

COLLOQUE NATIONAL SUR LE TRAITEMENT DU SIGNAL ET SES APPLICATIONS

NICE du 26 au 30 AVRIL 1977

ESTIMATION/IDENTIFICATION DE SYSTEMES LINEAIRES

BRUITES PAR FILTRAGE SOUS-OPTIMUM

MM. C.A. DARMON et J.J. FUCHS

Laboratoire d'Automatique - IRISA - Complexe Scientifique de RENNES Beaulieu - BP 25 A - 35031 RENNES Cédex

RESUME

Deux méthodes d'Estimation/Identification de systèmes linéaires bruités -à partir de la seule connaissance des entrées et des sorties- sont présentées ici. Ces deux méthodes réalisent une approximation à l'estimateur bayésien de l'état et des paramètres, et ont pour originalité d'utiliser un modèle (d'état) canonique du système :

$$X(t+1) = SX(t) - AY(t) + Bu(t) + Ke(t)$$

$$Y(t) = TY(t) + CX(t) + e(t)$$

où $Y(t)$, $u(t)$ et $e(t)$ sont respectivement les vecteurs de sortie, de commande et de bruit, S et C sont à priori connues, et A , B , K et T sont à estimer en même temps que $X(t)$.

La première méthode, l'EIVE 1 est un estimateur linéaire optimal sur un modèle d'état approché où à l'instant t , $e(t-1)$ est remplacé par la pseudo-innovation $\epsilon(t-1)$. La deuxième, l'EIVE 2, est un filtre de Kalman-étendu sur le modèle d'état exact.

Le présent article ne considère que le cas une entrée-une sortie -où T est nécessairement nul. Il est montré que l'EIVE 1 et l'EIVE 2 sont liées aux méthodes de PANUSKA et FUHRT dont les propriétés de stabilité et de convergence sont mathématiquement bien connues. Enfin, les résultats de simulations fournis confirment le bon comportement des EIVE 1 et EIVE 2 sur les plans identification des paramètres et convergence vers le filtre de Kalman.

Ce travail est partiellement supporté par un contrat CETHEDEC.

SUMMARY

Two methods for Estimation/Identification of linear stochastic systems -from inputs and outputs- are introduced in this paper. Both methods approximate the Bayes estimator of the state and parameters, upon a canonical state representation of the system :

$$X(t+1) = SX(t) - AY(t) + Bu(t) + Ke(t)$$

$$Y(t) = TY(t) + CX(t) + e(t)$$

where $Y(t)$, $u(t)$ and $e(t)$ are respectively the output, input, and noise ; S and C are known a-priori, and A , B , K , T are to be estimated as well as $X(t)$.

The first method, EIVE 1, is an optimal linear estimator of an approximate version of the state representation, where at time t , $e(t-1)$ is replaced by the pseudo-innovation $\epsilon(t-1)$. The second method, EIVE 2, is an extended Kalman filter of the time state representation.

This paper deals only with the cas one input-one output -where $T = 0$ -. It is shown that EIVE 1 and EIVE 2 are connected with the methods of PANUSKA and FUHRT, whose stability and convergence properties are well understood from a mathematical point of view. At last, the simulation results confirm the good performance of EIVE 1 and EIVE 2 in identification as well as in steady-state convergence to the Kalman Filter.



ESTIMATION/IDENTIFICATION DE SYSTEMES LINEAIRES

BRUITES PAR FILTRAGE SOUS-OPTIMUM

1. INTRODUCTION

Le présent article décrit deux méthodes de filtrage adaptatif pour des systèmes aléatoires linéaires, basées sur l'utilisation des techniques de filtrage non linéaire approché sur un modèle canonique d'état du système, augmenté du modèle d'évolution des paramètres. Les deux méthodes effectuent l'estimation conjointe de l'état et des paramètres et peuvent donc entre autres être considérées comme des méthodes d'identification récursive.

Etat de l'art en identification récursive :

Les différentes approches en identification récursive appartiennent à deux grandes classes :

(1) Algorithmes identifiant le modèle du système aléatoire sous sa forme autorégressive à moyenne glissante (ARMA) : Recursive least squares [9], Recursive generalized least squares [11], Instrumental Variables [12], Maximum Likelihood [2], Recursive maximum likelihood type 1 (RML1) [3], Recursive maximum likelihood type 2 (RML2) [4].

(2) Algorithmes basés sur le modèle d'état et convergent vers celui-ci par différentes méthodes de stabilité : Algorithmes de KUDVA-NARENDRA [5], d'ANDERSON [6], etc...

Les algorithmes de la classe (1) ont atteint un optimum de performance dans les méthodes de maximum de vraisemblance approché (RML1, RML2). Au contraire les algorithmes du type (2) disponibles jusqu'à l'heure actuelle ne fournissaient guère de meilleurs résultats que la méthode des moindres carrés et étaient en particulier biaisés en présence du bruit. Les algorithmes proposés ici comme, antérieurement, l'algorithme de SALUT [7], viennent établir un pont entre les méthodes du type 1 les plus évoluées et les représentations en variables d'état, à la base des méthodes du type 2.

Etat de l'art en filtrage adaptatif :

A notre connaissance, en dehors de l'algorithme d'estimation optimal de SALUT, et des algorithmes décrits ici, il n'existe pas encore de filtres adaptatifs pour des systèmes aléatoires linéaires. L'algorithme de SALUT, de nature strictement optimale, a eu le mérite de montrer la faisabilité des méthodes d'estimation/identification basées sur le principe de l'estimation bayésienne de paramètres. Cependant, par suite de son caractère strictement optimal, cet algo-

rithme reste très compliqué à implémenter si on le compare à des méthodes tels que le RML1 et le RML2 par exemple.

Philosophie des méthodes présentées :

L'approche suivie ici est de réaliser une approximation de l'estimateur bayésien (Filtre de Kalman étendu, par exemple) qui permet une simplification draconienne des calculs et rend celui-ci compétitif avec les méthodes du RML1 et du RML2 par exemple.

