad \text{SI } r(t) < r_1 \text{ ENTONCES}$$

$$(67) \quad v(t) = \frac{1 - \tau_3}{1 - \tau_3 z^{-1}} e(t - \Gamma)^2$$

SINO

$$(68) \quad v(t) = v(t - 1)$$

FINSI

$$(69) \quad \delta_d(t) = \frac{1}{z(t)^T P(t - 1)^2 z(t)} \left[ \frac{z(t)^T P(t - 1)^3 z(t)}{z(t)^T P(t - 1)^2 z(t)} - A \right]$$

$$(70) \quad \delta_d(t) = v(t)^{-1} + \frac{\delta_d(t)}{\delta_d(t) z(t)^T P(t - 1) z(t) - 1}$$

$$(71)$$

(71)

$$\alpha(t) = \begin{cases} 0 & \alpha_d \leq 0 \\ \alpha & 0 \leq \alpha_d \leq \frac{1}{z^T P z} \\ \frac{1}{z^T P z} & \frac{1}{z^T P z} < \alpha_d \leq v^{-1} + \frac{1}{z^T P z} \\ 0 & \alpha_d > v^{-1} + \frac{1}{z^T P z} \end{cases}$$

$$(72) \quad e(t) = y(t) - z(t)^T \theta(t-1)$$

$$(73) \quad \theta(t) = \theta(t-1) + \frac{P(t-1)z(t)e(t)}{v(t) + [1 - \alpha(t)v(t)]z(t)^T P(t-1)z(t)}$$

$$(74) \quad P(t) = P(t-1) - \frac{P(t-1)z(t)z(t)^T P(t-1)}{[v(t)^{-1} - \alpha(t)]^{-1} + z(t)^T P(t-1)z(t)}$$

En la fórmula (64), INT es la función parte entera, y SIGN es la función signo, que devuelve -1 si su argumento es negativo, 1 si es positivo y 0 si el argumento es nulo.  $\tau_1, \tau_2$  y  $\tau_3$  son los polos de tres filtros de primer orden utilizados por el método. En la ecuación (67),  $\Gamma$  indica el número de retardos a considerar en el error de predicción a utilizar, para evaluar a partir de él, el nivel del ruido del sistema, definido por  $v(t)$ . La constante A de la ecuación (69) representa el valor que deben alcanzar los elementos de la diagonal de la matriz de covarianzas. Los elementos fuera de la diagonal deben converger a cero.

Este método, es uno de los más complejos de los que se dispone actualmente. Probablemente, debido a ello no se ha empleado en aplicaciones industriales (en lo que los autores conocen). Si se desea obtener un estimador industrial, lo más sencillo y eficaz posible, es preferible emplear otros métodos más simples. Por otra parte, se desconoce su sensibilidad ante errores de modelado, perturbaciones, etc.

*Bertin, Bittanti y Bolzarn* (1986), han combinado el algoritmo de Fortescue *et al.*(1981) con el de Kulhavy *et al.*(1984) para obtener un factor de olvido variable y direccional. El algoritmo, viene dado por las siguientes ecuaciones:

$$(75) \quad e(t) = y(t) - z(t)^T \theta(t-1)$$

$$(76) \quad K(t) = \frac{P(t-1)z(t)}{1 + z(t)^T P(t-1)z(t)}$$

$$(77) \quad \theta(t) = \theta(t-1) + K(t)e(t)$$

$$(78) \quad \alpha(t) = 1 - \frac{1 - z(t)^T K(t)}{\Sigma_0} e(t)^2$$

$$(79) \quad \beta(t) = \text{MAX} \{ \beta_{\min}; \beta(t) \}$$

$$(80) \quad \text{SI } |z(t)| = 0 \text{ ENTONCES}$$

$$(81) \quad \mu(t) = 1$$

SINO

$$(82) \quad \mu(t) = \beta(t) - [1 - \beta(t)] [z(t)^T P(t-1)z(t)]^{-1}$$

FIN - SI

$$(83) \quad H(t) = \frac{P(t-1)z(t)}{\mu(t)^{-1} + z(t)^T P(t-1)z(t)}$$

$$(84) \quad P(t) = [I - H(t)z(t)^T] P(t-1)$$

Donde  $\beta_{\min}$  representa el valor mínimo, que puede tomar en cualquier instante el factor de olvido direccional  $\beta(t)$ .

