

FILTRAGE ET IDENTIFICATION <i>FILTERING AND IDENTIFICATION</i>
--

ALGORITHMES DE CALCUL DE MODELES MARKOVIENS POUR FONCTIONS ALEATOIRES

P. FAURRE

SAGEM et LABORIA (France)

1 - INTRODUCTION

1.1 - Au cours des trentes dernières années, la nécessité de réaliser des systèmes de plus en plus performants a conduit à considérer et à modéliser les signaux aléatoires ou les perturbations aléatoires intervenant dans ces systèmes.

De nouvelles disciplines comme le filtrage statistique, la théorie de la prédiction, la théorie de la détection, la commande optimale stochastique se sont développées. Les contributions initiales de WIENER et KOLMOGOROV furent considérables.

Dans l'approche la plus ancienne, les fonctions aléatoires sont décrites par leurs covariances ou leurs spectres, paramètres qui sont le résultat brut d'une analyse statistique d'échantillons.

Le développement des calculateurs numériques et des méthodes de variables d'état en automatique déterministe conduisit Kalman et Bucy à développer vers 1960 une approche différente fondée essentiellement sur une représentation par équations différentielles ou récurrentes, linéaires et stochastiques, ou encore, ce qui est équivalent, par processus gaussiens-markoviens (KALMAN (1) et (2)).

Kalman et Bucy appliquèrent cette technique au problème du filtrage statistique et de la prédiction. Les représentations gaussiennes-markoviennes s'avèrent aussi très puissantes pour le problème de lissage(BRYSON)et de la détection (SAGE).

1.2 - Ce papier résume quelques travaux (ANDERSON, FAURRE, KALMAN) faisant le lien entre les deux types de modèles. Plus précisément comment obtenir à partir de la covariance d'un processus gaussien vectoriel, une représentation markovienne ?

Un résultat mathématique connu sous le nom de lemme positif réel caractérise l'ensemble \mathcal{F}^{D} de toutes les réalisations markoviennes d'un processus gaussien décrit par sa covariance. Cet ensemble s'avère être convexe, fermé, borné avec deux points extrémaux max et min, P^* et P_* . Des algorithmes permettent de calculer effectivement ces réalisations extrémales.

De plus, l'une des réalisations extrémales, P_* , s'avère être le filtre optimal.

De nouveaux résultats pour le cas non stationnaire ont été obtenus récemment (voir ATTASI, CLERGET), mais on se limitera ici au cas stationnaire.

2 - ENSEMBLE DES REALISATIONS MARKOVIENNES D'UN

PROCESSUS STATIONNAIRE2.1 - Réalisations markoviennes

Nous considérerons dans la suite une échelle de temps continue et renvoyons à FAURRE (4) pour l'analyse dans le cas d'une échelle de temps discrète. Tous les processus seront supposés centrés.

La donnée du problème considéré sera un processus gaussien vectoriel stationnaire $y(t)$ de dimension m de covariance connue

$$E \{y(t+\tau)y'(t)\} = \Lambda(\tau) \quad (*) \quad (1)$$

et de spectre (transformée de Laplace de la covariance)

$$S(s) = \mathcal{L}\{\Lambda(\tau)\} = \int e^{-s\tau} \Lambda(\tau) d\tau \quad (2)$$

Ce processus sera dit admettre une représentation markovienne, s'il existe un processus markovien $x(t)$ de dimension finie n tel que

$$y(t) = Hx(t) + w(t) \quad (3)$$

où le terme éventuellement présent $w(t)$ est un bruit blanc.

On peut montrer qu'un processus gaussien stationnaire markovien vérifie nécessairement une équation différentielle linéaire excitée par un bruit blanc (FAURRE(4)) (on voit ainsi l'analogie entre des représentations d'état et les processus markoviens), ce qui conduit à la définition suivante :

DEFINITION :

On appelle réalisation markovienne (si elle existe) d'un processus gaussien stationnaire un modèle de la forme :

$$\dot{x} = Fx + v \quad (4)$$

$$y = Hx + w \quad (5)$$

où $\begin{bmatrix} v \\ w \end{bmatrix}$ est un bruit blanc de covariance

(*) La notation y' désigne le transposé de la matrice y .