Résumé des chapitres suivants :

Dans le chapitre 2, seront présentées les formes canoniques des systèmes aléatoires linéaires sur lesquelles sont implantées les méthodes décrites. Le chapitre 3 traitera le problème d'estimation bayésienne approchée de l'état et des paramètres. Deux méthodes y seront présentées : une première méthode, l'EIVE1 (Estimateur-Identificateur par variables d'état, n° 1), reposant sur l'approximation de l'innovation par la pseudo-innovation dans l'équation d'état ; une deuxième méthode l'EIVE2, qui est en fait un filtre de Kalman étendu. Le chapitre 4 indiquera les liens qui existent entre l'EIVE1 et l'EIVE2 et les méthodes de PANUSKA et FUHRT (RML1 et RML2). Enfin le chapitre 5 sera consacré à des simulations des algorithmes proposés, avec comparaison aux RML1 et RML2.

2. PRÉSENTATION DES SYSTÈMES LINÉAIRES
(MONOVARIABLES)2.1. Représentation canonique en variables d'état

Soit le système aléatoire linéaire stationnaire :

$$(2.1) \quad \begin{aligned} X(t+1) &= \Phi X(t) + \Gamma u(t) + v(t) & X \in \mathbb{R}^n \\ y(t) &= C X(t) + w(t) & u, y \in \mathbb{R} \end{aligned}$$

où $v(t)$ et $w(t)$ sont des bruits blancs échantillonnés de variance $V_{t-\tau}$ et $W_{t-\tau}$ supposé ici monovariable pour la convenance de l'exposé.

Par un changement de base dans l'espace d'état (cf. [7]), ce système admet comme formulation équivalente à (2.1), la nouvelle formulation :

$$(2.2) \quad \begin{aligned} X(t+1) &= (S-AC)X(t) + B \cdot u(t) + v(t) \\ y(t) &= CX(t) + w(t) \end{aligned}$$

ESTIMATION/IDENTIFICATION DE SYSTEMES LINEAIRES

BRUITES PAR FILTRAGE SOUS-OPTIMUM

sous réserve qu'il soit totalement observable par la sortie $y(t)$, les matrices S , A , B et C étant respectivement :

$$S = \begin{bmatrix} 0 & 1 & 0 \\ & \searrow 1 \\ 0 & & 0 \end{bmatrix} \quad A = \begin{bmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{bmatrix} \quad B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix}$$

Si les bruits blancs $v(t)$ et $w(t)$ sont nuls, la forme (2.2) est canonique, en ce sens que le nombre de paramètres représentant les matrices d'état $S-AC$, de sortie C , et de commande B , est minimum parmi toutes les formes équivalentes à (2.1).

La représentation (2.2) n'est généralement pas canonique si les bruits blancs $v(t)$ et $w(t)$ sont non nuls. L'on montre que, si le système est totalement observable, sa représentation canonique est donnée par le filtre de Kalman, associé à (2.2), d'équations :

$$\hat{X}^-(t+1) = (S-AC) \hat{X}^-(t) + Bu(t) + K(t) \epsilon(t)$$

où $\epsilon(t)$ est donné par :

$$\epsilon(t) = y(t) - \hat{C}\hat{X}^-(t)$$

soit en explicitant la sortie en fonction de l'entrée et de $\epsilon(t)$:

$$(2.3) \quad \begin{aligned} \hat{X}^-(t+1) &= (S-AC) \hat{X}^-(t) + Bu(t) + K(t) \epsilon(t) \\ y(t) &= C\hat{X}^-(t) + \epsilon(t) \end{aligned}$$

La représentation (2.3), appelée "forme filtre", est équivalente à (2.2), donc à (2.1), en ce sens qu'à une suite identique $\{X_0, u(t), 0 \leq t \leq T\}$ il correspond un même processus de sortie $y(t, \omega)$. Or, par suite de la forme optimale du filtre de Kalman, $\epsilon(t)$ est décorrélé des $\hat{X}(\tau)$, $\tau \leq t$ et de $\epsilon(\tau)$, $\tau < t$. Ainsi (2.3) constitue un cas particulier de (2.2) avec $v(t) = K(t)w(t)$.

En régime établi, les équations du système (2.3) deviennent :

$$(2.3') \quad \begin{aligned} \hat{X}^-(t+1) &= (S-AC)\hat{X}^-(t) + Bu(t) + Ke(t) \\ y(t) &= C\hat{X}^-(t) + \epsilon(t) \end{aligned}$$

et l'on montre que la matrice $S-AC-KC$ est asymptotiquement stable.

2.2. Forme autorégressive des systèmes linéaires monovariables

Posons au lieu de K , $K-A$; la forme canonique (2.3') est de la forme :

$$(2.4) \quad \begin{aligned} X(t+1) &= SX(t) - Ay(t) + Bu(t) + Ke(t) \\ y(t) &= CX(t) + \epsilon(t) \end{aligned}$$

Le système (2.4) peut être réécrit sous la forme de transformées en z :

$$(2.5) \quad A(z) y(z) = B(z) u(z) + K(z) \epsilon(z)$$

$$\text{où } A(z) = 1 + a_1 z^{-1} + \dots + a_n z^{-n}$$

$$B(z) = b_1 z^{-1} + \dots + b_n z^{-n}$$

$K(z) = 1 + k_1 z^{-1} + \dots + k_n z^{-n}$ polynôme asymptotiquement stable.

