Filtrage non linéaire en temps discret par convolution de particules

Vivien Rossi — Jean-Pierre Vila

Laboratoire d'Analyse des système et Biométrie
UMR INRA/ENSAM
2, place Pierre Viala
34060 Montpellier, France
{rossiv,vila}@ensam.inra.fr

RÉSUMÉ. *Nous présentons une nouvelle méthode pour traiter les problèmes de filtrage non linéaire en temps dicret. Notre approche s'appuie sur l'estimation de densité par noyaux de convolution et sur le principe des filtres particulaires. Plus précisément, l'approche particulaire classique estime le filtre optimal à l'aide de particules générées suivant le modèle du système dynamique étudié, pondérées suivant la vraisemblance des données. La qualité de l'approximation est liée à ces pondérations, calculées avec la densité du bruit associé aux observations. Nous proposons d'effectuer la pondération des particules avec des noyaux de convolution. En plus d'autoriser la prise en compte de systèmes dynamiques plus généraux, certains problèmes rencontrés par les filtres particulaires usuels peuvent être ainsi contournés. Nous explicitons la mise en oeuvre de cette méthode et les résultats théoriques qui assurent la convergence de ce nouveau filtre vers le filtre optimal ideal. Cette méthode affiche de bonnes performances face aux méthodes actuellements employées. Nous présentons les résultats pour différents systèmes.*

ABSTRACT. *This paper presents a new method to deal with nonlinear filtering problems in discrete time. Our approach is based on density estimation with convolution kernels and on the principle of particles filters. More precisely, the classical particle approach estimates the optimal filter with particles generated according to the dynamical system model of interest, weighted by the likelihood of observed data. The approximation obtained is directly dependent on these weightings. These ones strongly depend on the observation noise density. We propose to weight the particles with convolution kernels. In addition to allowing to take into account more general dynamical systems, some problems met by the usual particle filters can thus be circumvented. The implementation of this method and the theoretical results which ensure its consistency are developed. To illustrate the good performances in practice of this approach, results of comparison, on various dynamical systems, with the most currently used methods, are presented.*

MOTS-CLÉS : *filtrage non linéaire, filtres particulaires, estimateur à noyaux*

KEYWORDS: *nonlinear filtering, particle filters, kernel estimate*

1. Introduction

Tous les phénomènes physiques, chimiques ou biologiques pouvant être modélisés par un système de la forme (1) relèvent de la théorie du filtrage en temps discret.

$$\begin{cases} x_t = f_t(x_{t-1}, \varepsilon_t) \\ y_t = h_t(x_t, \eta_t) \end{cases} \qquad [1]$$

Description des variables : les $x_t \in \mathbb{R}^s$ sont les variables d'état non observées, les $y_t \in \mathbb{R}^d$ sont les variables d'observation, les ε_t sont les bruits blancs du modèle d'état et les η_t sont les bruits blancs du modèle d'observation.
Les bruits blancs sont indépendants des autres variables, ils servent à modéliser les aléas du système. L'objectif est d'estimer les variables d'états non observées, x_t connaissant $y^t = y_1, \cdots, y_t$ les valeurs des variables d'observations jusqu'à l'instant t.
Dans le cas où les fonctions f_t et h_t sont linéaires et les bruits blancs sont additifs et gaussiens le problème a été résolu par Kalman[4] en 1960.
Pour le cas général où les fonctions f_t et h_t sont non linéaires le problème reste ouvert. Le filtre de Kalman étendu (EKF), une généralisation empirique par linéarisation itérative du filtre de Kalman, est la seule méthode utilisée en pratique. Son comportement étant imparfait, la recherche sur ce problème est très active. Haykin([3]), Yee([7] ou Lo ([5]) proposent des méthodes basées sur les réseaux de neurones. Leurs résultats sont intéressants mais l'apprentissage des réseaux nécessite la disponibilité d'un grand échantillon $\{(x_t^i, y_t^i)_{t=1,\cdots,T}\}_{i=1,\cdots,n}$ alors que de part la nature du problème les x_t ne sont pas observés.
L'approche par filtre particulaire utilise uniquement la connaissance du système (1), elle respecte ainsi l'esprit du problème. De plus, elle s'appuie sur de nombreux résultats théoriques (cf. Del Moral & al. [2]). Mais elle peut présenter certains défauts de comportement en pratique. Une étude détaillée est réalisée par Oudjane([6]).
Nous proposons une méthode proche des filtre particulaires mais qui s'en distingue par le type de pondération utilisée et par son allègement des hypothèses sur le modèle.

2. Filtre par noyau de convolution pour un système dynamique quelconque

Avant de continuer précisons les objets supposés connus, à chaque instant t, dans le système S (1) : π_0 la distribution de probabilité de l'état initial x_0, f_t la fonction d'évolution du modèle d'état, h_t la fonction d'observation, $\mathcal{L}_{\varepsilon,t}$ la loi de probabilité du bruit sur le modèle d'état ε_t et $\mathcal{L}_{\eta,t}$ la loi de probabilité du bruit d'observation η_t.
Nous utilisons les notations employées pour les filtres particulaires : π_t désigne la mesure de probabilité de $x_t|y^t$ c'est à dire le filtre optimal. Pour notre approche nous supposons que cette mesure admet une densité que l'on note $p(x_t|y^t)$. Cette densité conditionnelle est définie par $p(x_t|y^t) = p_{XY}(x_t, y^t)/p_Y(y^t)$.
Nous estimons les densités p_{XY} et p_Y par la méthode des noyaux de convolution. Ce type d'estimateurs nécessitent beaucoup d'observations pour être de bonne qualité. En général, en filtrage on n'en possède pas mais on a suffisamment d'information sur le système (1) pour en générer. Soient $z_t = (x_t, y^t)$, et pour tout t soient μ_t et ν_t deux

