



**MODELE d'ETAT à COEFFICIENT DE RAPPEL ALEATOIRE
FILTRAGE PAR DENSITES APPROCHEES**

DENIS DE BRUCQ, VIRGINIE RUIZ

Université de Rouen - La3I/ LACIS-ITEPEA
UFR des Sciences et Techniques 76 821 Mont Saint Aignan Cedex
TdSI-GDR 134 CNRS

RÉSUMÉ

Le filtrage de Kalman-Bucy s'applique au modèle d'état comprenant des équations linéaires, bruitées, décrivant l'évolution de l'état et des équations, linéaires, bruitées d'observations. Ce filtrage consiste à calculer, façon récursive, la loi de probabilité, a posteriori, de l'état au vue de l'observation actuelle et des observations passées. Le filtrage par densités approchées permet de traiter des équations d'état non linéaires ou à bruits non Gaussiens.

Pour un coefficient de rappel aléatoire, cas typique d'une situation de changements de modèles, l'article introduit une famille de lois de probabilités, paramétrées, bimodales servant, par ajustement des paramètres, à approcher les lois a posteriori de l'état aux divers instants. Les paramètres sont calculés récursivement, lors des mises à jour et des prédictions.

I. INTRODUCTION

En raison de l'originalité de la méthode, nous nous appuyerons sur un modèle d'état explicite, simple, non linéaire scalaire.

L'évolution de l'état X , au cours du temps, est approchée par une équation linéarisée à laquelle, nous rajoutons un bruit blanc W gaussien. Le coefficient a de rappel est un scalaire. Afin d'avoir un système stable, $|1 - a| < 1$. Un biais constant b , est ajouté par ailleurs.

L'observation $Y(t)$ provient linéairement de l'état $X(t)$ auquel s'ajoute un bruit blanc, gaussien, noté $V(t)$ $t \in N$.

Pour ces raisons, le modèle d'état s'écrit:

$$X(t + 1) = X(t) - aX(t) + b + W(t)$$

$$Y(t) = X(t) + V(t)$$

Si de plus, l'état $X(0)$ est gaussien, alors le filtrage de Kalman-Bucy fournit la moyenne et la variance de la loi gaussienne de l'état $X(t)$, lorsque l'observation $y(1 \rightarrow t) \triangleq (y(1), \dots, y(t))$ est connue.

Dans certaines situations physiques relevant de changements de modèles aléatoires, les lois mises en jeu ne sont plus unimodales. Ces situations ne peuvent pas être traitées par le filtrage de Kalman étendu: les lois gaussiennes étant unimodales.

A titre d'exemple, supposons que le coefficient a ne soit pas constant, mais constitue une chaîne de Markov à nombre fini, m , d'états $A(t)$ $t \in N$. En raison du produit $A(t) X(t)$,

ABSTRACT

The Kalman Filtering applies to state models with noisy linear equations which describe the state evolution, and with noisy linear equations of observations. This filtering recursively computes the a posteriori state law given the present and past observations. The filtering by approximated densities allows to process either nonlinear state equations or equations with non-Gaussian noises.

For a random dynamical coefficient, typical situation of models abrupt changes, this paper introduces bi-modal parametered laws of probability which are used to approach the a posteriori state laws, at any time, by adapting parameters. These are recursively computed at each updating and prediction.

le caractère gaussien de la loi de l'état, est perdu! Soit le modèle

A(t) chaîne de markov

$$X(t + 1) = X(t) - A(t).X(t) + b + W(t)$$

$$Y(t) = X(t) + V(t)$$

Partant de $X(0)$ gaussien, alors $X(t)$ est un mélange de m^t gaussiennes dont le nombre croit avec le temps t .

L'article présente, la méthodologie du filtrage par densités approchées qui consiste à développer le logarithme des densités de probabilité, a priori, et, a posteriori, de l'état X sur des fonctions $\varphi_1, \varphi_2, \dots, \varphi_k$.

Pour le filtrage de Kalman étendu, seuls les résultats obtenus en simulation sont présentés, pour que les particularités du filtrage par densités approchées, soient mises en évidence.