En effet (2.4) s'écrit en transformée en z

$$(2.6) \quad \begin{aligned} X(z) &= z^{-1}SX(z) - z^{-1}Ay(z) + z^{-1}Bu(z) + z^{-1}Ke(z) \\ y(z) &= CX(z) + \epsilon(z) \end{aligned}$$

Soit, en détaillant (2.6) :

$$(2.6.1) \quad x_1(z) = z^{-1}x_2(z) - a_1 z^{-1}y(z) + b_1 z^{-1}u(z) + k_1 z^{-1}\epsilon(z)$$

$$(2.6.2) \quad x_2(z) = z^{-1}x_3(z) - a_2 z^{-1}y(z) + b_2 z^{-1}u(z) + k_2 z^{-1}\epsilon(z)$$

$$(2.6.n) \quad x_n(z) = -a_n z^{-1}y(z) + b_n z^{-1}u(z) + k_n z^{-1}\epsilon(z)$$

$$(2.6.n+1) \quad y(z) = \epsilon(z) + x_1(z).$$

En reportant (2.6.n), ..., (2.6.2) dans (2.6.1), l'on trouve :

$$x_1(z) = -\sum_1^n a_i z^{-i} y(z) + \sum_1^n b_i z^{-i} u(z) + \sum_1^n k_i z^{-i} \epsilon(z)$$

En reportant dans (2.6.n+1), on retrouve la formule demandée (2.5). De plus $K(z) = k_1 z^{-1}$ est asymptotiquement stable par suite de $S-KC$ asymptotiquement stable.

3. ESTIMATION/IDENTIFICATION PAR VARIABLES

D'ETAT

3.1. Estimation/Identification bayésiennes

L'estimation du vecteur d'état et l'identification des paramètres du système linéaire stochastique :

$$X(t+1) = SX(t) - Ay(t) + Bu(t) + Ke(t)$$

$$y(t) = CX(t) + \epsilon(t)$$

peut être effectuée par estimation bayésienne de l'état $X(t)$ et des paramètres $\theta(t) = [A^T, B^T, K^T]^T$



ESTIMATION/IDENTIFICATION DE SYSTEMES LINEAIRES

BRUITES PAR FILTRAGE SOUS-OPTIMUM

probabilisés préalablement.

En d'autres termes la colonne $\theta(t)$ est supposée vérifier l'équation aux différences :

$$\theta(t+1) = \theta(t),$$

et la colonne $[X^T(0), \theta^T(0)]^T = \underline{x}^T(0)$ est affectée d'une mesure $P(\underline{x}(0))$ de moyenne $[\hat{X}^T(0), \hat{\theta}^T(0)]^T$ et de variance $\Sigma(0)$.

L'estimateur bayésien :

$$\underline{x}(t) = E(\underline{x}(t) | y(1), \dots, y(t-1), u_1, \dots, u(t-1))$$

calculable en utilisant l'équation de KUSHNER-STRATONOVITCH (cf. par exemple [8]) est susceptible de fournir à tout instant t l'estimée optimale de l'état et des paramètres au sens de l'écart quadratique conditionnel minimum. G. SALUT [7] a montré les propriétés étonnantes de cet estimateur qui ne fait intervenir que les moments d'ordre inférieur ou égal à 4 de $\underline{x}(t)$.

(a) Dans le cas où $K=0$, l'estimateur bayésien devient linéaire et bénéficie des équations du filtre de Kalman (Méthode des moindres carrés).

La méthode des moindres carrés par variables d'état est simple et réalise les mêmes équations d'identification que la méthode des moindres carrés de YOUNG [9], cf. DARMON [10].

(b) Dans le cas où $K \neq 0$, l'estimateur bayésien est malheureusement fort complexe, et il s'avère plus raisonnable d'en rechercher une approximation par un filtre à deux moments : Filtre de Kalman étendu, filtres du deuxième ordre.

Les deux méthodes décrites ci-après sont basées sur le principe énoncé. L'une est une approximation très brutale de l'estimateur bayésien où la pseudo-innovation $y(t) - \hat{y}(t)$ est supposée être proche du bruit blanc de mesure. L'autre n'est en fait rien d'autre que le filtre de Kalman étendu approchant l'estimateur bayésien.

3.1. Méthode EIVE 1

3.1.1. Equations d'estimation et d'identification

Les équations d'état du modèle à identifier/estimer sont approchées en supposant qu'à l'instant $t+1$ les bruits $e(t), \dots, e(t-k)$ sont proches des erreurs $\epsilon(t), \dots, \epsilon(t-k)$ d'estimation/identification

$$\epsilon(t) = y(t) - \hat{y}(t)$$

Les équations d'état :

$$(3.2) \quad \begin{aligned} X(t+1) &= SX(t) - Ay(t) + Bu(t) + Ke(t) \\ y(t) &= CX(t) + e(t) \end{aligned}$$

sont ainsi approximées par :

$$(3.3) \quad \begin{aligned} X(t+1) &= SX(t) - Ay(t) + Bu(t) + Ke(t) \\ y(t) &= CX(t) + e(t) \end{aligned}$$

C'est sur le modèle (3.3) qu'est calculé l'estimateur "optimal". L'état est augmenté des paramètres A, B et K

$$(3.4) \quad \begin{aligned} X(t+1) &= SX(t) - Ay(t) + Bu(t) + Ke(t) \\ A(t+1) &= A(t) \\ B(t+1) &= B(t) \\ K(t+1) &= K(t) \\ y(t) &= CX(t) + e(t) \end{aligned}$$

Soit :

$$(3.5) \quad \begin{aligned} X(t+1) &= SX(t) + \psi(t) \theta(t) \\ \theta(t+1) &= \theta(t) \\ y(t) &= C \cdot \begin{bmatrix} X(t) \\ \theta(t) \end{bmatrix} + e(t) \end{aligned}$$

avec :

$$(3.6) \quad \theta(t) = [A^T, B^T, K^T]^T$$

$$(3.7) \quad \psi(t) = [-Iy(t), Iu(t), Ie(t)]$$

Les estimées sous-optimales de $X(t)$ et $\theta(t)$ sont données par application des équations du filtre de Kalman sur (3.5) :

$$(3.8) \quad \begin{aligned} \hat{X}^-(t+1) &= S\hat{X}^-(t) + \hat{\psi}(t) \hat{\theta}^-(t) + K_X(t) \epsilon(t) \\ \hat{\theta}^-(t+1) &= \hat{\theta}^-(t) + K_\theta(t) \epsilon(t) \\ \epsilon(t) &= y(t) - C\hat{X}^-(t) \end{aligned}$$

