

ESTIMACIÓ DE PARÀMETRES DE SISTEMES DINÀMICS  
UTILITZANT TÈCNiques DE FILTRATGE

P. BRUNET

El control de sistemes és tant més acurat com més precís és el model matemàtic del sistema controlat. En molts casos, això només es pot aconseguir amb un algorisme d'estimació en temps real dels seus paràmetres; i quan el sistema no és determinista degut a l'entrada de sorolls, s'han d'utilitzar tècniques de filtratge. Fins ara, s'han utilitzat molt sovint filtres subòptims, que no tenen assegurada la convergència. Concretament, l'aplicació de l'equació de Fokker Plank a les equacions diferencials de les estimacions permet concloure una no observabilitat dels paràmetres per part d'aquests filtres quan s'apliquen a una classe de sistemes quasilinears. Ara bé, els paràmetres poden ésser estimats correctament si es parteix del model discret del sistema. En aquest cas es pot utilitzar una modificació del filtre Extended Kalman, el teorema de partició o el filtrat òptim discret. A l'article es presenten els tres algorismes aplicats a un exemple d'un sistema lineal de segon ordre amb un paràmetre desconegut, i se'n discuteixen els avantatges de cada un d'ells.

1. INTRODUCCIÓ

L'estimació dels paràmetres d'un sistema del qual se'n coneix l'estructura és un aspecte molt important de la teoria de control, ja que l'estimació en temps real permet anar coneixent els paràmetres encara que varien lentament; així, el model del procés es va refinant de forma contínua.

Deixant de banda el cas de sistemes sense soroll a la dinàmica i únicament amb soroll a l'observació, que ja han sigut estudiats ampliament /1/, /2/, a aquest article ens concretarem als sistemes estocàstics que ho són sigui perquè l'entrada no és accessible i no més se'n coneixen les seves propietats estadístiques, sigui perquè a més de l'entrada el sistema rep sorolls.

En aquest cas, el model dinàmic continu del sistema si suposem que no depèn del temps serà de la forma:

$$dx = f(x,a)dt + G(x,a)d\beta \quad (1)$$

on

$x = (x_1 \dots x_n)^T$  és el vector estat, de dimensió  $n$ , i

$a = (a_1 \dots a_m)^T$  és el vector de  $m$  paràmetres a estimar;

el primer terme és la part determinista del model, i el segon inclou el soroll, modelitzat com un moviment brownià vectorial  $\beta_t$ , -- /3/, amb matriu de covariances  $\Omega$ :

$$E[\dot{\beta} \dot{\beta}^T] = \Omega dt \quad (2)$$

Els algorismes de filtratge ens donen una estimació de l'estat  $x$  a partir de una observació  $z(t)$  pertorbada per soroll i modelitzada normalment com

$$\left. \begin{aligned} dz &= h(x) dt + d\eta \\ \text{amb} & \\ E[\dot{\eta} \dot{\eta}^T] &= Rdt \end{aligned} \right\} \quad (3)$$

Si es prefereix, en canvi, un model discret del sistema, aquest relaciona l'estat en instants successius  $t_k$  i  $t_{k+1}$ ; el soroll ara es modelitza com soroll blanc vectorial gaussià  $u_k, v_k$ :

Model dinàmic

$$x_{k+1} = F(x_k, a) + G(x_k, a)u_k; E[u_k u_k^T] = Q \quad (4)$$

- P. Brunet del Laboratori de Càlcul, ETSEIB, i del Institut de Cibernètica, UPB. Diagonal, 647. Barcelona 28.  
- Article rebut l'Agost de 1978.

Model d'observació

$$z_k = H(x_k) + V_k; E[V_k V_k^T] = R \quad (4)$$

Tant en el cas continu com en el discret, - si es vol que els algorismes de filtratge estimin els paràmetres  $a_1 \dots a_m$ , aquests s'han d'incloure al vector estat, i així arriben als models dinàmics

$$\left. \begin{aligned} dy &= \left\{ \frac{dx}{da} \right\} = \left\{ \begin{matrix} f(x, a) \\ 0 \end{matrix} \right\} dt + \left\{ \begin{matrix} G(x, a) \\ 0 \end{matrix} \right\} d\beta \\ y_{k+1} &= \left\{ \begin{matrix} x_{k+1} \\ a_{k+1} \end{matrix} \right\} = \left\{ \begin{matrix} F(x_k, a_k) \\ a_k \end{matrix} \right\} + \left\{ \begin{matrix} G(x_k, a_k) \\ 0 \end{matrix} \right\} u_k \end{aligned} \right\} (5)$$

on  $y = (x_1 \dots x_n \ a_1 \dots a_m)^T$ .

Encara que (1), (4) fossin models lineals en l'estat, (5) no ho és en general degut als productes entre paràmetres i variables  $x_k$ . Per tant, ens trobem davant d'un cas de filtratge no lineal, i no es pot aplicar el filtre òptim de Kalman Bucy.

En molts casos s'han utilitzat filtres subòptims, que s'obtenen truncant el desenvolupament en sèrie de Taylor de  $f(x, a)$ ,  $F(x_k, a)$ ,  $G$  i  $H$ , i menyspreant els moments d'ordre superior de la densitat de probabilitat de les estimacions. Aquests filtres no tenen garantida la seva convergència, i a més, diferents simulacions, /4/, havien mostrat divergència en les estimacions als casos concrets estudiats.

A l'apartat següent veurem com l'aplicació de l'Equació de Fokker Planck permet deduir analíticament aquest mateix resultat. Als apartats posteriors, s'exposen no obstant tres mètodes que convergeixen a l'estimació correcta, tot partint del model discret del sistema.

2. ESTIMACIÓ DE PARÀMETRES PER FILTRATGE CONTINU SUBÒPTIM

Suposem un sistema continu (1) i un model d'observació (3), tots dos lineals en l'estat:

$$\left. \begin{aligned} dx &= F(a) x dt + G(a) d\beta; E[d\beta d\beta^T] = Qdt \\ dz &= H(a) x dt + d\eta; E[d\eta d\eta^T] = Rdt \end{aligned} \right\} (6)$$

si els paràmetres es coneixen i, per tant, - només es vol estimar l'estat  $x=(x_1 \dots x_n)^T$ , - el filtre de Kalman és el que ens dona l'estimació  $\hat{x}(t)$  òptima en el sentit de minimitzar l'experança matemàtica de  $\|\hat{x}-x\|^2$  conegudes les observacions. El filtre de Kalman ve donat, /3/, per:

$$\left. \begin{aligned} d\hat{x} &= F\hat{x} dt + PHR^{-1}(dz - H\hat{x} dt) \\ dP &= (FP + PF^T + GQG^T - PH^TR^{-1}HP) dt \end{aligned} \right\} (7)$$

(l'evolució temporal de l'estimació  $\hat{x}$  té dos termes; l'un és la part determinista de l'equació del sistema, i l'altre hi introdueix una correcció proporcional al terme d'innovació o diferència entre l'observació real  $dz$  i la prevista,  $H\hat{x} dt$ ). Aquestes equacions -- ens donen l'evolució de la mitjana  $\hat{x}$  i de la matriu de covariàncies  $P$  de la densitat de probabilitat de l'estat  $x$  conegudes les observacions  $\{Z(z)/z \leq t\}$ . Es pot demostrar que aquesta densitat de probabilitat és gaussiana, i per tant les equacions (7) permeten conèixer totalment la seva evolució temporal; en cada moment l'estimació  $\hat{x}$  de  $x$  és la seva mitjana.