El factor de olvido, depende del error de predicción  $e(t)$ ; en determinadas circunstancias esto puede producir un comportamiento extraño de dicho factor de olvido. Bertin, Bittanti y Bolzern (1986), proponen dos modificaciones de la ecuación de obtención de  $\beta(t)$  para mejorar el comportamiento del algoritmo. La primera, consiste en ponderar el error de estimación:

$$(85) \quad \alpha(t) = 1 - \frac{1 - z(t)^T K(t)}{\Sigma_0} \left\{ \sum_{i=1}^m e(t-i+1)^2 / m \right\}$$

La otra ley de adaptación, de tipo heurístico, es la siguiente:

$$(86) \quad \alpha(t) = 1 - [1 - z(t)^T K(t)] \frac{e(t)^2}{\Sigma_0 + e(t)^2}$$

En la fórmula (86), se verifica que  $0 < \alpha(t) \leq 1$ , para todo  $t$  y la sensibilidad de  $\alpha(t)$  con respecto a  $e(t)^2$ , es reducida.

El método de Bertin *et al.*(1986), comparte los problemas de los métodos de Fortescue *et al.*(1981) y de Kulhavy *et al.*(1984). Bertin *et al.*(1986) indican que se pueden obtener comportamientos extraños del algoritmo en función del error de predicción. Por lo tanto, si se desea obtener un estimador industrial fiable, esto es inaceptable.

Kulhavy (1987), desarrolló un mecanismo para modificar el parámetro  $\beta$  del algoritmo direccional presentado anteriormente por él mismo, obteniendo el denominado algoritmo de olvido direccional restringido; su objetivo, es lograr una mayor adaptatividad del estimador:

$$(87) \quad e(t) = y(t) - z(t)^T \theta(t-1)$$

$$(88) \quad \zeta(t) = z(t)^T P(t-1)z(t)$$

$$(89) \quad K(t) = \frac{P(t-1)z(t)}{1 + \zeta(t)}$$

$$(90) \quad \theta(t) = \theta(t-1) + K(t)e(t)$$

$$(91) \quad e_n(t) = \frac{e(t)^2}{\sigma^2 [1 + \zeta(t)]}$$

$$(92) \quad \beta(t)^{-1} = 1 + (1 + \Omega) \left[ \ln(1 + \zeta(t)) + \frac{\zeta(t)}{1 + \zeta(t)} (e_n(t) - 1) \right]$$

$$(93) \quad \text{SI } \zeta(t) = 0 \text{ ENTONCES}$$

$$(94) \quad \mu(t) = 1$$

SINO

$$(95) \quad \mu(t) = \beta(t) - [1 - \beta(t)] \zeta(t)^{-1}$$

FIN - SI

$$(96) \quad H(t) = \frac{P(t-1)z(t)}{\mu(t)^{-1} + \zeta(t)}$$

$$(97) \quad P(t) = [I - H(t)z(t)^T] P(t-1)$$

Si se prevé que los parámetros desconocidos no varíen muy rápidamente, valores de  $\Omega < 0,5$ , son suficientes para un adecuado seguimiento de los mismos. En este caso, el valor en régimen estacionario de  $\zeta(t)$ , es lo suficientemente

pequeño para aproximar el logaritmo natural del modo siguiente:

$$(98) \quad \ln [1 + \zeta(t)] \approx \frac{\zeta(t)}{1 + [\zeta(t)/2]}$$

Usando esta aproximación, Kulhavy (1987) demostró que el valor esperado de  $\beta(t)$  en régimen estacionario cumple la siguiente relación:

$$(99) \quad \beta(t)^{-1} = 1 + 2\Omega$$

Se puede emplear esta última relación para seleccionar con facilidad el valor de  $\Omega$ . Kulhavy (1987), indica que este método de cálculo del factor de olvido no se puede considerar como universalmente aceptable. No obstante, las ideas aportadas en algunos algoritmos previos pueden introducirse fácilmente en este esquema. Por ejemplo, el algoritmo propuesto por Fortescue *et al.*(1981) se puede implementar dentro de este esquema del siguiente modo:

$$(100) \quad \beta(t)^{-1} = 1 + \frac{e_N(t)^2}{\Sigma_0}$$

Obsérvese que la fórmula (100) no es, numéricamente, idéntica a la (101) propuesta por Fortescue *et al.*(1981),

$$(101) \quad \beta(t) = 1 - \frac{e_N(t)^2}{\Sigma_0}$$

sino que tan sólo es equivalente a ella, cualitativa o conceptualmente. Seguidamente, se comparan los algoritmos de Fortescue *et al.*(1981) y Kulhavy (1987). Para ello se emplea la fórmula de Kulhavy (1987) una vez aproximado el logaritmo y tomando  $\zeta/2 \approx \zeta$ , comparados ambos con la unidad:

$$(102) \quad \beta(t)^{-1} = 1 + (1 + \Omega) \frac{\zeta(t)}{1 + \zeta(t)} e_n(t)^2$$

Relacionando la fórmula (102) con la (103):

$$(103) \quad \beta(t)^{-1} = 1 + \frac{e_n(t)^2}{\Sigma_0}$$

se deduce que:

$$(104) \quad \Sigma_0 \approx \frac{1 + \zeta(t)^{-1}}{1 + \Omega}$$

Por tanto, es obvio que en el algoritmo de Fortescue *et al.*(1981) la longitud de la memoria del estimador es fijada a priori por el usuario, mientras que con el algoritmo de Kulhavy (1987), es dependiente de la incertidumbre de los parámetros estimados a través de  $\zeta(t)$ . Esto implica un mayor grado de adaptatividad del algoritmo.

El método de olvido direccional restringido de Kulhavy (1987), incorpora parámetros, como  $\Omega$ , que son de muy difícil ajuste. Por otra parte, este algoritmo es bastante complejo y no ha sido experimentado industrialmente; en principio, no resulta aconsejable para construir un estimador adaptativo industrial de tipo general.

### 3.8 Otras Técnicas

*Anderson*(1985), ha propuesto un método de identificación recursiva basado en el empleo en paralelo de varios algoritmos de mínimos cuadrados recursivos (o filtros de Kalman). Además, emplea otro algoritmo para el cálculo de probabilidades a posteriori y determinada lógica para saltar de un estimador a otro. El método es bastante complejo, y las mejoras que aporta con respecto a otros no son considerables, siendo el coste computacional muy elevado. Por lo tanto, no será considerado como un método ampliamente aplicable en algoritmos adaptativos de amplia difusión en la industria.

## 4. NUEVOS MÉTODOS PARA LA ESTIMACIÓN ADAPTATIVA DE SISTEMAS VARIANTES

Para superar los problemas que presentan los métodos analizados anteriormente se han desarrollado dos nuevos algoritmos, que verifican los requisitos necesarios para su utilización en aplicaciones reales: una seguridad total en su

funcionamiento, fácilmente configurables por un usuario de nivel medio, presentando una gran capacidad de adaptación a los cambios del sistema y siendo computacionalmente eficientes.

El primero de ellos, se basa en el *principio de mantener constante la información almacenada en el estimador*. Para ello, se mantienen a un nivel constante los elementos de la diagonal de  $P(t)$ , o lo que es lo mismo, la traza de la matriz de covarianzas.

Para lograr este último objetivo, se utiliza un vector de factores de olvido variables; así, cuando falta excitación en alguno de los componentes del vector de parámetros el factor de olvido correspondiente tiende a uno, y cuando es excesiva, el factor de olvido alcanza y se mantiene en un valor mínimo prefijado por el usuario. De este modo, se evita que los parámetros fluctúen excesivamente con las señales de excitación. El mecanismo, se puede implementar del siguiente modo:

1. Actualizar  $P(t)$  sin olvido
2. PARA  $i := 1$  HASTA  $\text{Dim}[P(t)]$  HACER
  - SI  $P_{ii}(t) \geq P(i, i)_0$  ENTONCES
    - $\beta_i(t) := 1.0$
  - SINO
    - $\beta_i(t) := \text{MAX} \{P(i, i)/P(i, i)_0; \beta_{\min}\}$
  - FIN - SI
3. Dividir  $P_{ij}(t)$  por  $\beta_i \cdot \beta_j$