$$(3.9) \quad \begin{bmatrix} K_X(t) \\ K_\theta(t) \end{bmatrix} = \Phi(t) P^+(t) \mathcal{C}^T = \Phi(t) \begin{bmatrix} K_X^+(t) \\ K_\theta^+(t) \end{bmatrix}$$

$$(3.10) \quad P^+(t) = P^-(t) - P^-(t) \mathcal{C}^T (1 + \mathcal{C} P^-(t) \mathcal{C}^T)^{-1} \mathcal{C} P^-(t)$$

$$(3.11) \quad P^-(t+1) = \Phi(t) P^+(t) \Phi^T(t)$$

où $P(t) = \frac{\Sigma(t)}{W}$ où W est la covariance de $e(t)$ et $\Sigma(t)$ la variance d'erreur d'estimation de la colonne $\begin{bmatrix} X(t) \\ \theta(t) \end{bmatrix}$ (satisfaisant le modèle approché (3.5)).

$$\Phi(t) = \begin{bmatrix} S & \psi(t) \\ 0 & I \end{bmatrix}$$

3.2. Méthode EIVE 2

3.2.1. Principe

ESTIMATION/IDENTIFICATION DE SYSTEMES LINEAIRES

BRUITES PAR FILTRAGE SOUS-OPTIMUM

Les équations d'état du modèle à identifier/estimer sont tout d'abord transformées de façon à éliminer le bruit $e(t)$ sur l'état.

Les équations (2.4) :

$$\begin{aligned} X(t+1) &= SX(t) - Ay(t) + Bu(t) + Ke(t) \\ y(t) &= CX(t) + e(t) \end{aligned}$$

sont modifiées en :

$$\begin{aligned} (3.12) \quad X(t+1) &= (S-KC)X(t) + (K-A)y(t) + Bu(t) \\ y(t) &= CX(t) + e(t) \end{aligned}$$

L'estimateur/identificateur EIVE 2 est un filtre de Kalman étendu sur le modèle (3.12) : Les équations en sont établies en linéarisant les équations d'état autour de l'estimée $(\hat{X}^-(t), \hat{K}^-(t))$ sous la forme :

$$\begin{aligned} (3.13) \quad X(t+1) &= (S-\hat{K}^-(t)C)X(t) + (\hat{K}^-(t)-K)\hat{C}\hat{X}^-(t) \\ &\quad + (K-A)y(t) + Bu(t) \\ y(t) &= CX(t) + e(t) \end{aligned}$$

Posons :

$$(3.14) \quad \theta^T = [A^T, B^T, K^T], \text{ et } \hat{\psi}(t) = [-Iy(t), Iu(t), Ie(t)]$$

$$(3.15) \quad \Phi(t) = \begin{bmatrix} S-\hat{K}^-(t)C & \hat{\psi}(t) \\ 0 & I \end{bmatrix}$$

Les équations approchées (3.13) donnent sous forme compacte :

$$\begin{aligned} (3.16) \quad X(t+1) &= (S-\hat{K}^-(t)C)X(t) + \hat{\psi}(t)\theta(t) + \hat{K}^-(t)\hat{C}\hat{X}^-(t) \\ \theta(t+1) &= \theta(t) \\ y(t) &= CX(t) + e(t) \end{aligned}$$

3.2.2. Équations de l'estimateur/identificateur EIVE 2

Les estimées optimales de $X(t)$, $\theta(t)$ sont fournies par les équations récurrentes :

$$(3.17) \quad \hat{X}^-(t+1) = (S-\hat{K}^-(t)C)\hat{X}^-(t) + \hat{\psi}(t)\hat{\theta}^-(t) + \hat{K}^-(t)\hat{C}\hat{X}^-(t) + K_X(t)\varepsilon(t)$$

$$(3.18) \quad \hat{\theta}^-(t+1) = \hat{\theta}^-(t) + K_\theta(t)\varepsilon(t)$$

Elles sont strictement de la même forme que celles de l'EIVE 1. La seule différence entre l'EIVE 1 et l'EIVE 2 est la forme des équations donnant $P^+(t)$ et $P^-(t)$.

4. LIENS DE L'EIVE 1 ET L'EIVE 2 AVEC LES MÉTHODES DU RML1 ET DU RML2 [3, 4]

L'on se propose ici d'expliquer les liens des algorithmes EIVE 1 et EIVE 2 avec les deux méthodes du RML1 et du RML2 [3, 4].

Les méthodes du RML1 et du RML2 réalisent l'identification des paramètres du modèle autorégressif à moyenne glissante (ARMA) équivalent au modèle d'état (2.4)):

$$(4.1) \quad y(t) = -a_1y(t-1) - \dots - a_ny(t-n) + b_1u(t-1) + \dots + b_nu(t-n) + e(t) + k_1e(t-1) + \dots + k_ne(t-n).$$

Elles peuvent toutes deux être vues, soit comme des méthodes de Newton-Raphson (approchées) convergeant vers le maximum de vraisemblance, soit comme des méthodes bayésiennes (approchées) où les paramètres $\theta(t)$ à identifier sont supposées affectées d'une loi de probabilité à priori, et obéissent à l'équation :

$$(4.2) \quad \theta(t) = \theta(t-1).$$

Dans une approche unifiée, c'est la deuxième présentation qui est donnée ici.