En els casos no lineals, com per exemple el model (5) continu, l'anterior densitat de probabilitat ja no és gaussiana, i el sistema (7) que dona el filtre òptim es converteix en un sistema d'infinites equacions diferencials acoblades per l'evolució dels moments de la densitat de probabilitat, impossible d'integrar. A partir d'aquí i mitjançant hipòtesis simplificatives que no tenen suport teòric, s'han deduït els filtres subòptims, dels quals aquí tractarem l'Extended Kalman, el Gaussià i el Truncat, /3/. Tots ells, -- aplicats al sistema (5) amb observació lineal  $dz=Hx dt + d\eta$  es poden expressar en la forma:

$$\left. \begin{aligned} d\hat{x} &= [f(\hat{x}, \hat{a}) + \frac{1}{2}c(P\delta^2 f)] dt + P_x H^T R^{-1} (dz - H\hat{x} dt) \\ d\hat{a} &= P_{ax} H^T R^{-1} (dz - H\hat{x} dt) \\ dP_x &= [ \hat{F}_x P_x + P_x \hat{F}_x^T + \hat{F}_a P_{ax} + P_{ax}^T \hat{F}_a^T + \hat{G}Q\hat{G}^T + cA - \\ &\quad - P_x H^T R^{-1} H P_x ] dt \\ dP_{xa} &= [ \hat{F}_x P_{xa} + \hat{F}_a P_{ax} - P_x H^T R^{-1} H P_{xa} ] dt \\ dP_a &= -P_{ax} H^T R^{-1} H P_{ax} dt \end{aligned} \right\} (8)$$

on  $c=0$  a l'Extended Kalman i  $c=1$  als altres.

A més,

$\hat{x}$  és l'esperança matemàtica de  $x$  coneixent  $z(\tau)$ ,  $\tau \leq t$  que dóna el filtre

$\hat{a}$  és l'esperança matemàtica de  $a$  coneixent  $z(\tau)$ ,  $\tau \leq t$

$$P_x = E\{(x-\hat{x})(x-\hat{x})^T / z(\tau), \tau \leq t\}$$

$$P_{xa} = P_{ax}^T = E\{(x-\hat{x})(a-\hat{a})^T / z(\tau), \tau \leq t\}$$

$$P_a = E\{(a-\hat{a})(a-\hat{a})^T / z(\tau), \tau \leq t\}$$

$F_x$  és la matriu jacobiana  $(F_x)_{ij} = \frac{\delta f_i}{\delta x_j}$

$F_a$  és la matriu de  $n \times m$   $(F_a)_{ij} = \frac{\delta f_i}{\delta a_j}$

$\hat{F}_x$ ,  $\hat{F}_a$ ,  $\hat{G}$  són les matrius  $F_x$ ,  $F_a$ ,  $G$  amb  $x=\hat{x}$ ,  $a=\hat{a}$

$(P\delta^2 f)$  és un vector amb

$$(P\delta^2 f)_k = \text{tr} \left\{ P_x^T \frac{\delta^2 f_k}{\delta x_i \delta x_j} + 2P_{xa}^T \frac{\delta^2 f_k}{\delta x_i \delta a_j} + P_a^T \frac{\delta^2 f_k}{\delta a_i \delta a_j} \right\}_{x=\hat{x}, a=\hat{a}}$$

$$(A)_{ij} = \sum_{k_1 k_2} G_{ik_1} Q_{k_1 k_2} \sum_{k_3 k_4} P_{k_3 k_4} \left[ \frac{\delta^2 G_{jk_2}}{\delta y_{k_3} \delta y_{k_4}} \right]_{x=\hat{x}, a=\hat{a}} +$$

$$+ \sum_{k_1 k_2} Q_{k_1 k_2} \sum_{k_3 k_4} P_{k_3 k_4} \left[ \frac{\delta G_{ik_1}}{\delta y_{k_3}} \frac{\delta G_{jk_2}}{\delta y_{k_4}} \right]_{x=\hat{x}, a=\hat{a}} +$$

$$+ \gamma \frac{1}{4} \sum_{k_1 k_2 k_3 k_4 k_5 k_6} Q_{k_1 k_2} (P_{k_3 k_4} P_{k_5 k_6} +$$

$$+ P_{k_3 k_5} P_{k_4 k_6} + P_{k_3 k_6} P_{k_4 k_5}) \left[ \frac{\delta^2 G_{ik_1}}{\delta y_{k_3} \delta y_{k_4}} \right]_{x=\hat{x}, a=\hat{a}}$$

$$\left. \frac{\delta^2 G_{ik_2}}{\delta y_{k_5} \delta y_{k_6}} \right]_{x=\hat{x}, a=\hat{a}}$$

i  $\gamma=0$  al filtre truncat;  $\gamma=1$  al Gaussià.

Com es pot veure, el sistema (8) d'equacions diferencials dels filtres només dóna l'evolució dels dos primers moments degut a les hipòtesis simplificatives. Ara per tal d'estudiar la convergència o no de l'estimació  $\hat{a}$  al valor correcte  $a$ , és convenient estudiar el conjunt sistema-observació-filtre, equacions (3), (5) i (8). L'aplicació de l'equació de Fokker-Planck a aquest sistema conjunt, /5/, permet arribar, sota les hipòtesis

$$- \text{rang}(H) = n \text{ i donat } T, \hat{F}_a x = 0 \Rightarrow x = 0 \quad \forall t \geq T$$

-  $f_i(x, a)$  depèn linealment de  $x_i$  i com a màxim del cub de  $x_k$ ,  $k \neq i$ .

-  $G(x, a)$  és únicament funció dels paràmetres  $G(a)$

i després de certes manipulacions, a:

$$\frac{\delta p}{\delta t} = -p \left\{ (m+n+1) \text{tr} \left[ \hat{F}_x - H^T R^{-1} H P_x \right] + \frac{1}{2} c \cdot \right.$$

$$\left. \sum_{k=1}^n \text{tr} \left[ P_x^T \left( \frac{\delta^3 f_k}{\delta x_i \delta x_j \delta x_k} \right) \right] \right\}_{x=v, a=\alpha} - \sum_{i=1}^n \sum_{j=1}^n \frac{\delta p}{\delta \pi_{ij}}$$

$$\left[ \hat{F}_x P_x + P_x \hat{F}_x^T + \hat{G} Q \hat{G}^T - P_x H^T R^{-1} H P_x \right]_{ij}$$

equació que ens dóna l'evolució de la densitat de probabilitat  $p$  de les estimacions donades pel filtre:

$$p(\alpha_1 \dots \alpha_m \pi_{11} \dots \pi_{nn}) d\alpha_1 \dots d\alpha_m d\pi_{11} \dots d\pi_{nn} d\pi_{11}^1 \dots$$

$$\dots d\pi_{nn}^m d\pi_{11}^{11} \dots d\pi_{nn}^{mm} = \text{Prob} \left\{ \hat{a}_i \in \{\alpha_i, \alpha_i + d\alpha_i\}, \right.$$

$$P_x)_{ij} \in \{\pi_{ij}, \pi_{ij} + d\pi_{ij}\}, P_a)_{ij} \in \{0, d\pi_{ij}^j\},$$

$$P_a)_{ij} \in \{0, d^{ij}\} \left. \right\}$$

Ara bé, es pot veure, /5/, que l'equació (9) implica la no convergència de  $\hat{a}$  als paràmetres correctes a ja que aquests no hi apareixen. Per tant, sota les hipòtesis vistes anteriorment, els filtres subòptims estudiats són inadequats per l'estimació de paràmetres, i hem de concloure que cal utilitzar algoris

mes diferents d'aquests.