Para realizar un análisis comparativo de este método con el de Saelid y Foss (1983), se estiman los parámetros del proceso variante siguiente:

$$(106) \quad (1 + a_1 z^{-1}) y(t) = b_1 z^{-1} u(t) + v(t)$$

$$\begin{array}{lll} 0 & \leq t < 149 & a_1 = -0.8 \quad b_1 = 1.0 \\ 150 & \leq t < 299 & a_1 = -0.4 \quad b_1 = 1.0 \\ 300 & \leq t < 500 & a_1 = -0.4 \quad b_1 = 0.3 \end{array}$$

cuando es excitado con una señal  $u(t)$  de tipo PRBS de amplitud variable y una perturbación estocástica  $v(t)$  de desviación típica 0.02. En la figura 5, se muestran estas señales, así como los parámetros estimados, los elementos de la diagonal de la matriz de covarianzas y los factores de olvido cuando se utiliza el

método de Saelid y Foss (1983). En la figura 6, se muestran las mismas cuando se utiliza el nuevo método.

En la figura 5, se puede observar como el método de Saelid y Foss (1983) elimina el problema del “blow-up”, ya que lleva los factores de olvido a la unidad en ausencia total de excitación ( $50 \leq t < 99$ ), pero “congela” al estimador tras el período de excitación elevada ( $100 \leq t < 140$ ).

El nuevo método, propuesto por García (1988), ante períodos de excitación elevada, trabaja con factores de olvido reducidos y no es capaz de mantener la traza requerida, pero ante la falta de excitación, evita el “blow-up”, limitando los valores que puede tomar  $P(t)$ . Obsérvese, como en el intervalo  $100 \leq t < 300$ , la excitación es lo suficientemente elevada en todo el vector de observaciones, como para olvidar siempre el máximo permitido. En este caso, los elementos de  $P(t)$  se mantienen por debajo del nivel de consigna. A partir de  $t = 300$ , la excitación a que está sometido el elemento  $b_1$  asociado a  $u(t - 1)$ , sigue siendo lo suficientemente elevada para que el factor de olvido correspondiente continúe siendo mínimo, mientras que el elemento  $a_1$  asociado a  $y(t - 1)$ , se encuentra menos excitado, adecuándose el factor de olvido correspondiente al mismo, para que el elemento asociado de  $P(t)$ , no supere el valor límite especificado.

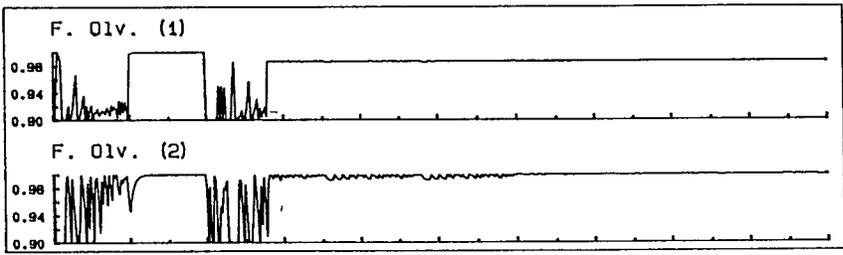
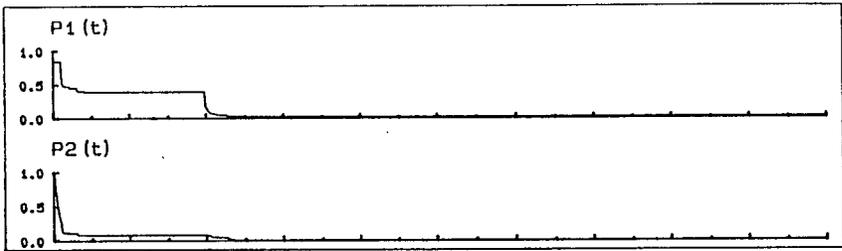
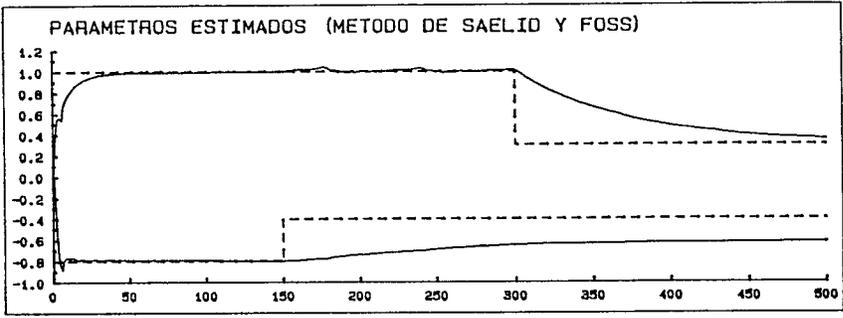
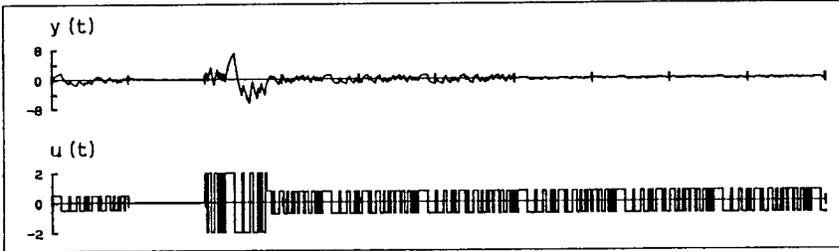