4.1. Méthode du maximum de vraisemblance type 1

Dans la méthode du RML1, due originellement à PANUSKA [3], l'équation (4.1) est écrite sous la forme :

$$(4.3) \quad y(t) = \phi^T(t) \theta(t) + e(t)$$

où $\phi^T(t) = [-y(t-1), \dots, -y(t-n), u(t-1), \dots, u(t-n), e(t-1), \dots, e(t-n)]$

et $\theta^T(t) = [a_1, a_2, \dots, a_n, b_1, \dots, b_n, k_1, \dots, k_n]$

Si $e(t-1), \dots, e(t-n)$ étaient connus à l'instant t , les équations (4.3) et (4.2) donneraient lieu à l'utilisation d'un filtre optimum linéaire. Comme en fait $e(t-1), \dots, e(t-n)$ ne sont pas connus, on les suppose égaux à $\varepsilon(t-1) \dots \varepsilon(t-n)$ où $\varepsilon(t) = y(t) - \hat{y}(t)$.

$$\begin{aligned} \hat{y}(t) &= -\hat{a}_1y(t-1) - \dots - \hat{a}_ny(t-n) \\ &\quad + \hat{b}_1u(t-1) + \dots + \hat{b}_nu(t-n) \\ &\quad + \hat{k}_1\varepsilon(t-1) + \dots + \hat{k}_n\varepsilon(t-n) \end{aligned}$$

Soit :

$$(4.4) \quad \varepsilon(t) = -\sum_{i=1}^n \hat{k}_i\varepsilon(t-i) + y(t) + \sum_{i=1}^n \hat{a}_i y(t-i) - \sum_{i=1}^n \hat{b}_i u(t-i)$$



ESTIMATION/IDENTIFICATION DE SYSTEMES LINEAIRES

BRUITES PAR FILTRAGE SOUS-OPTIMUM

L'équation approchée obtenue en posant $\epsilon(t-i) = \epsilon(t-i)$ dans (4.3) est :

$$(4.5) \quad y(t) = \hat{\phi}^T(t)\theta + e(t)$$

avec :

$$(4.6) \quad \hat{\phi}^T(t) = [-y(t-1), \dots, -y(t-n), u(t-1), \dots, u(t-n), \epsilon(t-1), \dots, \epsilon(t-n)]$$

et fournit les équations de filtrage suivantes :

$$(4.7) \quad \hat{\theta}(t) = \hat{\theta}(t-1) + K_{\theta}\epsilon(t)$$

$$(4.8) \quad \epsilon(t) = y(t) - \hat{\phi}(t) \cdot \hat{\theta}(t-1)$$

$$(4.9) \quad K_{\theta}(t) = P(t)\hat{\phi}(t)$$

$$(4.10) \quad P(t) = P(t-1) - \frac{P(t-1)\hat{\phi}(t)\hat{\phi}^T(t)P(t-1)}{1 + \hat{\phi}^T(t)P(t-1)\hat{\phi}(t)}$$

(équation équivalente à $P(t)^{-1} = P(t-1)^{-1} + \hat{\phi}(t)\hat{\phi}^T(t)$ où $P(t)$ est défini par $P(t) = \sum(t)/W$).

Forme de $P(t)$ et de $K_{\theta}(t)$:

Par résolution de l'équation (4.10), il vient:

$$(4.11) \quad P(t) = [P^{-1}(n-1) + \sum_n^t \hat{\phi}(\tau)\hat{\phi}^T(\tau)]^{-1}$$

D'où, par (4.9) :

$$(4.12) \quad K_{\theta}(t) = [P^{-1}(n-1) + \sum_n^t \hat{\phi}(\tau)\hat{\phi}^T(\tau)]^{-1}\hat{\phi}(t)$$

4.2. Méthode du maximum de vraisemblance type 2

Dans la méthode du maximum de vraisemblance type 2, due à FUHRT [4], une meilleure approximation du filtre optimum est réalisée :

$\epsilon(t)$ défini par l'équation (4.4) peut se développer au 1er ordre autour des paramètres estimés $\hat{\theta} = \hat{\theta}(t-1)$:

$$(4.13) \quad \epsilon(\hat{\theta}, t) = \epsilon(\theta, t) - \left[\frac{\partial \epsilon}{\partial \theta}(\hat{\theta}, t) \right]^T (\theta - \hat{\theta})$$

or, en supposant $K(z)$ asymptotiquement stable :

$$\epsilon(\theta, t) = e(t)$$

D'autre part, en supposant $\hat{\theta}$ presque constant (cf. ASTROM-BOHLIN [2]) :

$$\frac{\partial \epsilon}{\partial \theta}(\hat{\theta}, t) = \hat{\phi}^T(t)$$

avec :

$$(4.14) \quad \hat{\phi}^T(t) = [-\hat{y}(t-1), \dots, -\hat{y}(t-n), \hat{u}(t-1) \dots \hat{u}(t-n), \hat{\epsilon}(t-1), \dots, \hat{\epsilon}(t-n)]$$

et :

$$\hat{y}(z) = \frac{y(z)}{\hat{K}(z)}, \quad \hat{u}(z) = \frac{u(z)}{\hat{K}(z)}, \quad \hat{\epsilon}(z) = \frac{\epsilon(z)}{\hat{K}(z)}$$

d'où la formule :

$$(4.15) \quad \epsilon(t) = \epsilon(\hat{\theta}, t) = \hat{\phi}^T(t)\theta - \hat{\phi}^T\hat{\theta} + e(t)$$

La grandeur $\epsilon(t)$ étant maintenant prise comme mesure à l'instant t sur le processus constant :

$$(4.2) \quad \theta(t) = \theta(t-1)$$

les équations de filtrage donnant $\hat{\theta}$ sont :

$$(4.16) \quad \hat{\theta}(t) = \hat{\theta}(t-1) + K_{\theta}(t)[\epsilon(t) - \hat{\phi}(t)]$$

$$(4.17) \quad K_{\theta}(t) = P(t)\hat{\phi}(t)$$

$$(4.18) \quad P(t) = P(t-1) - \frac{P(t-1)\hat{\phi}(t)\hat{\phi}^T(t)P(t-1)}{1 + \hat{\phi}^T(t)P(t-1)\hat{\phi}(t)}$$

(équation équivalente à $P(t)^{-1} = P(t-1)^{-1} + \hat{\phi}(t)\hat{\phi}^T(t)$).