$$E \left\{ \begin{bmatrix} v(t) \\ w(t) \end{bmatrix} \begin{bmatrix} v'(s)w'(s) \end{bmatrix} \right\} = \begin{bmatrix} Q & S \\ S' & R \end{bmatrix} \delta(t-s) \quad (6)$$

et pour lequel les conditions suivantes sont vérifiées

$$F \quad : \quad \text{matrice asymptotiquement stable} \quad (7)$$

$$(F,L) \quad : \quad \text{complètement commandable avec } Q = LL' \quad (8)$$

$$(H,F) \quad : \quad \text{complètement observable} \quad (9)$$

On peut voir que la covariance du processus $y(t)$ ainsi représenté :

$$\Lambda(\tau) = E\{y(t+\tau)y'(t)\} \quad (10)$$

$$= E\left\{ \left[H \int_0^\infty e^{F\alpha} v(\tau-\alpha) d\alpha + w(\tau) \right] \left[H \int_0^\infty e^{F\beta} v(-\beta) d\beta + w(0) \right]' \right\}$$

est de la forme

$$\Lambda(\tau) = He^{F\tau} G \varepsilon(\tau) + G'e^{-F'\tau} H' \varepsilon(-\tau) + R\delta(\tau) \quad (11)$$

où l'on a posé

$$G = PH' + S \quad (12)$$

avec

$$P = E\{x(t)x'(t)\} \quad (13)$$

solution de l'équation

$$FP + PF' = -Q \quad (14)$$

et

$$\varepsilon(\tau) = \begin{cases} 0 & \text{si } \tau < 0 \\ 1/2 & \text{si } \tau = 0 \\ 1 & \text{si } \tau > 0 \end{cases} \quad (15)$$

et où $\delta(\tau)$ est l'impulsion de Dirac.

Le spectre de $y(t)$, transformée de Laplace de (11) est donc

$$S(s) = R + H(sI-F)^{-1}G + G'(-sI-F')^{-1}H' \quad (16)$$

On remarque immédiatement que $S(s)$ est rationnel en s et parahermitien ($S(s) = S'(-s)$)

Ainsi, nous voyons déjà apparaître deux conditions nécessaires (qu'on verra ultérieurement suffisantes) pour l'existence de réalisations markoviennes :

i) la covariance $\Lambda(\tau)$ doit être de type positif (comme toute covariance) ou de façon équivalente le spectre doit vérifier (théorème de Bochner)

$$S(j\omega) \geq 0 \text{ pour tout } \omega \text{ réel } (*) \quad (17)$$

ii) le spectre doit être une fonction rationnelle de la fréquence s , ou encore la covariance $\Lambda(\tau)$ doit être de la forme (11).

Pour les applications, la covariance $\Lambda(\tau)$ vient en général d'une estimation statistique. Il est donc important de pouvoir exprimer $\Lambda(\tau)$ par une formule du type (11). Il s'agit en fait de méthodes d'approximation du type approximation de Padé, et un algorithme utilisable pour cela est l'algorithme de HO (voir HO, IRVING, RISSANEN, FAURRE (4)).

Nous supposons donc pour la suite que les données du problème, à savoir la covariance $\Lambda(\tau)$, sont en fait les 4 matrices $\{H, F, G, R\}$. Les inconnues sont alors $\{P, Q, S\}$.

2.2 - Ensemble \mathcal{P} de toutes les réalisations markoviennes

Une réalisation markovienne particulière sera parfaitement définie par la matrice P correspondante : en effet les matrices Q et S s'en déduisent par les relations (14) et (12).

On appellera \mathcal{P} l'ensemble des matrices P correspondant à une réalisation markovienne de $\Lambda(\tau)$ définie par (11).

On peut montrer que cet ensemble \mathcal{P} est caractérisé par les relations

$$\left\{ \begin{array}{l} FP + PF' = -Q \\ G - PH' = S \end{array} \right. \quad (18)$$

$$\left. \begin{array}{l} G - PH' = S \\ \left[\begin{array}{cc} Q & S \\ S' & R \end{array} \right] \geq 0, P \geq 0 \end{array} \right\} \quad (19)$$

$$\left. \begin{array}{l} \left[\begin{array}{cc} Q & S \\ S' & R \end{array} \right] \geq 0, P \geq 0 \end{array} \right\} \quad (20)$$

Les relations (18) et (19) ont déjà été vues et (20) est évident car les matrices correspondantes sont des covariances ; la réciproque peut être démontrée (voir FAURRE (4)).

Ces relations (18), (19) et (20) interviennent dans un résultat mathématique important connu sous le nom de lemme positif réel (POPOV) qui est le suivant
THEOREME (Lemme positif réel)

La fonction $\Lambda(\tau)$ donnée par l'expression (11) est de type positif si, et seulement s'il existe trois matrices P, Q, S (P et Q symétriques) vérifiant (18), (19) et (20).

(*) La notation $A \geq 0$ signifie que la matrice A est non négative définie.