Nous pouvons aisément, nous mettre dans une situation où la loi de l'état X est multimodale et ne peut donc pas être approchée par une gaussienne qui est unimodale. Ceci correspond aux cas où à chaque instant t , un changement de modèle est plausible.

Pour fixer les idées, supposons que $1 - A(t)$ prenne l'une des valeurs $1 - a_1 = 0.25$ et $1 - a_2 = 0.75$, de façon aléatoire. Nous simulerons un processus markovien avec une probabilité ρ de changement d'état.

L'article met en oeuvre sur cet exemple, le filtrage par den-



sités approchées en utilisant un mélange de deux lois gaussiennes, ce qui introduit cinq paramètres à suivre récursivement.

II. METHODOLOGIE

Soit le modèle non linéaire décrit par les équations bruitées suivantes: $\forall t \in N$,

$$X(t+1) = g(X(t), W(t)) \quad (II.1)$$

$$Y(t) = h(X(t)) + V(t) \quad (II.2)$$

où $W(t)$, $V(t)$ $t \in N$, sont des bruits blancs indépendants de loi quelconque.

Le filtrage par densités approchées consiste à introduire des contraintes linéaires α_i , sur la loi de la variable $X(t+1)$ lorsque l'observation $Y(1 \rightarrow t)$ est connue.

$$\begin{aligned} \alpha_i &= E[\varphi_i(X(t+1)/Y(1 \rightarrow t))] \\ &= \int \varphi_i(x) f_{X(t+1)/Y(1 \rightarrow t)}(x/Y(1 \rightarrow t)) dx \end{aligned}$$

L'équation d'état (1) est utilisée, avec g non linéaire quelconque pour calculer les contraintes α_i . A l'aide du théorème de transfert des probabilités, la loi connue au mieux de l'état $X(t)$, à la date t , est utilisée.

$$\begin{aligned} \alpha_i &= E[\varphi_i(g(X(t), W(t)))] \\ &= \int \int \varphi_i(g(x, w)) f_{X(t)/Y}(x/Y) f_W(w) dx dw \end{aligned}$$

Les fonctions φ_i peuvent être des indicatrices d'ensembles, des fonctions puissances etc.. La loi approchée pour $X(t+1)$ est alors définie comme l'unique loi qui rend maximum l'entropie sous les contraintes α_i ; il ne s'agit pas de rajouter une information autre que celle retenue par les contraintes. La loi déterminée, est une loi de type exponentiel. L'étude directe de l'équation d'état (1) en l'absence d'observation montre que la loi, pour t grand, de l'état peut-être multimodale. La probabilité est concentrée autour des zéros de l'équation $x=g(x,0)$.

Les logarithmes des densités sont développés linéairement sur des fonctions φ_i , à choisir astucieusement. Les coefficients λ_i des développements linéaires varient dans la famille de densités et sont recalculés pour chaque mise à jour et pour chaque prédiction.

A chaque date t , l'équation d'observation (2) est utilisée avec V bruit blanc d'observation de loi de probabilité de type exponentiel. La formule de Bayes à partir d'un état $X(t)$ et d'un bruit $V(t)$ de type exponentiel permet le calcul de la loi de $X(t)$ conditionnellement à l'observation $Y(1 \rightarrow t)$.

La mise en oeuvre sur l'exemple simple de l'article, de cette méthodologie, permet de rendre concrètes les étapes qui viennent d'être décrites.

III. MODELE NON LINEAIRE

Une approche phénoménologique est essentielle. Partons d'un processus gaussien dont l'évolution est décrite par un

modèle d'état bruité, à coefficient a_j , ($j=1,2$) de rappel et comportant un biais b

$$X(t+1) = X(t) - a_j X(t) + b + W(t) \quad (III.1)$$

$$Y(t) = X(t) + V(t) \quad (III.2)$$

où les bruits W et V , gaussiens, centrés, de variance Q et R sont, par hypothèse, indépendants.