### 3. UTILITZACIÓ DEL FILTRE EXTENDED KALMAN MODIFICAT

Fins ara hem estudiat els algorismes de filtratge aplicats a sistemes amb model continu; si en canvi volem estimar els paràmetres de sistemes modelitzats de forma discreta (4), una de les possibilitats és aplicar el filtre Extended Kalman discret; si a més, el sistema a estimar és lineal

$$\left. \begin{aligned} x_{k+1} &= F(a) x_k + G(a) u_k \\ z_k &= H x_k + v_k \end{aligned} \right\} \quad (9)$$

el filtre ve donat per:

$$\left. \begin{aligned} \hat{x}_{k+1} &= F(a) \hat{x}_k + K_k (z_k - H \hat{x}_k) \\ \hat{a}_{k+1} &= \hat{a}_k + L_k (z_k - H \hat{x}_k) \\ K_k &= \left[ F P_{1k} H^T + M_k P_{2k}^T H^T \right] S_k^{-1} \\ M_k &= \left[ \begin{array}{c} \delta \\ -\delta a \end{array} F(a) \hat{x}_k \right]_{a=\hat{a}} \\ S_k &= H P_{1k} H^T + R \\ L_k &= P_{2k}^T H^T S_k^{-1} \\ P_{1k+1} &= F P_{1k} F^T + F P_{2k} M_k^T + M_k P_{2k}^T F^T + M_k P_{3k} M_k^T - \\ &\quad - K_k S_k K_k^T + Q \\ P_{2k+1} &= F P_{2k} + M_k P_{3k} - K_k S_k L_k^T \\ P_{3k+1} &= P_{3k} - L_k S_k L_k^T \end{aligned} \right\} \quad (10)$$

on la matriu de covariàncies del vector

$$y = (x_1 \dots x_n a_1 \dots a_m)^T$$

és

$$\begin{bmatrix} P_1 & P_2 \\ P_2^T & P_3 \end{bmatrix}$$

L'algorisme (10), igualment com la seva versió contínua (8), tampoc convergeix a l'estimació correcta dels paràmetres, com ha demostrat Ljung /6/. De tota manera, i en el cas que estem tractant d'estimació de paràmetres

de sistemes lineals, Ljung demostra que modificant únicament la definició de la matriu  $M_k$  a les equacions (10), es pot assegurar la convergència amb probabilitat 1 de l'algorisme al valor correcte dels paràmetres. En aquest algorisme Extended Kalman modificat, es fa dependre  $M_k$  del terme d'innovació:

$$M_k = \left[ \begin{array}{c} \delta \\ \delta a \end{array} F(a) \hat{x}_k \right]_{a=\hat{a}} + \beta_k (z_k - H \hat{x}_k) \quad (11)$$

i el tensor  $\beta_k$  es pot anar calculant a cada pas de forma recurrent:

$$\beta_k = \left[ \begin{array}{c} \delta \\ \delta a \end{array} F(a) P_{1k} H^T + F \pi_k H^T \right]_{a=\hat{a}} - K_k S_k^{-1} \sigma_k$$

$$\sigma_k = H \pi_k H^T \quad (12)$$

$$\begin{aligned} \pi_{k+1} &= \left[ \begin{array}{c} \delta \\ \delta a \end{array} F \right]_{a=\hat{a}} P_{1k} F^T(\hat{a}) + F(\hat{a}) \pi_k F^T(\hat{a}) + \\ &\quad + F(\hat{a}) P_{1k} \left[ \begin{array}{c} \delta \\ \delta a \end{array} F^T(a) \right]_{a=\hat{a}} + \\ &\quad + \left[ \begin{array}{c} \delta \\ \delta a \end{array} \{G O G^T\} \right]_{a=\hat{a}} - \beta_k S_k K_k^T - \\ &\quad - K_k \sigma_k K_k^T - K_k S_k \beta_k^T \end{aligned}$$

Encara que, com ho hem dit, la convergència queda assegurada, podem veure que molt sovint les equacions de recurrència pels tensors (12) poden arribar a consumir gran part del temps de càlcul.

Per tal de poder veure un exemple d'aplicació d'aquest algorisme, així com per tenir la possibilitat de comparar-lo amb els que veurem als apartats següents, s'ha plantejat l'estimació del paràmetre  $a$  (coeficient d'esmortiment) a un sistema oscil·lant de dimensió 2 pertorbat per soroll.

$$\begin{Bmatrix} x_{k+1} \\ y_{k+1} \end{Bmatrix} = \begin{Bmatrix} x_k \\ y_k \end{Bmatrix} + \begin{bmatrix} 0 & 1 \\ -1/2 & -2a \end{bmatrix} \begin{Bmatrix} x_k \\ y_k \end{Bmatrix} + \begin{Bmatrix} u_{k1} \\ u_{k2} \end{Bmatrix} \quad (13)$$

del qual n'observem solament la variable  $x_k$ :

$$z_k = x_k + v_k \quad (14)$$

Es suposa que els tres sorolls blancs  $u_{k1}$ ,  $u_{k2}$  i  $v_k$  són mútuament independents i amb variances respectivament  $q_1$ ,  $q_2$ ,  $r$ .