Dans l'équation (4.16), $\hat{\epsilon}(t)$ représente l'estimée à priori de $\epsilon(t)$ soit :

$$\begin{aligned} \hat{\epsilon}(t) &= E(\epsilon(t)|t-1) = E(\epsilon(t)) - \left(\frac{\partial \epsilon}{\partial \theta} \right)^T [\theta - \hat{\theta}(t-1)]|t-1 \\ &= E(\epsilon(t)) - \left(\frac{\partial \epsilon}{\partial \theta} \right)^T (\hat{\theta}(t-1) - \hat{\theta}(t-1)) = 0 \end{aligned}$$

par suite l'équation (4.16) se simplifie en :

$$(4.16') \quad \hat{\theta}(t) = \hat{\theta}(t-1) + K_{\theta}(t)\epsilon(t)$$

devenant par là-même identique à l'équation d'estimation (4.7) de la méthode du RML1. Les autres équations, donnant $K_{\theta}(t)$ et $P(t)$, diffèrent cependant de celles obtenues par la méthode du RML1 : $\hat{\phi}(t)$ est ici remplacée par $\hat{\phi}(t)$ colonne des grandeurs filtrées (par un filtre de la forme $\frac{1}{\hat{K}(z)}$).

Forme de $P(t)$ et de $K_{\theta}(t)$:

Par résolution de l'équation (4.18), il vient:

$$(4.19) \quad P(t) = [P^{-1}(n-1) + \sum_n^t \hat{\phi}(\tau)\hat{\phi}^T(\tau)]^{-1}$$

d'où, par (4.17) :

$$(4.20) \quad K_{\theta}(t) = [P^{-1}(n-1) + \sum_n^t \hat{\phi}(\tau)\hat{\phi}^T(\tau)]^{-1}\hat{\phi}(t)$$

4.3. Liens de l'EIVE 1 avec le RML1

4.3.1. Théorème

L'algorithme EIVE 1 est identique au RML1 pour $t \geq n$.

Démonstration : cf. DARMON [10].

4.3.2. Un théorème fondamental pour le RML1 (SODERSTROM [1])

Soit $\phi(t)$ et $\epsilon(t)$ solution en régime établi au point θ des équations

Posons : $f(\theta) = E(\phi(t)\epsilon(t))$ et $G(\theta) = E(\phi(t)\phi^T(t))$,

$$D_p = \{\theta \mid S-KC \text{ asymptotiquement stable}\},$$

$$D_I = \{\theta \mid \exists [G(\theta)]^{-1}\}$$

ESTIMATION/IDENTIFICATION DE SYSTEMES LINEAIRES

BRUITES PAR FILTRAGE SOUS-OPTIMUM

Théorème

Soit le système (4.1) avec des entrées $u(t)$ stationnaires et indépendantes des $e(t)$, et supposons que pour tout $p > 0$ $E(|e(t)|^p) < \infty$

a) Alors $\theta(t) \rightarrow \theta^* \in D_p \cap D_I$ avec une probabilité non nulle seulement si :

$$i) \quad f(\theta^*) = 0$$

ii) $G^{-1}(\theta^*) \left[\frac{\partial}{\partial \theta} (f(\theta)) \right]_{\theta^*}$ a toutes ses valeurs propres non positives.

b) Si $\hat{\theta}(t)$ appartient à un sous-ensemble compact de D infiniment souvent p.s. et $(\theta, R) = (\theta^*, G(\theta^*))$ ($\theta^* \in D_I$) est un point globalement asymptotiquement stable des systèmes d'équations différentielles :

$$(4.21) \quad \frac{d\theta(\tau)}{d\tau} = G^{-1}(\theta(\tau))f(\theta(\tau))$$

$$(4.21') \quad \frac{dR(\tau)}{d\tau} = -R(\tau) + G(\theta(\tau))$$

Alors $\hat{\theta}(t) \xrightarrow[t \rightarrow \infty]{} \theta^*$ p.s.

Démonstration

La démonstration est fournie dans SODERSTROM [1] et repose sur la constatation que la solution aux équations (4.7) à (4.10) est "peu différente" de la solution de (4.21) et (4.21') en posant

$$R(t) = \frac{P^{-1}(t)}{t}$$

et (approximativement) $\tau = \log t$.

4.3.3. Application à l'EIVE 1

L'application du théorème 4.3.2 à l'EIVE 1 est immédiate par suite du théorème 4.3.1 et fournit le résultat suivant (cf. SODERSTROM [1]) :

L'EIVE 1 ne peut converger dans $D_p \cap D_I$ qui vers des points θ^* où :

$$(4.22) \quad E(\epsilon(t) \phi(t)) = 0$$

c'est-à-dire, $1 \leq k \leq n$:

$$E(\epsilon(t)e(t-k)) = 0, \quad E(\epsilon(t)y(t-k)) = 0, \quad E(\epsilon(t)u(t-k)) = 0$$

Les points θ^* satisfaisant la condition ci-dessus ont été étudiés par SODERSTROM [1] qui énonce la proposition suivante :

Proposition

Si une des conditions (a) ou (b) est réalisée :

$$(a) \quad u(t) = 0$$

$$(b) \quad E \left[\begin{bmatrix} -y(t-1) \\ -y(t-n) \\ u(t-1) \\ u(t-n) \end{bmatrix} \left[-y(t-1) \dots -y(t-n) \quad u(t-1) \dots u(t-n) \right] \right] > 0,$$

$$E(u(t)e(t)) = 0 \quad \text{et} \quad E(u^2(t)) \ll E(e^2(t))$$

Alors le seul point θ^* satisfaisant (4.22) est :

$$\theta^* = \theta = [a_1, a_2 \dots a_n, b_1, b_2 \dots b_n, k_1, \dots, k_n]^T$$

Néanmoins θ ne satisfait pas nécessairement la condition de stabilité locale a) du théorème

4.3.2. Plusieurs contre-exemples à cette stabilité locale pour le RML1 (donc l'EIVE 1) ont ainsi été présentés dans [1]. Heureusement, dans "la plupart des cas", la stabilité locale existe et permet de vérifier que s'il y a convergence des algorithmes EIVE 1 et RML1 avec les conditions de la proposition 4.3.3, les paramètres trouvés sont bien ceux du modèle.