A la limite, la loi asymptotique de l'état X est gaussienne. Pour chacun des coefficients a_1 et a_2 , les lois asymptotiques sont gaussiennes de variances σ_1^2, σ_2^2 et de moyennes m_1, m_2 différentes de zéro, en raison du biais b .

Supposons maintenant que le coefficient $A(t)$ de rappel, constant au cours du temps, soit inconnu. Le modèle devient:

$$A(t+1) = A(t) \quad (III.3)$$

$$X(t+1) = X(t) - A(t).X(t) + b + W(t) \quad (III.4)$$

$$Y(t) = X(t) + V(t) \quad (III.5)$$

Celui-ci n'est plus linéaire. Le filtrage de Kalman étendu effectue la linéarisation des équations d'état. Ce filtrage se limite à introduire des approximations gaussiennes des lois de probabilité. En simulation, prenons $A(0)$ aléatoire entre les deux valeurs a_1 et a_2 . Au cours du temps, quelque soit la distribution $p.\delta_{a_1} + (1-p).\delta_{a_2}$ avec $p \in [0, 1]$, a priori, sur $A(0)$, la loi de $X(t)$ approche l'une des deux lois gaussiennes asymptotiques et l'approximation par une gaussienne fournie par le filtrage de Kalman étendu, est convenable.

Dans une troisième étape, prenons un rappel ($A(t), t \in N$), aléatoire, non constant au cours du temps. En simulation, nous supposons que A constitue une chaîne de Markov, symétrique, à deux états a_1 et a_2 , pour laquelle nous donnons la probabilité ρ de transition:

$$P[A(t+1) = a_j / A(t) = a_j] = 1 - \rho \quad (III.6)$$

$$P[A(t+1) = a_i / A(t) = a_j] = \rho$$

Le modèle d'état (4) et (5) comprend en plus, l'évolution aléatoire (6) du rappel A .

La variable aléatoire $A(t)$ prend les valeurs a_1 et a_2 avec des probabilités $p^1(t)$ et $p^2(t) = 1 - p^1(t)$, évaluées en simulation.

Les deux valeurs possibles de $A(t)$ conduisent à des lois de probabilités limites différentes et par conséquent une corrélation est à attendre, à chaque instant, entre $X(t)$ et $A(t)$.

La loi de ($A(t), X(t)$) conditionnellement à $y(1 \rightarrow t)$ dépend de la valeur numérique actuelle $Y(t) = y(t)$, en raison de l'équation (5) même si $A(t)$ n'y est pas explicite.

L'étape de prédiction consiste à partir des lois à la date t et des équations (6) et (4) d'évolution, à prédire la loi de ($A(t+1), X(t+1)$).

La méthodologie consiste à faire le choix d'une famille de lois de type exponentiel, pour approcher les lois, a priori, et, a posteriori, de l'état ($A(t), X(t)$) aux diverses dates.



IV. FILTRE FINI A CINQ PARAMETRES

Les lois gaussiennes sont de type exponentiel, en effet:
 $\forall x \in N,$

$$f(x) = \frac{1}{\sqrt{2\pi\sigma}} \exp\left(-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2\right) \quad \text{ou} \quad (IV.1)$$

$$f(x) = \exp(\lambda_0 + \lambda_1 x + \lambda_2 x^2) \quad \text{avec} \quad \int f(x) dx = 1$$

La base de fonctions φ vaut $\varphi_0(x)=1(x), \varphi_1(x)=x$
 $\varphi_2(x)=x^2$, avec des paramètres $\lambda_0, \lambda_1, \lambda_2$ qui s'obtiennent
 bijectivement, dans le cas gaussien, à partir de la moyenne
 m et de la variance σ^2 [cf. de DRUCQ & al 1991].

La loi de probabilité de A(t) est une mesure: $\forall a \in N,$

$$P[A(t) = a_1] \cdot \delta_{a_1}(a) + P[A(t) = a_2] \cdot \delta_{a_2}(a) \quad (IV.2)$$

$$= p^1 \cdot \delta_{a_1}(a) + p^2 \cdot \delta_{a_2}(a)$$

Introduisons la probabilité de l'état (A(t), X(t)).