En aquest cas, l'algorisme (10)-(12) queda més simplificat pel fet que només hi ha un paràmetre a estimar, pel fet que la matriu  $G$  és la identitat i que  $H = [1, 0]$ . Això porta a fer que no es tingui de treballar amb cap tensor, sinó que  $\pi_k$  i  $P_{1k}$  són matrius de  $2 \times 2$ ;  $K_k$ ,  $P_{2k}$ ,  $M_k$  i  $\beta_k$  són vectors de dimensió 2; i la resta són escalars. En concret, les equacions del filtre modificat són:

$$\begin{Bmatrix} \hat{x}_{k+1} \\ \hat{y}_{k+1} \end{Bmatrix} = F \begin{Bmatrix} \hat{x}_k \\ \hat{y}_k \end{Bmatrix} + K_k (z_k - \hat{x}_k);$$

$$F = \begin{bmatrix} 1 & 1 \\ -1/2 & 1-2\hat{a}_k \end{bmatrix}$$

$$\hat{a}_{k+1} = \hat{a}_k + L_k (z_k - \hat{x}_k) \quad (15)$$

$$K_k = (FP_{1k} + M_k P_{2k}^T) H^T \frac{1}{S_k}$$

$$S_k = p_{11} + r$$

$$L_k = p_{13} (p_{11} + r)$$

$$P_{1k+1} = FP_{1k} F^T + FP_{2k} M_k^T + M_k P_{2k}^T F^T + M_k P_{3k} M_k^T - K_k S_k K_k^T + Q$$

$$P_{2k+1} = FP_{2k} + M_k P_{3k} - K_k S_k L_k$$

$$P_{3k+1} = P_{3k} - L_k^2 S_k$$

$$M_k = \begin{Bmatrix} 0 \\ -2\hat{y}_k \end{Bmatrix} + \beta_k (z_k - \hat{x}_k)$$

$$\beta_k = \frac{1}{p_{11} + r} \begin{bmatrix} 0 \\ -2p_{12} \end{bmatrix} + F \pi_k H^T \begin{bmatrix} 1 \\ -K_k \sigma_k S_k \end{bmatrix}$$

$$\sigma_k = H \pi_k H^T$$

$$\pi_{k+1} = \begin{bmatrix} 0 & 0 \\ 0 & -2 \end{bmatrix} P_{1k} F^T + F \pi_k F^T - \beta_k S_k K_k^T - K_k S_k \beta_k^T - K_k \sigma_k K_k^T$$

on  $p_{ij}$  són les components de la matriu de co

QÜESTIÓ - v.3, nº1 (març 1979)

variances

$$\begin{bmatrix} p_1 & p_2 \\ p_2^T & p_3 \end{bmatrix}, \quad i \quad Q = \begin{bmatrix} q_1 & 0 \\ 0 & q_2 \end{bmatrix}$$

La simulació de l'algorisme anterior s'ha implementat amb els següents passos:

1) Inicialitzacions

$$x_0 = y_0 = 0$$

$$q_1 = q_2 = r = .01$$

$$a = 0.3$$

$$\hat{x}_0 = \hat{y}_0 = \hat{a}_0 = 0$$

2)  $k = 0$

3) Simulació del sistema, (13)-(14) utilitzant variables pseudoaleatòries gaussianes. Obtenició de  $z_k$ ,  $x_{k+1}$ ,  $y_{k+1}$ .

4) Simulació del filtre, (15). Obtenició de les estimacions i resta de variables a l'instant  $k+1$ . (Cal tenir en compte que les equacions de  $K_k$ ,  $M_k$  i  $\beta_k$  s'han de resoldre simultàniament).

5)  $k = k+1$

6) Tornar a 3) si  $(\hat{a}-a)^2$  és massa gran.

Els resultats de la simulació es poden observar a les figures 1, 2 i 3. A la figura 1, es té l'evolució real del sistema i de la seva observació  $z_k$ ; a les figures 2 i 3, les estimacions  $\hat{a}_k$  sense introduir la modificació al filtre Extended Kalman ( $\beta_k \equiv 0$ ) i introduint-les respectivament.

#### 4. FILTRATGE ÒPTIM DISCRET

Quan es tracta amb sistemes modelitzats en forma discreta, i encara que no siguin lineals, eq. (4), és possible, almenys teòricament, obtenir l'estimació òptima per aplicació recursiva de l'equació de Chapman Kolmogorof per processos markowians i del teorema de Bayes, /7/. D'aquesta forma, es va obtenir l'evolució temporal de la densitat de probabilitat de les variables a estimar con-

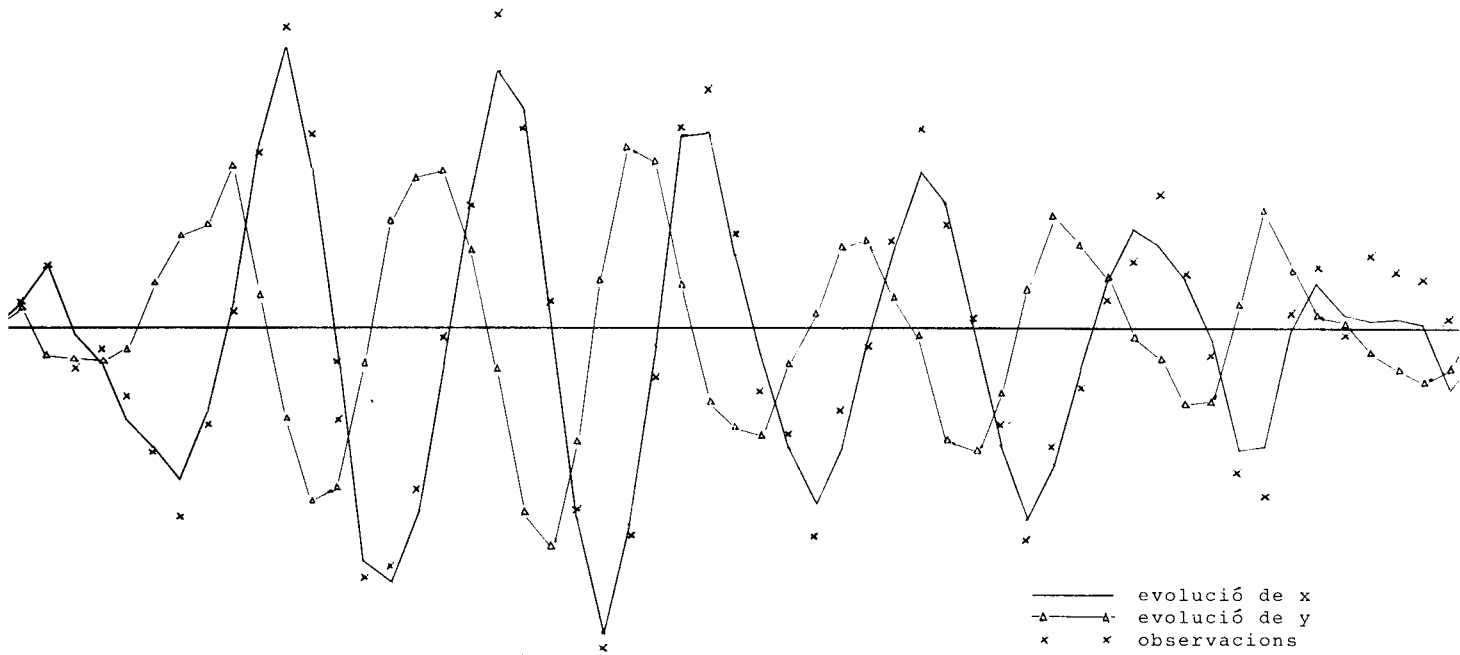


Fig. 1.  
Evolució temporal de les variables d'estat del sistema i de la observació