Figura 5: Análisis del mecanismo de olvido de información propuesto por Saelid y Foss.

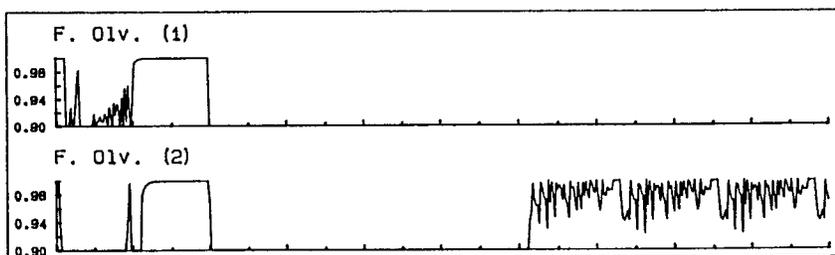
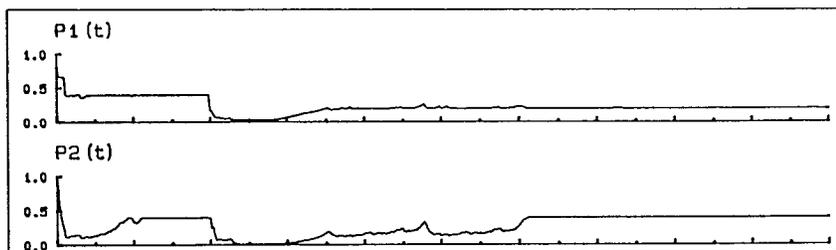
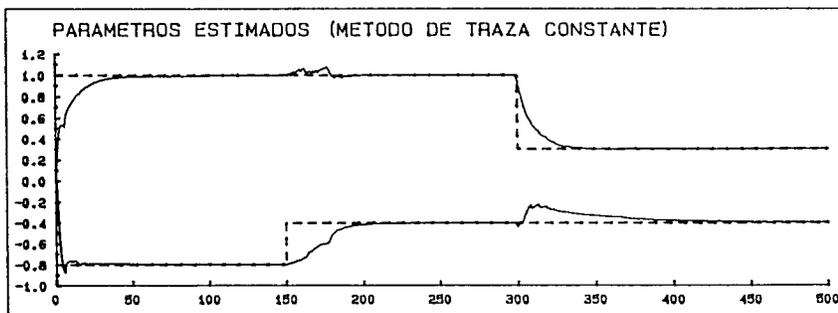
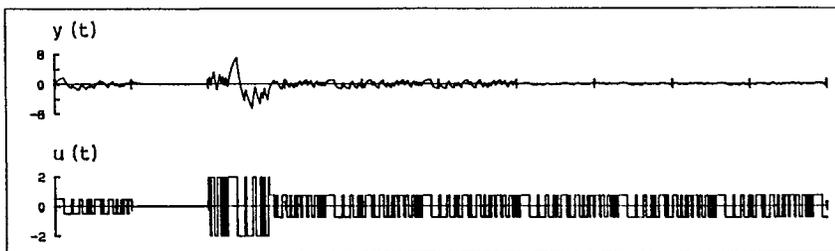


Figura 6: Análisis del mecanismo de olvido de información basado en mantener la traza de  $P(t)$  constante.