4.4. Liens de l'EIVE 2 avec le RML2

Il a été établi heuristiquement dans DARMON [10] que si l'EIVE 2 converge avec une probabilité non nulle vers un θ^* appartenant à D_p , il présente asymptotiquement les mêmes équations d'évolution que le RML2. Une étude mathématique est en cours dans ce sens, dont le résultat attendu est que :

les points possibles de convergence $\theta^* \in D_p \cap D_I$ de l'EIVE 2 satisfont (comme pour le RML2) :

$$i) \quad E(\epsilon(t) \tilde{\phi}(t)) = 0 \quad (\tilde{\phi}(t) \text{ défini en (4.14)})$$

$$ii) \quad \{E(\tilde{\phi}(t) \tilde{\phi}(t))\}^{-1} \left. \frac{\partial}{\partial \theta} E(\epsilon(t) \tilde{\phi}(t)) \right|_{\theta^*} \text{ a toutes ses valeurs propres non positives.}$$

Ces points correspondent à des minimums locaux en θ de $E(\epsilon^2(t))$ et contiennent dans le cas gaussien le point à maximum de vraisemblance (où $E(\epsilon^2(t))$ est un minimum absolu) représenté par θ la colonne des paramètres du modèle.

De plus, SODERSTROM donne dans [1] des conditions suffisantes pour que le seul θ^* possible soit θ . Celles-ci sont données par :

Proposition

Si une des trois conditions (a), (b) ou (c) est réalisée :

$$(a) \quad u(t) = 0$$

$$(b) \quad E \left[\begin{bmatrix} -y(t-1) \\ -y(t-n) \\ u(t-1) \\ u(t-n) \end{bmatrix} \left[-y(t-1) \dots -y(t-n) \quad u(t-1) \dots u(t-n) \right] \right] > 0,$$



ESTIMATION/IDENTIFICATION DE SYSTEMES LINEAIRES

BRUITES PAR FILTRAGE SOUS-OPTIMUM

$$E(u(t)e(\tau)) = 0 \quad \text{et} \quad E(u^2(t)) \ll E(e^2(t))$$

$$(c) \quad E \begin{bmatrix} -y(t-1) \\ -y(t-n) \\ u(t-1) \\ u(t-n) \end{bmatrix} [-y(t-1) \dots -y(t-n), u(t-1) \dots u(t-n)] > 0$$

$$E(u(t)e(\tau)) = 0 \quad \text{et} \quad E(u^2(t)) \gg E(e^2(t))$$

Alors le seul θ^* possible est :

$$\theta^* = \theta = [a_1 \dots a_n, b_1 \dots b_n, k_1 \dots k_n]^T$$

5. RESULTATS DE SIMULATION

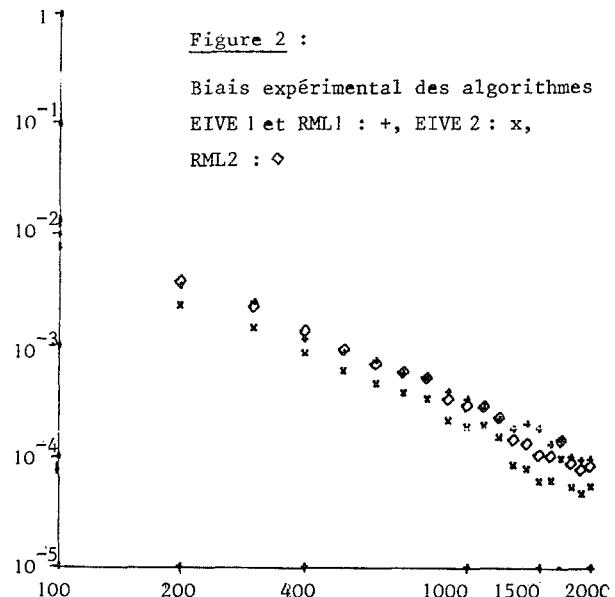
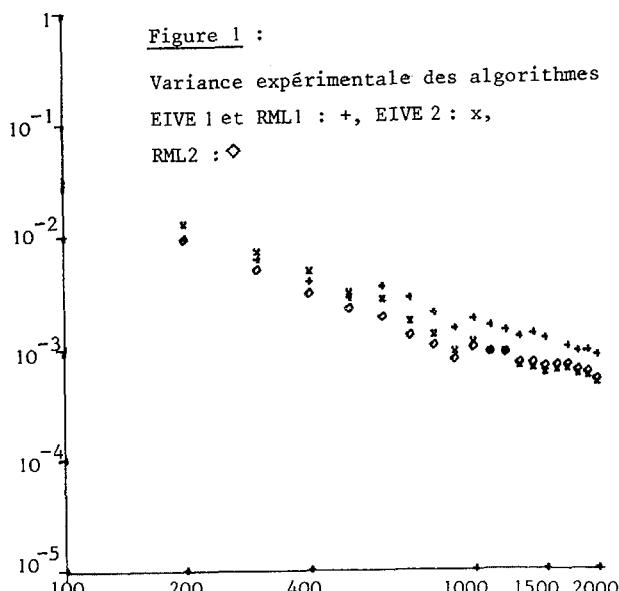
L'EIVE 1 et l'EIVE 2 sont simulés ici sur le modèle autorégressif suivant, d'ordre 1 :

$$(5.1) \quad y(t) - 0,8y(t-1) = u(t-1) + e(t) + 0,7e(t-1)$$

$e(t)$ est une séquence gaussienne blanche de variance $E(e^2(t)) = 1$. Le signal d'entrée $u(t)$ est caractérisé par une suite de nombres aléatoires indépendants équirépartis sur $[-1, +1]$. Les deux signaux $u(t)$ et $e(t)$ sont décorrélés. Les résultats fournis ici portent :

1) Sur la qualité de l'EIVE 1 et de l'EIVE 2 en tant qu'organes d'identification, comparés au RML1 et au RML2.