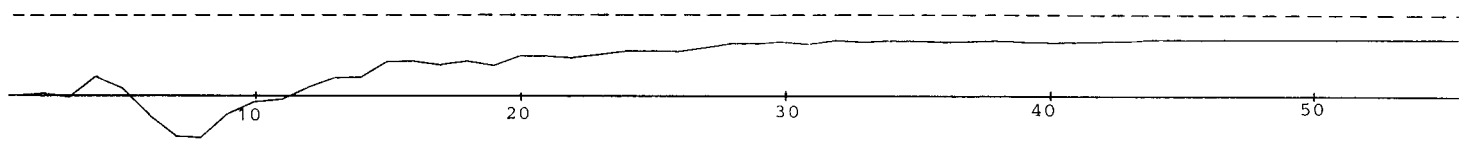


Fig. 2.  
Filtre Extended Kalman ( $\beta_k \equiv 0$ ). Evolució de  $\hat{a}_k$

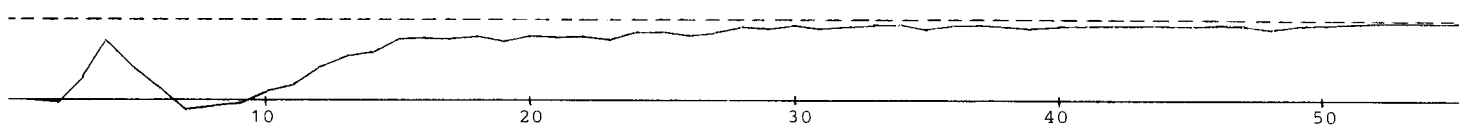


Fig. 3.  
Filtre Extended Kalman modificat. Evolució de  $\hat{a}_k$

dicionada per les observacions; i a partir - d'aquesta, com ja ho hem vist, és immediat - trobar les estimacions i les seves covariances, que no són més que els seus moments.

En concret, suposem coneguda la densitat de probabilitat condicionada pel conjunt d'observacions fins l'instant  $t_k$ ,

$$\text{Prob} \{x_k = x_1, a = a_1 / \{z_0 \dots z_k\}\}$$

podem predir  $x$  i  $a$  a l'instant  $t_{k+1}$  utilitzant només l'equació de la dinàmica:

$$\text{Prob} \{x_{k+1} = x_2, a = a_1 / \{z_0 \dots z_k\}\} = \int \dots \int \text{Prob} \{x_k = x_1, a = a_1 / \{z_0 \dots z_k\}\} \cdot$$

$$\cdot \text{Prob} \left\{ \begin{matrix} x_{k+1} = x_2 \\ a = a_1 \end{matrix} \middle/ \begin{matrix} x_k = x_1 \\ a = a_1 \end{matrix} \right\} dx_1$$

on la

$$\text{Prob} \{x_{k+1} = x_2 / x_k = x_1\} = \text{Prob} \{G(x_1, a_1) u_k = x_2 - F(x_1, a_1)\} \quad (16)$$

depèn de les característiques del soroll. Ara, a partir d'aquesta densitat de probabilitat, podem calcular la que obtindrem en incloure la nova observació  $z_{k+1}$  que ha arribat:

$$\text{Prob} \{x_{k+1} = x_2, a = a_1 / \{z_0 \dots z_{k+1}\}\} = \frac{\text{Prob} \{z_{k+1} / x_{k+1} = x_2, a = a_1\} \cdot \text{Prob} \{x_{k+1} = x_2, a = a_1 / \{z_0 \dots z_k\}\}}{\text{Prob} \{z_{k+1} / \{z_0 \dots z_k\}\}} \quad (17)$$

com podem veure, incloure la informació d'una nova observació no és més que multiplicar la densitat de probabilitat de la predicció per la

$$\text{Prob} \{z_{k+1} / x_{k+1} = x_2, a = a_1\} = \text{Prob} \{v_{k+1} = z_{k+1} - H(x_2)\},$$

i normalitzar-la (ja que el denominador és una constant respecte a les variables  $x_{k+1}$ ,  $a$ ). En resum, l'aplicació alternada de (16) i (17) permet d'anar obtenint l'evolució de la densitat de probabilitat, i per tant de les estimacions.

Les dificultats principals que apareixen en

l'aplicació de l'anterior algorisme provenen, d'una banda de l'emmagatzemament de la densitat de probabilitat, i d'altre banda, de la complexitat inherent a l'equació (16), que requereix una integral per trobar cada un dels punts de la nova densitat de probabilitat. Aquests aspectes seran discutits més àmpliament a l'apartat 6.

Concretament, en l'aplicació que s'ha fet del present algorisme a l'estimació del paràmetre de l'exemple de l'equació (13)-(14), s'ha emmagatzemat la densitat de probabilitat a cada un dels punts d'una malla cúbica de dimensió  $10 \times 10 \times 10$ . No obstant, hi han mètodes més refinats per guardar-la, com pot ser treballar amb una aproximació per splines, /8/, amb lo qual calen menys punts.

En el cas particular de l'exemple que estem veient, l'equació de Chapman Kolmogorof queda:

$$\text{Prob} [\bar{x}_{k+1} = \xi, y_{k+1} = \eta, a_{k+1} = \zeta / z_0 \dots z_k] = \int_{\tau} \int_{\omega} \text{Prob} [\bar{x}_k = \tau, y_k = \omega, a_k = \zeta] \cdot \text{EXP} \left[ -\frac{1}{2} \frac{(\xi - \tau - \omega)^2}{q_1} - \frac{1}{2} \frac{(\eta - \alpha \tau - 2 \zeta \omega)}{q_2} \right] d\tau d\omega \quad (18)$$

I el teorema de Bayes,

$$\text{Prob} [\bar{x}_{k+1} = \xi, y_{k+1} = \eta, a_{k+1} = \zeta / z_0 \dots z_{k+1}] = K \cdot \text{Prob} [\bar{x}_{k+1} = \xi, y_{k+1} = \eta, a_{k+1} = \zeta / z_0 \dots z_k] \cdot e^{-\frac{1}{2r} (z_{k+1} - \xi)^2} \quad (19)$$

on  $K$  és la constant de normalització.

En aquest cas, la simulació de l'algorisme consta dels següents passos:

1) Inicialitzacions.

$$q_1 = q_2 = r = .01$$

$$x_0 = y_0 = 0$$

$$a = .3$$

Inicialització de la densitat de probabilitat.

- 2)  $k = 0$
- 3) Simulació del sistema (13)-(14) obtenint  $z_{k+1}, x_{k+1}, y_{k+1}$ .
- 4) Aplicació de l'equació (18) per obtenir la predicció a  $t_{k+1}$  coneixent les observacions fins  $t_k$ . (Això és calcular numèricament la integral per cada un dels 1000 punts de la malla).
- 5) Multiplicació de cada un dels punts de la densitat de probabilitat per

$$e^{-\frac{1}{2r}(z_{k+1}-\xi)^2}, \text{ equació (19)}.$$

- 6) Normalització de la densitat de probabilitat de manera que la integral (suma) sigui 1.

- 7) Càlcul de les estimacions com moments de la densitat de probabilitat.
- 8) Tornar a 3) si  $(\hat{a}-a)^2$  és massa gran.