El experimento anterior, ha mostrado claramente como esta técnica resuelve el problema del “blow-up” y el de la falta de uniformidad de la excitación en el espacio de parámetros. Por tanto, el mecanismo de olvido desarrollado es adecuado para ser incorporado en un estimador adaptativo industrial. El ajuste del algoritmo de olvido, resulta muy sencillo, ya que tan sólo es preciso seleccionar dos parámetros:

- El valor deseado de los elementos de la diagonal de  $P(t)$
- El factor de olvido mínimo

Cuanto mayor es la traza, mejor es el seguimiento de los parámetros, pero el estimador es más sensible al ruido. Es preciso establecer un compromiso entre estos dos requisitos contrapuestos y el factor de olvido mínimo, determina la máxima capacidad de seguimiento del estimador.

El otro método propuesto, consiste en la *utilización de una matriz  $Q(t)$  que se añade directamente a la matriz de covarianzas*, con objeto de mantener la traza constante a un nivel  $\text{Tr}_0$  seleccionado por el usuario. No obstante, para evitar que el estimador tienda a olvidar demasiada información durante etapas de elevada excitación del proceso, es preciso limitar el valor que puede adquirir la matriz  $Q(t)$  a un valor  $Q_{\max}$  seleccionado por el usuario. Debido a esta limitación los elementos de la matriz de covarianzas disminuyen, pero recuperan el valor de referencia cuando se reduce la excitación. El algoritmo es el siguiente:

$$(107) \quad \begin{aligned} & 1) \text{ Actualizar } P(t) \text{ sin olvidar información alguna} \\ & 2) Q(t) = \text{MAX} \{K(t)z(t)^T P(t-1); Q_{\max}\} \\ & 3) P_{ii}(t) = \text{MIN} \{P_{ii}(t) + Q(t); \text{Tr}_0\} \end{aligned}$$

Recuérdese que si  $Q(t) = K(t)z(t)^T P(t-1)$ , la diagonal de la matriz  $P(t)$  permanece inalterable.

Otra variante de este último método, consiste en emplear un valor de  $Q(t)$  fijo, denominada  $Q$ . En este caso,  $Q$  representa el incremento fijo que se le da en cada paso a cada uno de los elementos de la diagonal de la matriz de covarianzas.

Además, es preciso comprobar que los elementos de la diagonal no crezcan por encima de un determinado valor seleccionado por el usuario. El algoritmo se puede resumir del modo siguiente:

$$(108) \quad \begin{aligned} & 1) \text{ Actualizar } P(t) \text{ sin olvidar información alguna} \\ & 2) P_{ii}(t) = \text{MIN} \{P_{ii}(t) + Q; \text{Tr}_0\} \end{aligned}$$

En aplicaciones reales, puede ocurrir que los elementos de la matriz de covarianzas tengan magnitudes muy dispares. En este caso, puede ser contraproducente el requerir que todos los elementos de la diagonal de  $P$  tiendan al mismo valor; entonces, es posible definir el vector al que se desea que tienda la diagonal de  $P(t)$ . No obstante, un enfoque más práctico suele ser el fijar un valor para los elementos de la diagonal de  $P$  asociados a  $u(t)$ , otro para los asociados a la perturbación estocástica y si es preciso, otro para los asociados al error. Esto complica ligeramente la configuración del estimador.

Estos algoritmos de modificación directa de  $P(t)$ , funcionan correctamente y son fiables, ya que todos incorporan explícitamente el control de la matriz de covarianzas. En realidad, la utilización de los métodos de adición de  $Q(t)$ , bien sea fija o variable limitada a un máximo, no son más que una nueva implementación del principio incorporado en el método de traza constante, presentado anteriormente.

Para ilustrar el funcionamiento de este mecanismo, se repite el experimento realizado anteriormente para comparar los métodos de Saelid y Foss (1983) y el de traza constante. Los resultados se muestran en la figura 7; el comportamiento de los parámetros estimados es muy bueno y el efecto de “blow-up” completamente eliminado. El método de traza constante es sencillo de usar, pues sólo es preciso seleccionar dos parámetros; el valor deseado para la diagonal de  $P(t)$  y la matriz  $Q(t)$  o su valor máximo.