La structure de l'ensemble \mathcal{P} est intéressante. Il est assez facile de voir que \mathcal{P} est convexe fermé dans l'espace des matrices symétriques nxn. De plus on peut montrer (voir FAURRE (4)) que relativement à la relation d'ordre

$$A \leq B \iff B - A \text{ non négative définie,} \quad (21)$$

il admet un maximum P^* et un minimum P_* , c'est-à-dire que

$$P \in \mathcal{P} \implies P_* \leq P \leq P^* \quad (22)$$

3 - ALGORITHMES DE CALCUL DES REALISATIONS EXTREMALES

Un résultat théorique utile introduit dans FAURRE (2) est la liaison existant entre la réalisation extrémale P et la covariance $\Lambda(\tau)$, qu'on peut exprimer ainsi :

$$x'P^*x = \inf_{u(\cdot) \in \mathcal{E}(x)} \int_{-\infty}^0 \int_{-\infty}^0 u'(\alpha) \Lambda(\beta-\alpha) u(\beta) d\alpha d\beta \quad (23)$$

où l'ensemble $\mathcal{E}(x)$ est défini par

$$\mathcal{E}(x) = \{u(\cdot) \mid x = \int_{-\infty}^0 e^{-F'\alpha} H'u(\alpha) d\alpha\} \quad (24)$$

La matrice P^* apparaissant comme infimum d'un critère quadratique sous contrainte linéaire, on peut montrer que P^* est la valeur limite d'une équation de Riccati.

Dans le cas régulier (R définie positive) les algorithmes sont les suivants :
ALGORITHME

Le maximum P^* s'obtient en intégrant

$$\begin{cases} \Sigma(o) = 0 \\ \dot{\Sigma} = F'\Sigma + \Sigma F + (H' - \Sigma G)R^{-1}(H - G'\Sigma) \end{cases} \quad (25)$$

jusqu'à convergence vers la valeur d'équilibre Σ_∞ , alors

$$P^* = \Sigma_\infty^{-1} \quad (26)$$

Le minimum P_* s'obtient en intégrant

$$\begin{cases} \Omega(o) = 0 \\ \dot{\Omega} = F\Omega + \Omega F' + (G - \Omega H')R^{-1}(G' - H\Omega) \end{cases} \quad (27)$$

jusqu'à convergence vers la valeur d'équilibre Ω_∞ , alors

$$P_* = \Omega_\infty \quad (26)$$

Ces algorithmes sont stables et efficaces pour calculer des réalisations markoviennes (voir GERMAIN). Les méthodes initiales et plus anciennes pour trouver une réalisation markovienne sont dérivées des idées de factorisation spectrale. Alors que pour un processus scalaire les deux approches sont comparables, dans le cas multivariable les algorithmes que nous proposons semblent l'emporter nettement sur la complexité des factorisations de matrices rationnelles.

4 - APPLICATIONS

4.1 - Filtre statistique optimal

Rappelons que le problème du filtrage statistique consiste à calculer la meilleure estimée $\hat{x}(t)$ de $x(t)$ à partir des observations passées $y(\alpha)$, $\alpha \leq t$:

$$\hat{x}(t) = E\{x(t) \mid y(\alpha), \alpha \leq t\} \quad (29)$$

Pour un modèle markovien de la forme (4) à (6), l'algorithme de filtrage porte le nom de filtre de Kalman (voir KALMAN (1) et (2)) et s'écrit

$$\dot{\hat{x}} = F\hat{x} + K[y - H\hat{x}] \quad (30)$$

$$\text{avec } K = (\Sigma H' + S)R^{-1} \quad (31)$$

et Σ étant l'unique solution définie positive de l'équation de Riccati stationnaire du filtre.

$$F\Sigma + \Sigma F' - (\Sigma H' + S)R^{-1}(H\Sigma + S') + Q = 0. \quad (32)$$

Si l'on écrit les équations (30) et (5) sous la forme

$$\dot{\hat{x}} = F\hat{x} + TW \quad (33)$$

$$y = H\hat{x} + W \quad (34)$$

on voit que le filtre est une réalisation markovienne particulière caractérisée par

$$\begin{bmatrix} \dot{v} \\ \dot{w} \end{bmatrix} = \begin{bmatrix} TW \\ W \end{bmatrix} \quad (35)$$

Il est possible de montrer (voir FAURRE (4)) que cette réalisation markovienne particulière est P_* .

Ainsi donc la réalisation markovienne minimale P_* présente-t-elle la particularité d'être le filtre de Kalman.

4.2 - Implications pratiques

Lorsqu'on se trouve devant un problème concret de filtrage statistique pour lequel les modèles sont mal connus, le premier problème est celui de les obtenir par des méthodes statistiques.