2) Sur la convergence des estimées de l'état pour l'EIVE 1 et l'EIVE 2 vers l'état du filtre de Kalman représenté par (5.1).



Résultats d'identification :

Ces résultats sont résumés dans les figures 1 et 2 qui représentent les moyennes sur 15 passages de 2 000 équations chacun de :

- la variance expérimentale de $\hat{\theta}(t)$: $\frac{1}{N} \sum_{n=1}^N \|\hat{\theta}_n(t) - \theta\|^2$,
- le biais expérimental normalisé de $\hat{\theta}(t)$: $\frac{1}{N} \sum_{n=1}^N \|\hat{\theta}_n(t) - \theta\|^2 / \|\theta\|^2$,

pour les méthodes de l'EIVE 1 et l'EIVE 2 comparées aux RML1 et RML2, tous les paramètres étant initialisés à 0 dans chacun des cas.

Les courbes obtenues indiquent que :

1°) l'EIVE 1 et le RML1 sont identiques comme annoncé dans 4.3.1.

2°) le RML2 et l'EIVE 2 sont meilleurs en précision (variance expérimentale) et biais que le RML1 et l'EIVE 1.

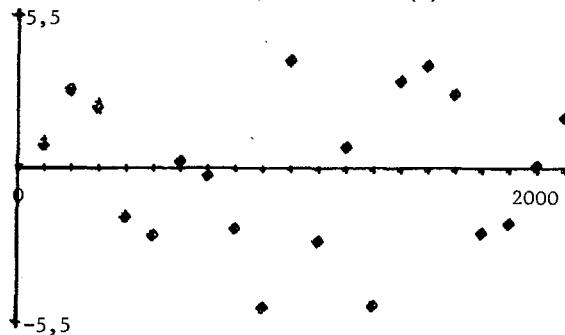
3°) l'EIVE 2 est légèrement plus performant que le RML2.

Résultats d'estimation :

La figure 3 montre pour une réalisation particulière de (5.1) l'évolution en fonction du temps de $X(t)$ état du filtre de Kalman, $\hat{X}_1(t)$ état de l'estimateur EIVE 1, $\hat{X}_2(t)$ état de l'estimateur EIVE 2. On vérifie que $\hat{X}_1(t)$ et $\hat{X}_2(t)$ convergent bien vers $X(t)$ pour $t \rightarrow \infty$.

ESTIMATION/IDENTIFICATION DE SYSTEMES LINEAIRES
BRUITES PAR FILTRAGE SOUS-OPTIMUM

Figure 3 : Etat et estimation de l'état
par EIVE1 (*) et EIVE2 (\diamond)



6. CONCLUSION

Deux méthodes d'estimation et d'identification de systèmes linéaires bruités à partir de la seule connaissance des entrées et des sorties ont été présentées. La première -l'EIVE 1- est un filtre de Kalman sur la représentation canonique d'état augmentée, approchée en supposant le résidu $\epsilon(t)$ égal au bruit blanc de la représentation. La deuxième méthode -l'EIVE 2- est un filtre de Kalman étendu sur le modèle canonique augmenté, après modification de celui-ci de façon à faire disparaître le terme de bruit sur l'état.

L'on a indiqué que l'EIVE 1 s'identifie au RML1 de PANUSKA, et que l'EIVE 2 admet asymptotiquement, sous certaines conditions, les mêmes trajectoires que le RML2 de FUHRT dans l'espace des paramètres.

Enfin des résultats expérimentaux ont été présentés qui permettent de comparer les performances des EIVE aux RML. Ces résultats s'avèrent très encourageants.

Notons en conclusion que l'EIVE 1 et l'EIVE 2 s'appliquent également à l'estimation/identification des systèmes multidimensionnels (cf. [10]) et qu'ils sont (comme le RML1 et le RML2) facilement modifiables en des algorithmes de type gradient stochastique beaucoup moins coûteux en temps de calcul.

Remerciements :

Les auteurs remercient B. LE PARMENTIER, stagiaire de DEA de la Faculté de Rennes, pour sa collaboration en général et notamment dans l'obtention des résultats expérimentaux.

7. BIBLIOGRAPHIE

- [1] SODERSTROM-LJUNG-GUSTAVSSON "A comparative study of recursive identification methods" - Rapport n° 7 424 Division of Automatic Control, Lund Institute of Technology (1974).
- [2] ASTROM-BOHLIN "Numerical identification of linear dynamic systems from normal operating records" IFAC Symposium on Self-Adaptative Systems, Teddington (1965).
- [3] PANUSKA "A stochastic approximation method for identification of linear systems using adaptative filtering". hoc. IACC (1968).
- [4] FUHRT "New estimator for the identification of dynamic processes". Rapport IBK, Inst. Boris Kidric Unica, Belgrade (1973).
- [5] KUDVA-NARENDRA "An identification procedure for discrete multivariable systems". IEEE Trans. AC, vol. 19, oct. 1974.
- [6] ANDERSON "Adaptative identification of multiple-input multiple-output plants". Proc. IEEE - Conf. on Déc. Control, (1974).
- [7] SALUT "Identification optimale des systèmes linéaires stochastiques". Thèse de Doctorat d'Etat, Université P. Sabatier, Toulouse (1976).
- [8] JAZWINSKY "Stochastic processes and filtering theory" N.Y. Academic Press (1970).
- [9] YOUNG "The use of linear regressive and related procedures for the identification of dynamic processes". Symp. on Adapt. Processes, UCLA (1968).
- [10] DARMON "Deux méthodes d'Estimation/Identification de systèmes linéaires bruités multidimensionnels", Revue du CETHEDEC, 4e trimestre 1976.
- [11] HASTINGS JAMES-SAGE "R.G.L.S. procedure for on-line identification of process parameters" IEE Proceedings, vol. 116, pp 2057-2062 (1969).
- [12] WONG-POLAK "Identification of linear discrete time system using the instrumental variable method". IEEE-TAC-12, pp. 707-718 (1967).