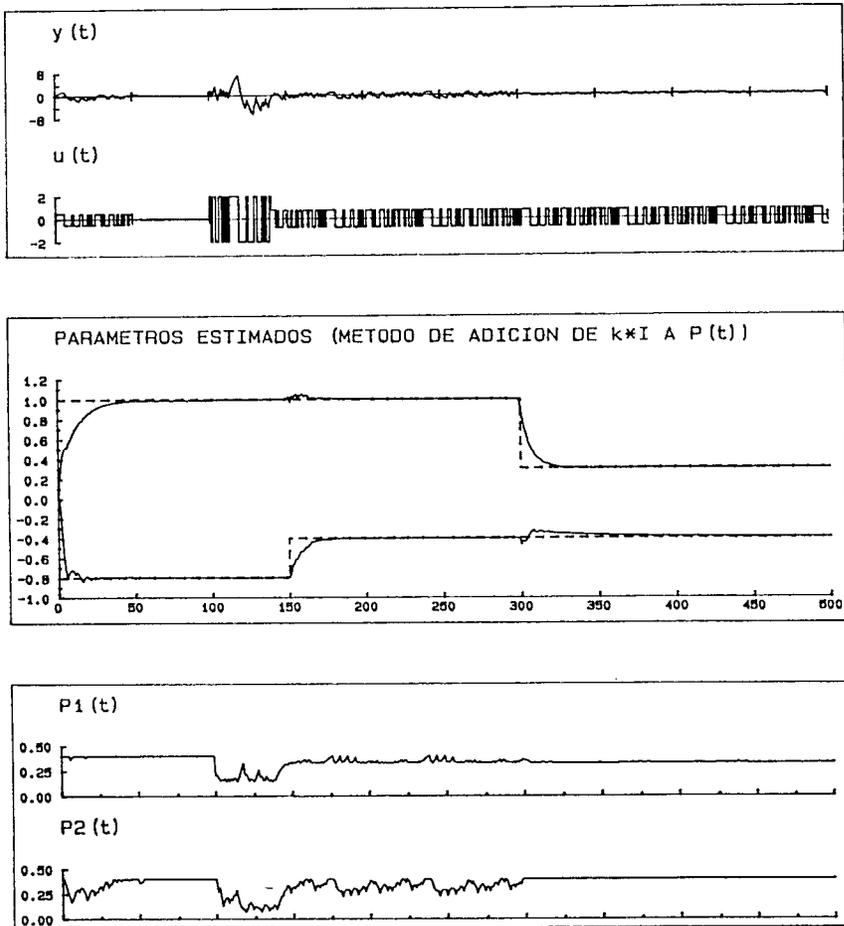


Figura 7: Análisis del mecanismo de olvido basado en la adición de una matriz constante a  $P(t)$ .

## 5. CONCLUSIONES

En el artículo, se han definido los requisitos que debe cumplir un estimador adaptativo para que pueda ser utilizado en aplicaciones reales. También se ha realizado un análisis, de los problemas fundamentales que se plantean en un proceso de estimación adaptativa de parámetros variantes.

En base a los análisis anteriores, se ha realizado un detallado estudio de los métodos de estimación adaptativa disponibles en la actualidad. En función de este estudio los autores se inclinan por la utilización de factores de olvido exponenciales vectoriales, modificados adecuadamente (nuevo método) para permitir una recuperación sin problemas, después de etapas de excitación muy elevada. También se recomienda, la utilización de los métodos basados en la modificación directa de la matriz de covarianzas, aunque éstos se emplearán preferentemente en situaciones de reinicialización del estimador.

Cuando la excitación es elevada, estos mecanismos de olvido permiten que la matriz de covarianzas, tome valores reducidos actuando el estimador mínimo-cuadrático como tal, presentando su típica "elevada rapidez de convergencia". Si la excitación es reducida, estos mecanismos controlan la matriz de covarianzas, comportándose entonces el estimador mínimo-cuadrático como uno de tipo gradiente, con menor velocidad de convergencia pero sin presentar problemas de blow-up.