A les figures 4 i 5 es veu l'evolució de la densitat de probabilitat marginal del paràmetre a donada pel filtre. A la figura 4, un dels punts de la malla coincideix amb el valor correcte del paràmetre, .3, i el filtre arriba a convergir totalment al valor correcte. En canvi, a la figura 5 veiem una simulació en què cap dels punts de la malla és el correcte.

### 5. TEOREMA DE PARTICIÓ

El teorema de partició de Lainiotis, /9/, permet, tant en el cas de models continus --

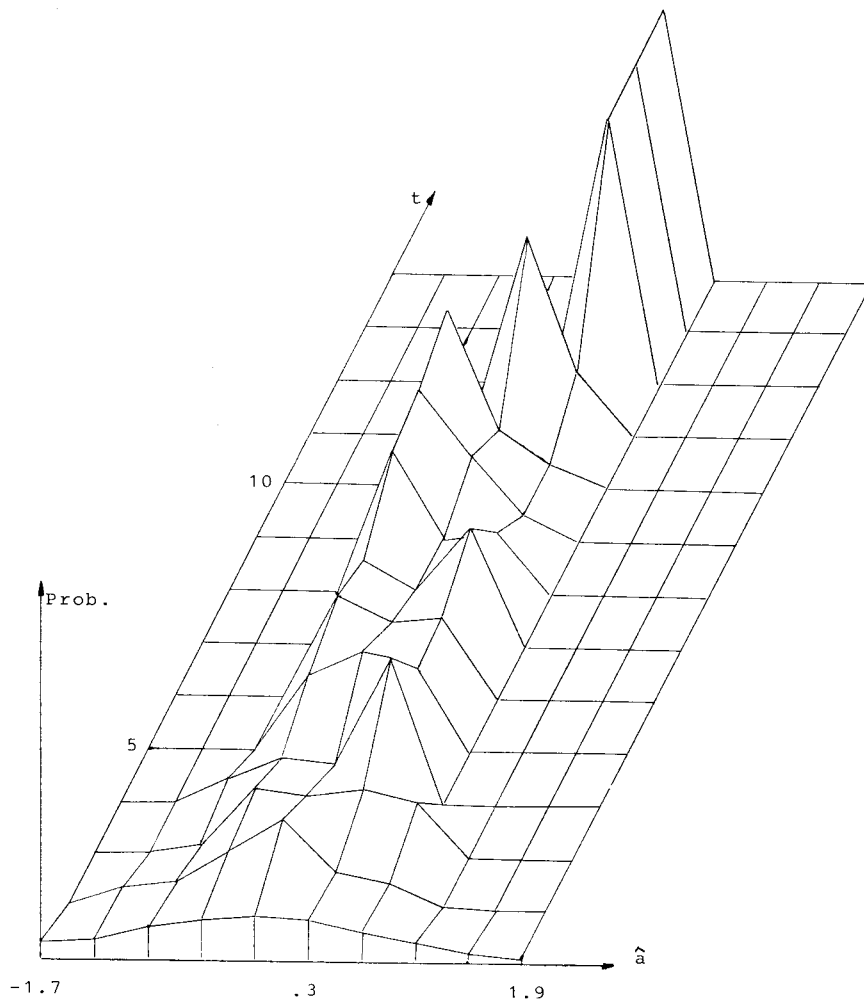


Fig. 4.  
Densitat de probabilitat del paràmetre. Malla des de -1.7 a 1.9



com discrets de sistemes lineals amb paràmetres desconeguts, l'estimació d'aquests. -- Aquí exposarem únicament la versió discreta, per tal de poder comparar amb els algorismes que hem vist anteriorment.

Suposant, per tant, que ens trobem davant -- del problema de l'estimació de paràmetres -- d'un sistema modelat per (10), l'algorisme -- consisteix a elegir N jocs de paràmetres --  $a^1 \dots a^N$  i aplicar el filtre òptim de Kalman als N sistemes lineals descrits pel model -- (10) amb cada un dels conjunts de paràmetres  $a^1$  (és clar que el filtre no es òptim en -- tots els casos ja que el model només és correcte com a màxim en un dels N sistemes). -- Cal tenir en compte que en cas d'estimació -- de l'estat  $x_k$  del sistema (10), el filtre de Kalman ve donat per

$$\hat{x}_{k+1} = F(a^1) \hat{x}_k + N_k (z_k - H\hat{x}_k) \quad (20)$$

$$N_k = F(a^1) P_k H^T (HP_k H^T + R)^{-1}$$

$$P_{k+1} = F(a^1) P_k F(a^1)^T + G(a^1) Q G^T(a^1) - N_k (HP_k H^T + R) N_k^T$$

A partir de les estimacions donades pels N - filters, es poden calcular les relacions de versemblança o probabilitat de tenir l'observació que realment hi ha hagut, sabent que -- es té un joc determinat l de paràmetres:

$$L_1(z_k/a^1) = \frac{1}{\det(P_z)} \exp\left\{-\frac{1}{2}(z_k - H\hat{x}_k^1)^T P_z^{-1} (z_k - H\hat{x}_k^1)\right\} \quad (l=1 \dots N) \quad (21)$$

on

$$P_z = HP_k^1 H^T + R$$

i  $\hat{x}_k^1, P_k^1$  són l'estimació i matriu de covariàncies donades pel filtre l a l'instant  $t_k$ .

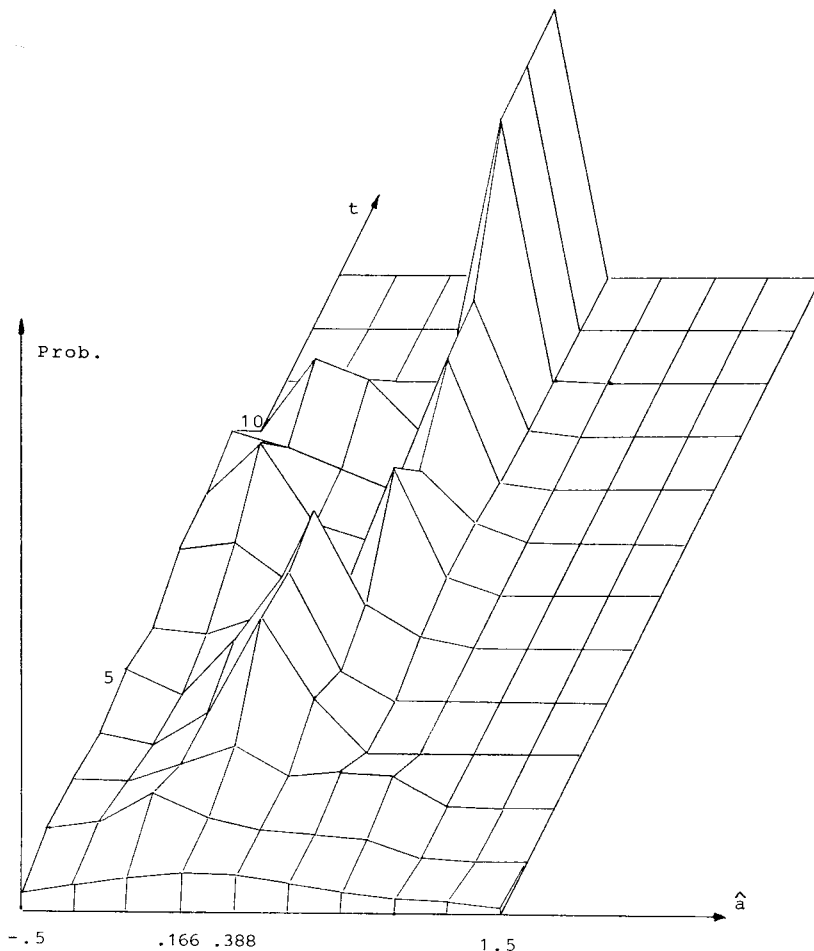


Fig. 5.  
Densitat de probabilitat del paràmetre. Malla des de -0.5 a 1.5

Aquestes relacions de versemblança permeten, mitjançant el teorema de Bayes, anar posant al dia la densitat de probabilitat dels diferents jocs de paràmetres conegudes les observacions:

$$\begin{aligned} \text{Prob} [\bar{a}=a^1/z_0\dots z_k] &= \\ &= k L_1(z_k/a^1) \cdot \text{Prob}[\bar{a}=a^1/z_0\dots z_{k-1}] \end{aligned} \quad (22)$$

Es pot veure que la mitjana en qualsevol instant d'aquesta densitat de probabilitat és l'estimació òptima del joc de paràmetres (i per tant té assegurada la convergència) si el nombre de filtres  $N$  tendeix a infinit. Per tant, la dificultat torna a trobar-se en l'implementació pràctica, on  $N$  ha d'ésser petit.

Com en els casos anteriors, l'algorisme d'estimació s'ha simulat conjuntament amb el sistema, equacions (13)-(14). Aquesta simulació consta dels següents passos:

1) Inicialitzacions.

$$x_0 = y_0 = 0$$

$$a = .3$$

$$q_1 = q_2 = r = .01$$

$$N = 10$$

Inicialització dels 10 valors  $a^1\dots a^{10}$  i de la densitat de probabilitat (plana).

2)  $k = 0$

3) Simulació del sistema (13)-(14) obtenint

$$z_k, x_{k+1}, y_{k+1}$$

4) Càlcul d'un pas d'estimació en cada un dels filtres de Kalman, equacions (20) per  $l = 1\dots 10$ .

5) Càlcul de les relacions de versemblança (21).

6) Càlcul, segons (22), de la nova densitat de probabilitat del paràmetre mitjançant producte per  $L_1$  i normalització.

7) Tornar a 3) si  $(\hat{a}-a)^2$  és gran

A les figures 6 i 7 podem veure l'evolució de la densitat de probabilitat del paràmetre. Les malles o conjunts de paràmetres a... són els mateixos que a les figures 4 i 5, respectivament.

## 6. DISCUSIÓ I CONCLUSIONS

A part dels filtres subòptims, que com s'ha vist són totalment desaconsellables per a l'estimació de paràmetres de sistemes dels quals se'n coneix l'estructura, a aquest article s'han presentat tres possibles algorismes d'estimació correcta quan el sistema ha estat modelitzat en forma discreta (la qual cosa no introdueix cap restricció). Les simulacions fetes permeten de fer una comparació entre ells que pot ésser interessant de cara a l'elecció en un cas concret.

Quant al filtre Extended Kalman modificat, encara que a l'exemple concret presentat aquí porta a un algorisme prou ràpid, això és degut al fet que hi ha un sol paràmetre a estimar i una sola variable observada. Si en els dos casos n'hi ha més d'un,  $\beta_k$  i  $\pi_k$  passen a ésser tensors de dimensió 3, i això fa que la quantitat de càlculs a fer creixi ràpidament. Per altra banda comparat amb els altres algorismes, és de convergència més lenta i sols ens dona l'estimació que convergeix al paràmetre amb probabilitat 1; però pel fet de no donar-nos la seva densitat de probabilitat no podem saber, per exemple, l'interval de confiança de les estimacions en un moment donat (les covariàncies que dona el filtre no són les veritables ja que la modificació sols assegura la convergència de les estimacions). D'altra banda, l'algorisme és aplicable únicament a sistemes lineals amb paràmetres desconeguts.

Quant al filtre òptim discret, el principal avantatge és la de no estar restringit a cap tipus de sistemes, amb la qual cosa es poden estimar perfectament paràmetres de sistemes no lineals. En canvi, el nombre d'operacions a fer a cada pas i la quantitat d'informació a guardar són grans, i creixen molt ràpidament amb la dimensió del sistema i el nombre de paràmetres a estimar (en canvi quasi no depèn del nombre de variables observades); si a la densitat de probabilitat es volen guardar  $N_p$  punts en cada dimensió, cal emma-