Lorsqu'on cherche à obtenir un modèle markovien, deux approches sont utilisables :

- i) une approche paramétrique : des hypothèses de structure et de dimension étant faites sur le processus, on estime au mieux les paramètres
- ii) l'approche utilisant les algorithmes présentés dans ce papier, qui ne fait aucune autre hypothèse a priori que la stationnarité et qui consiste :

- à estimer la covariance du processus par des moyens d'analyse statistique
- à utiliser des algorithmes du type HO, RISSANEN qui génèrent les matrices $\{H, F, G, R\}$ représentant la covariance
- à utiliser les algorithmes présentés ci-dessus pour calculer une ou deux réalisations extrémales P^* ou P_* , ou toute combinaison convexe de ces deux là.

Ayant ainsi obtenu une réalisation markovienne, il est possible de calculer explicitement le filtre de Kalman.

On pouvait faire reproche à cette procédure d'avoir à résoudre successivement deux équations de Riccati, une pour obtenir la réalisation markovienne et une autre pour le filtre.

Or il ressort du résultat nouveau cité plus haut que la réalisation P_* n'est rien d'autre que le filtre, et donc que la résolution d'une seule équation de Riccati suffit pour résoudre le problème global réalisation markovienne et filtre.

BIBLIOGRAPHIE

B.D.O. ANDERSON - The Inverse Problem of Stationary Covariance Generation, J. of Statistical Physics, Vol. 1, No 1, P. 133-147, 1969.

S. ATTASI - Systèmes linéaires homogènes à deux indices, Rapport LABORIA (à paraître)

A.E. BRYSON, M. FRAZIER - Smoothing for Linear and Nonlinear Dynamic Systems, Proc. Optimum System Synthesis conf., U.S. AIR Force Tech. Dept ASD-TDR, p. 63-119, Feb. 1963.

M. CLERGET - Systèmes linéaires positifs non stationnaires, Rapport LABORIA (à paraître).

P. FAURRE, J.P. MARMORAT (1) - Un algorithme de réalisation stochastique, C.R. Acad. Sc. Paris, T. 268, Sér. A, p. 978-981, Avril 1969.

P. FAURRE (2) - Identification par minimisation d'une représentation markovienne de processus aléatoire, Symp. on Optimization, Nice, June 1969 édité par Springer-Verlag, Lecture Notes in Mathematics 132.

P. FAURRE, P. CHATAIGNER (3) - Identification en temps réel et en temps différé par factorisation de matrices de Hankel, Colloque Franco-Suédois sur la Conduite des Procédés, IRIA, Octobre 1971.

P. FAURRE (4) - Réalisations Markoviennes de Processus Stationnaires, Rapport LABORIA No 13, Mars 1973 (IRIA, Voluceau, 78-ROCQUENCOURT, FRANCE).

F. GERMAIN - Algorithmes continus de calcul de réalisations markoviennes. Cas singulier et stabilité, Rapport LABORIA (à paraître).

B.L. HO, R.E. KALMAN - Effective Construction of Linear State Variable Models from Input/Output Data, Proc. 3 rd Allerton Conference, p. 449-459, 1965.

E. IRVING - Identification des Systèmes, II, Note EDF, HR 8202, 1968.

R.E. KALMAN (1) - A New Approach to Linear Filtering and Prediction Problems, Journal of Basic Engineering, p. 35-45, March 1960.

R.E. KALMAN, R.S. BUCY (2) - New Results in Linear Filtering and Prediction Theory, Journal of Basic Engineering, p. 95-108, March 1961.

R.E. KALMAN(3) - Linear Stochastic Filtering - Reappraisal and Outlook, Symposium on System Theory, Polytechnic Institute of Brooklyn, p. 197-205, April 1965.

T. KAILATH - An innovation approach to least square, estimation Part I-IV, IEEE Trans. AC. 13, 13, 16 et 16.

A.N. KOLMOGOROV - Sur l'interpolation et l'extrapolation des suites stationnaires C.R. Acad. Sc. Paris, 208, p. 2043-45, 1939.

V. POPOV - L'hyperstabilité des systèmes automatiques, Dunod, 1973.

J. RISSANEN - Recursive Identification of Linear Systems, J. SIAM Control, Vol. 9, No 3, p. 420-430, August 1971.

A.P. SAGE, J.L. MELSA - Estimation Theory with Applications to Communications and Control, Mc Gray Hill, 1971.

N. WIENER - Extrapolation, Interpolation and Smoothing of Stationary Time Series, The M.I.T. Press, 1949.