Pour A(t)= a_j , prenons pour densité de probabilité de X(t):

$$f^j(x) = e^{\lambda_0^j + \lambda_1^j x + \lambda_2^j x^2}, \quad t q \int f^j(x) dx = 1 \quad (IV.3)$$

Finalement, la loi du couple (A(t), X(t)), mélange de produits
 de Diracs et de gaussiennes, est caractérisée par :
 $\forall a, x \in \mathbb{R},$

$$f(a, x) = \sum_{j=1,2} p^j \cdot \delta_{a_j} \cdot f^j(x) \quad (IV.4)$$

Nous sommes en présence d'une loi dégénérée de type exponentiel,
 en raison des deux Diracs; il suffit de remplacer chacun des Diracs
 par des indicatrices d'intervalles petits autour de a_1 et a_2 pour obtenir
 l'écriture canonique (1) à l'aide de l'exponentielle [cf. RUIZ 93].

Cette loi de probabilité bimodale pour X(t) constitue une approximation
 puisque chacune des séquences A(0), ..., A(t-1) conduit à une moyenne et à une variance
 différente donc à une loi différente. A partir d'une loi gaussienne pour X(0),
 il faudrait considérer un mélange de 2^t lois gaussiennes pour écrire la loi exacte
 de X(t).

Mise à jour:

A l'aide de l'équation (III.5) et avec des notations évidentes, la formule de Bayes
 fournit, la loi du couple (A(t), X(t)) à la date t conditionnellement à l'observation
 $y(1 \rightarrow t-1)$ et y(t):

$$f_{(A,X)/Y}(a, x/y) = C \cdot f_{A,X}(a, x) \cdot f_V(y-x) \quad (IV.5)$$

$$\text{avec } C^{-1} = \int \int f_{A,X}(\alpha, \xi) \cdot f_V(y-\xi) d\alpha d\xi \quad (IV.6)$$

Le conditionnement maintient le caractère de type exponentiel, de la loi a posteriori.

Ainsi, pour A(t)= a_j , (j=1,2) les lois de densité $f_{X/Y}^j$ de

X(t) conditionnelle à $y(1 \rightarrow t)$ sont encore de type exponentiel. Nous nous limitons, en simulation, à des lois gaussiennes. De l'expression (5), il vient:

$$f(a, x/Y) = C.$$

$$\sum_{j=1,2} p^j \delta_{a_j}(a) e^{\lambda_0^j + \lambda_1^j x + \lambda_2^j x^2} \cdot e^{l_0 + l_1(y-x) + l_2(y-x)^2}$$

où les paramètres l_0, l_1, l_2 de la densité du bruit d'observation V, s'obtiennent à partir de sa moyenne m=0 et de sa variance R.

En additionnant les coefficients des fonctions 1, x et x^2 , nous obtenons, pour A(t)= a_j , (j=1,2) les coefficients λ_i^j , (i=0,1,2) des densité $f_{X/Y}^j(x/y)$ de X(t), conditionnellement à $y(1 \rightarrow t)$; la normalisation est assurée, ici, à l'aide de deux coefficients α_0^j . D'où la densité:

$$f(a, x/y) = C \cdot \sum_{j=1,2} p^j \alpha_0^j \delta_{a_j} \cdot \frac{1}{\alpha_0^j} e^{\lambda_0^j + \lambda_1^j x + \lambda_2^j x^2} \quad (IV.7)$$

Après intégration en x (7), les probabilités de la loi de A(t) actualisée conditionnellement à l'observation sont:
 pour $i=1,2 \quad i \neq j$

$$p_{i/t}^j = P[A(t) = a_j] = \frac{p_{i/t-1}^j \alpha_0^j}{p_{i/t-1}^j \alpha_0^j + p_{i/t-1}^i \alpha_0^i} \quad (IV.8)$$

La loi du couple (A, X), à la date t, conditionnellement à l'observation $y(1 \rightarrow t)$ connue, est caractérisée par (4) avec les p^j et les λ_i^j , (i=0,1,2) actualisés.

Prédiction:

La simulation choisie, qui limite les valeurs possibles de A à a_1 et à a_2 , maintient, à la date t+1, le caractère mélange de produits de Diracs par des gaussiennes.