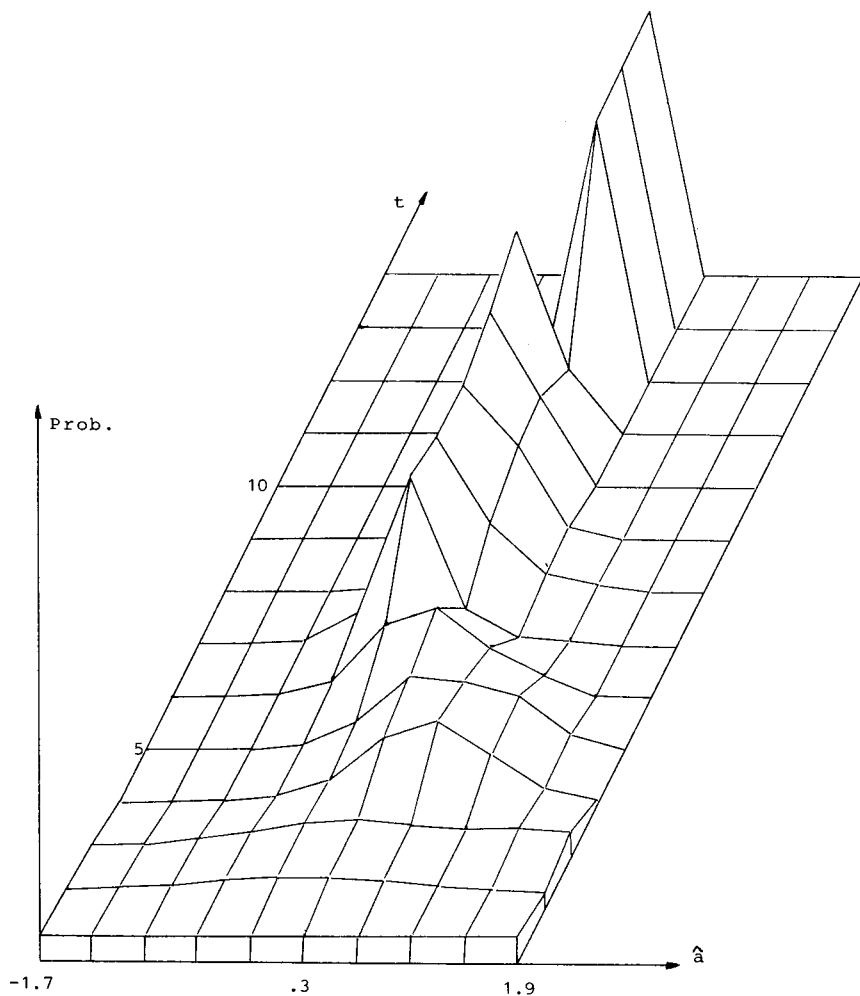


Fig. 6.  
 Densitat de probabilitat del paràmetre.  
 Malla de -1.7 a 1.9

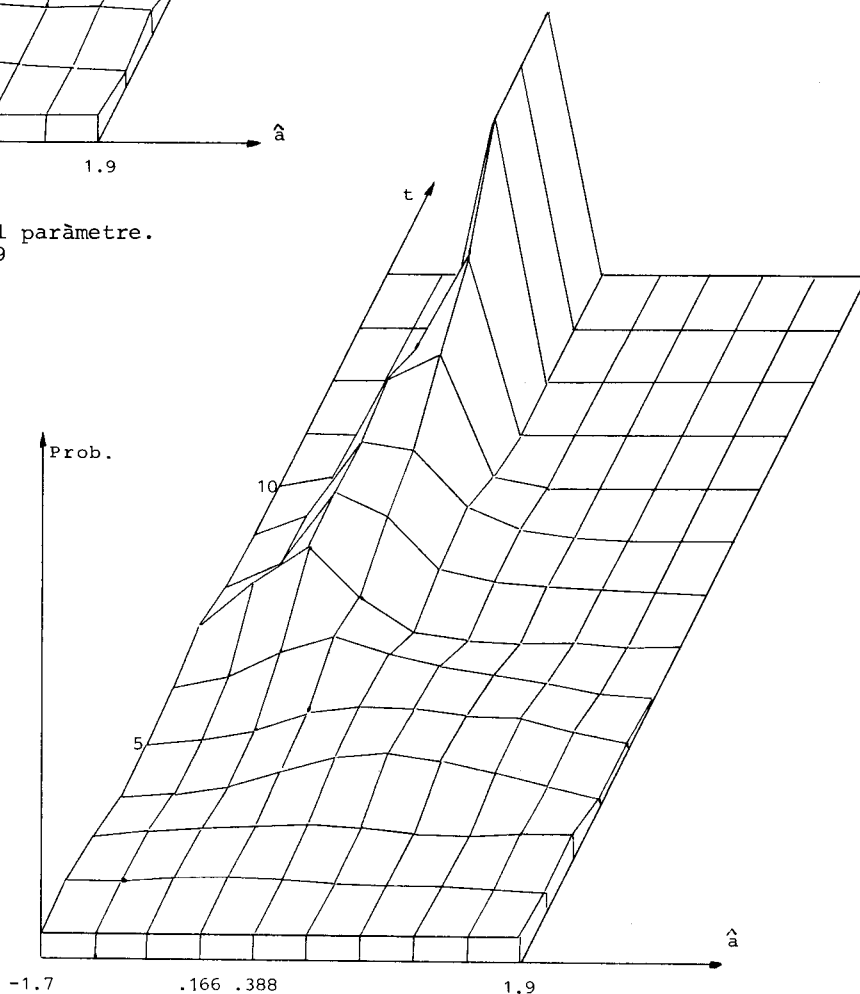


Fig. 7.  
 Densitat de probabilitat del paràmetre.  
 Malla de -.15 a 1.5

gatzemar  $N_p^{(n+m)}$  valors de la densitat de probabilitat, i calcular  $n \cdot N_p^{(n+m)}$  integrals ( $N_p^{(2n+m)}$  sumes) a cada pas. Això fa que aquest filtre no es pugui utilitzar exceptuant si  $n$  i  $m$  són petits. Per altra banda, els problemes d'elecció de la malla (vegeu diferència entre les figures 4 i 5) se solucionen utilitzant una malla que es vagi adaptant a la densitat de probabilitat a mesura que la variància d'aquesta disminueix:

- 1) Iniciliatzaions; adopció d'una malla inicial.
- 2) Fer un pas de l'algorisme del filtre òptim (apartat 4).
- 3) Modificar la malla (normalment es pren com a nova malla la que té com a límits  $\pm M\sigma$  amb  $M$  prefixat i  $\sigma$  desviació tipus de  $a$ ).
- 4) Trobar, per interpolació, els valors de la densitat de probabilitat als punts de la nova malla.
- 5) Tornar a 2).

Finalment, quant al teorema de partició, té l'avantatge d'ésser molt més ràpid que el filtre òptim sempre que el nombre  $N$  de filtres no sigui molt gran. El problema (figures 6 i 7) de l'elecció dels jocs de paràmetres es pot solucionar de forma idèntica al canvi de malla que acabem de veure. De tota manera, quan creix el nombre de paràmetres a estimar,  $N$  pot fer-se molt gran ja que cal, als diferents jocs de paràmetres, anar variant de forma independent els paràmetres  $a_1 \dots a_m$ . L'inconvenient que té aquest mètode, d'altra banda, és que s'aplica a sistemes lineals, i en el cas dels no lineals no s'ha demostrat la seva convergència.

Per tant, es pot concloure que en el cas de sistemes lineals, és més convenient utilitzar el teorema de partició o el filtre Extended Kalman modificat. En canvi, en sistemes no lineals s'haurà d'utilitzar el filtre òptim discret encara que sigui més lent. A part d'això, la complexitat de tots els algorismes creix molt ràpidament en augmentar el nombre de variables  $n$  i el de paràmetres  $m$ .

## 7. REFERÈNCIES

- /1/ SAGE, A.P., MELSA, J.L. "System Identification". Academic Press, 1971.
- /2/ NIHTLÄ, M. "A Program Package For Estimation of Parameters in Non-Linear Differential Equation Systems". Helsinki Univ. of Technology, Control Eng. Lab., Report 7, 1977.
- /3/ JAZWINSKY, A.H. "Stochastic Processes & Filtering Theory". Academic Press, 1970.
- /4/ Mc. GREAVY, C., GILL, P.J. "Estimation of Model Parameters in Noisy, non-linear Systems". Proc. of the 6th IFAC Congress, Boston, 1975.
- /5/ BRUNET, P. "Contribución al Estudio de la Estabilidad de Filtros Subóptimos y de Sistemas Estocásticos de Dimensión Elevada". Tesis Doctoral, U.P.B., 1976.
- /6/ LJUNG, L. "The Extended Kalman Filter as a Parameter Estimator for Linear Systems". Dept. of Electrical Engineering, Linköping Univ., Sweden, Report LITH-ISY-I-0154.
- /7/ BUCY, R.A. "Bayes Theorem and Digital Realizations of Non Linear Filters". J.A.A.S., 2, 1970.
- /8/ BUCY, R.S., YOUSSED, H. "Spline Functions". Proc. 5th Symp. Nonlinear Estimation and its Applications, Sept. 1974.
- /9/ LAINIOTIS, D.G. "Partitioned Estimation Algorithms, I: Nonlinear Estimation". Estimation Theory, American Elsevier Pu. Co. 1974, pp. 13-46.