## 6. BIBLIOGRAFÍA

- [1] **Anderson, P.** (1985) "Adaptative forgetting in recursive identification through multiple models". *INT. J. CONTROL*, Vol.42, No.5, pp.1175-1193.
- [2] **Astrom, K.J.** (1980) "Design principles for self-tuning regulators". En *Methods and Applications in Adaptive Control*, SPRINGER-VERLAG, Mayo-80.
- [3] **Bergmann, S.** (1982) "Some improvements for the application of digital parameter-adaptative control algorithms", IEE Conference on Applications of Adaptive and Multivariable Control, Hull, GB.
- [4] **Bertin, D., Bittanti, S y Bolzern, P.** (1986) "Tracking of non-stationary systems by means of different prediction-error directional forgetting techniques", 2ND-IFAC Workshop on Adaptive Systems in Control and Signal Processing, Lund, Sweden.

- [5] **Clarke, D. y Gawthrop, P.J.** (1979) "Self-tuning control". IEE PROCEEDINGS, Vol.126, No.6, pp.633-640.
- [6] **Cordero, A.O. y Mayne, D.Q.** (1981) "Deterministic convergence of a self-tuning regulator with variable forgetting factor". IEE PROCEEDINGS Vol.128, Pt. D, No.1.
- [7] **Fortescue, T.R., Kershenbaum, L. y Ydstie, B.E.** (1981) "Implementation of self-tuning regulators with variable forgetting factors". AUTOMATICA, Vol.17, No.6, pp.831-835.
- [8] **García, D.F.** (1988) "Control adaptativo de procesos monovariantes y multivariantes. Evaluación y diseño de controladores autoajustables de aplicación industrial generalizada". Tesis. Universidad de Oviedo.
- [9] **Goodwin, G.C., Elliot, H. y Teoh, E.K.** (1983) "Deterministic convergence of a self-tuning regulator with covariance resetting". IEE PROCEEDINGS, Vol.130, Pt. D, No.1.
- [10] **Goodwin, G.C., Hill, D.J. y Palaniswami, M.** (1985) "Towards an Adaptive robust controller". 7th IFAC/IFORS Symposium on Identification and System Parameter Estimation, Vol.1, pp.997-1002.
- [11] **Hagglund, T.** (1983) "New estimation techniques for adaptive control". Report 1025, Lund Institute of Technology. Thesis.
- [12] **Irving, E.** (1979) "New developments in improving power control network stability with adaptive generator control". En APPLICATIONS OF ADAPTATIVE CONTROL, ACADEMIC PRESS.
- [13] **Isermann, R.** (1985) "Parameter-adaptive control systems, a review on methods and applications". 7th IFAC/IFORS Symposium on Identification and System Parameter Estimation, York, UK, pp.373-378.
- [14] **Kulhavy, R. y Karny, M.** (1984) "Tracking of slowly varying parameters by directional forgetting". 9th-IFAC World Congress, Budapest, Hungary.
- [15] **Kulhavy, R.** (1987) "Restricted exponential forgetting in real time identification". Submitted to AUTOMATICA for publication.
- [16] **Saelid, S. y Foss, B.** (1983) "Adaptive controllers with a vector variable forgetting factors". 22ND Conference on Decision and Control, San Antonio, Texas, pp.1488-1494.
- [17] **Saelid, S., Egeland, O. y Foss, B.** (1985) "A solution to the blow-up problem in adaptive controllers". MODELLING, IDENTIFICATION AND CONTROL, VOL.6, No.1, pp.39-56.
- [18] **Schumann, R., Lachmann, K.H. y Isermann, R.** (1981) "Towards applicability of parameter-adaptive control algorithms". 8th IFAC World Congress, Kyoto, Japan.
- [19] **Tuffs, P.S.** (1985) "Self-tuning control: algorithms and applications". Oxford University, OUEL Report No.1567/85, D. Phil Thesis.

- 
- [20] **Wellstead, P.E. y Sanoff, S.P.** (1981) "Extended self-tuning algorithm". INT J. CONTROL, Vol.34, No.3, pp.433-455.
  - [21] **Ydstie, B.E. y Co, T.** (1985) "Recursive estimation with adaptive divergence control". IEE PROCEEDINGS, Vol.132, Pt.D, No.3.
  - [22] **Young, P.** (1984) "Recursive estimation and time-series analysis". SPRINGER-VERLAG.