Une évolution de A peut avoir lieu, suivant l'équation (III.6) et les probabilités p^j de la loi de A(t+1), sont définies par:
 $i=1,2 \quad i \neq j$

$$P[A(t+1) = a_j] = P[A(t) = a_j](1-\rho) \quad (IV.9)$$

$$+ P[A(t) = a_i]\rho$$

Suivant l'équation d'équation d'état (III.4), les calculs de prédiction sur X, se font avec A(t), sans changement. Il s'agit de calculer, sous chacune des hypothèses A(t)= a_j , les contraintes linéaires $\alpha_0^j, \alpha_1^j, \alpha_2^j$ associées aux fonctions 1, x, x^2 . Le transfert des probabilités de la date t+1 à la date t est effectué avec (III.4), et pour A(t)= a_j , les contraintes valent:

$$\alpha_0^j = \int \int p^j e^{\lambda_0^j + \lambda_1^j x + \lambda_2^j x^2} f_W(w) dx dw \quad (IV.10)$$

$$\alpha_1^j = \int \int (x - a_j x + b + w) p^j e^{\lambda_0^j + \lambda_1^j x + \lambda_2^j x^2} f_W(w) dx dw$$

$$= \alpha_0^j ((1 - a_j) m_{t/t}^j + b)$$



$$\alpha_2^j = \int \int (x - a_j x + b + w)^2 p^j e^{\lambda_0^j + \lambda_1^j x + \lambda_2^j x^2} f_W(w) dx dw$$

$$= \alpha_0^j (b^2 + Q + (1 - a_j)^2 (m_{t/t}^{j2} + \sigma_{t/t}^{j2}) + 2b(1 - a_j) m_{t/t}^j)$$

Ces contraintes déterminent les paramètres de Lagrange λ_i^j , ($i=0,1,2$) des densités approchées de probabilité $f_{X|Y(1 \rightarrow t)}^j$ qui réalise le maximum d'entropie. Et la loi du couple (A,X) , à la date $t+1$ conditionnellement à l'observation $y(1 \rightarrow t)$, s'écrit à nouveau avec (4).

V. SIMULATION

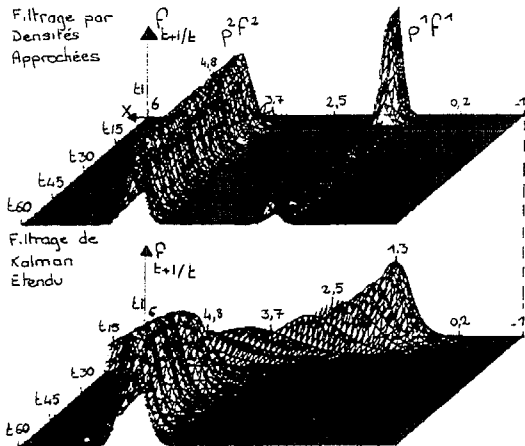
La loi de $X(0)$ est la loi stationnaire provenant de (III.4), $x(0) = 8/3$. Les coefficients sont pris tels que $1 - a_1 = 1/4$ et $1 - a_2 = 3/4$.

Nous pouvons, ici, décrire les phénomènes, à l'aide du vocabulaire des tests d'hypothèse et l'observation $Y(t+1)$ est de loi dépendant de l'hypothèse.

Changements rares de dynamiques:

Les deux filtrages sont rapidement très voisins dans les périodes longues de stationnarité. La loi de X se stabilise sur l'une des lois asymptotiques en dehors de durées assez courtes suivant immédiatement les changements de dynamique. Nous sommes en situation d'évolution linéaire avec ruptures de dynamique.

Prenons tout d'abord la situation d'un coefficient $\rho=0,05$. Le filtre par densités approchées sépare les deux dynamiques possibles.



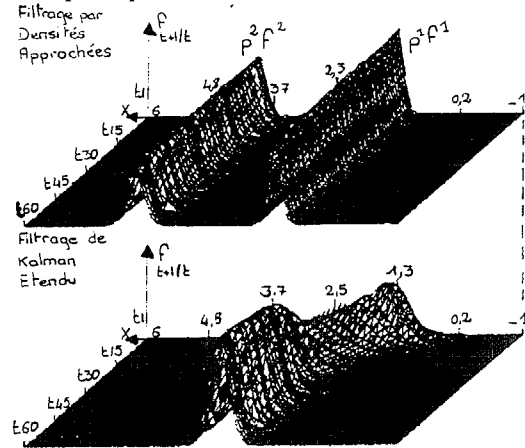
Le filtrage de Kalman étendu met de nombreuses itérations pour faire apparaître la nouvelle valeur de la moyenne m tandis que le filtrage par densités approchées en trois itérations, explicite le changement de dynamique sur les valeurs de p^1 et p^2 .

Changements fréquents de dynamiques:

Prenons maintenant $\rho=0,5$. La connaissance sur l'évolution devient très partielle et le filtre de Kalman étendu traduit cette perte d'information en fournissant une moyenne intermédiaire entre les deux moyennes des situations asymptotiques et une variance agrandie. L'aspect bimodal est évidemment masqué par la modélisation!

Tout au contraire, le filtrage par densités approchées permet le suivi des deux lois de probabilité correspondant au deux cas $a_1=3/4$ et $a_2=1/4$. Ces deux lois sont suivies régulièrement au cours du temps et permettent le calcul par maximum de vraisemblance de la meilleure prédiction

x de l'état du système. De même, les changements de modèles proviennent lorsque y est donnée, du suivi des deux probabilités p^1 et p^2 .



VI. CONCLUSION

Le filtrage de Kalman étendu conduit à d'excellents résultats lorsque les lois inconnues de probabilité de l'état X peuvent être approchées par des lois gaussiennes.

On s'écarte évidemment de ces conditions lorsqu'on étudie les phénomènes de ruptures. Le filtrage par densités approchées constitue une avancée technique importante. Il est clair que le suivi récursif de cinq caractéristiques de distributions statistiques conduit à une meilleure approximation que le suivi récursif de deux caractéristiques. L'exemple choisi en simulation utilise un mélange de deux lois gaussiennes. La méthodologie s'étend sans difficulté aux cas de mélanges non gaussiens [cf. RUIZ 93].

La modélisation doit permettre le transfert de la loi inconnue de probabilité de l'état, à la date $t+1$, vers celle, connue au mieux, de l'état et du bruit, à la date t , afin de permettre le calcul des contraintes. Le filtrage par densités approchées, s'applique à des situations plus nombreuses que le filtrage de Kalman étendu. Les bruits peuvent être quelconque. La connaissance d'un moment d'ordre supérieur à deux est une contrainte linéaire et conduit par maximum d'entropie, à une loi de type exponentiel. L'évolution de l'état X n'a plus besoin d'être décrit par une équation d'état! Il faut cependant avoir une procédure de prédiction de la loi de l'état à la date $t+1$, ici, Markovienne.

BIBLIOGRAPHIE

- I. CSISZAR, Sanov Property, Generalized L-Projection and a conditional limit theorem, The annals of Probability, Vol. 12, No 3, pp 768-793, 1984.
- D. de BRUCQ, E. BUSVELLE, P. COURTELLEMONT, V. RUIZ, Evolution d'état non-linéaire et filtrage approché par maximum d'entropie, GRETSI, pp 37-40, 1991.
- D. de BRUCQ, Extension du Filtre de Kalman, Approximation quadratique, Sém. de Math., Université de Rouen, Compte Rendu de séances 1990-1991.
- F. GAMBOA, Méthode du maximum d'entropie sur la moyenne et applications, Thèse Paris Sud ORSAY, 1990.
- G. JUMARIE, Non linear filtering, A weighted mean squares approach and a bayesian one via the maximum entropy principle, Signal Processing 21, pp 323-328, 1990.
- V. RUIZ, Estimation et prédiction d'un système évoluant de façon non linéaire, Filtrage par Densités Approchées, Thèse Université de ROUEN, 1993.