mesures telles que $z_t \sim \mu_t$ et $y^t \sim \nu_t$
A l'aide d'un algorithme simple, on peut simuler n réalisations de z_t, $\tilde{z}_t^i = (\tilde{x}_t^i, \tilde{y}_i^t)$ qui sont des représentants de μ_t. Nous construisons alors les estimations empiriques de μ_t et ν_t qu'il suffit de convoluer avec des noyaux K^1, K^2 pour obtenir une estimation de $p(x_t|y^t)$:

$$p^n(x_t|y^t) = \frac{\sum_{i=1}^n K_h^1(z_t, \tilde{z}_t^i)}{\sum_{i=1}^n K_h^2(y^t, \tilde{y}_i^t)}$$

Donnons tout d'abord un résultat de convergence ponctuelle de cette estimateur. La démonstration s'appuie sur des résultats de Youndjé([8]) et Bosq([1]) :

Théorème 1 (convergence en moyenne quadratique) *Si les noyaux K^1 et K^2 sont de Parzen-Rosenblatt, si p_Y est positive et continue au point y^t et si $p(x_t|y^t)$ est bornée alors*

$$\left.\begin{array}{l} \lim\limits_{n\to\infty} h_n = 0 \\ \lim\limits_{n\to\infty} nh_n^{td+s} = \infty \end{array}\right\} \implies \lim_{n\to\infty} E[p^n(x_t|y^t) - p(x_t|y^t)]^2 = 0$$

Il est aussi possible de montrer des résultats de convergences uniforme :

Théorème 2 (convergence L_1 intégrée)

$$\left.\begin{array}{l} \lim\limits_{n\to\infty} h_n = 0 \\ \lim\limits_{n\to\infty} nh_n^{td+s} = \infty \end{array}\right\} \implies \lim_{n\to\infty} E\Big[\int |p^n(x_t|y^t) - p(x_t|y^t)| \mathrm{d}x_t\Big] \to 0$$

Remarque : En suivant la même démarche on peut construire un estimateur convergent de $E[x_t|y^t]$, $\widehat{x}_t = \frac{\sum_{i=1}^n K_h^2(y^t, \tilde{y}_i^t)}{\sum_{i=1}^n K_h^2(y^t, \tilde{y}_i^t)}$. Nous l'utilisons pour les applications.
Pour une estimation de qualité, plus la dimension des observations est grande plus n doit être grand. Or, la dimension des observations augmente avec le temps. Pour contourner ce problème, nous avons développé plusieurs variantes non détaillées ici. Nous présentons uniquement les résultats obtenus par certaines de ces variantes dans la partie suivante.

3. Applications : Comparaison avec d'autres filtres

Les filtres comparés sont pour les nôtres deux variantes du filtre général présenté ci-dessus, un filtre à mémoire limitée noté LMF, un filtre à oubli exponentiel du passé noté RBF et pour les autres, l' EKF le filtre de Kalman étendu, l'IEKF une version robustifiée de l'EKF, le RBN de Yee([7]) un réseau à bases radiales et le NFFN de Lo([5]) un réseaux récurrent. Nous utilisons la même mesure de qualité que Lo et Yee pour comparer les filtres : la RMSE (racine carrée de l'erreur quadratique moyenne) calculée sur $N = 500$ trajectoires de taille 120 :

$$\mathrm{RMSE}(t) = (\frac{1}{N}\sum_{i=1}^N \|(x_{t,i} - \widehat{x}_{t,i}\|^2)^{1/2} \quad t = 1, \cdots, 120.$$

Considérons les deux systèmes dynamiques suivant :

$$S_1 = \begin{cases} x_{t+1} = & 1.1\exp(-2x_t^2) - 1 + \frac{\varepsilon_t}{2} \\ y_t = & x_t^3 + 0.1 * \eta_t \end{cases} \qquad S_2 = \begin{cases} x_{t+1} = & A * x_t + \frac{\varepsilon_t}{5} \\ y_t = & [1 \;\; 1]x_t + 0.1\eta_t \end{cases}$$

avec ε_t, η_t des bruits gaussiens centrés réduits indépendants, $x_0 \sim N(-0.5, 0.1^2)$ pour S_1, $x_0 \sim N(0,\, 0.1^2 I_2)$ pour S_2 et $A = \begin{bmatrix} 0.2 & 0.2 \\ 0.5 & -0.5 \end{bmatrix}$. Le système S_1 est aussi traité par Lo([5]) et Yee([7]). Le système S_2 étant linéaire le filtre de Kalman est optimal, il sert donc de référence. Les moyennes des 120 RMSE pour les différents filtres sont rassemblées dans les tableaux ci-dessous :

système S_1	RMSE
EKF	1.3905
IEKF (Lo)	0.2806
RBN (Yee)	0.2260
LMF	0.2174
RBF	0.2161
NFFN (Lo)	0.2120

système S_2	RMSE
RBF	0.2351
LMF	0.2355
Kalman	0.2321

4. Conclusion

Nos filtres ont globalement un bon comportement, surtout par rapport à l'EKF pour le système 1. De plus, il est important de noter que Lo à utilisé 200000 données pour l'apprentissage de son réseau NFFN, le temps de mise en oeuvre n'est donc pas comparable.

5. Bibliographie

[1] D. Bosq, J-P. Lecoutre, *Théorie de l'estimation fonctionnelle*. Economica, Paris (1987).

[2] P. Del Moral, J. Jacod, P. Protter, "The Monte-Carlo method for filtering with discrete-time observations". Probab Theory Relat. Fields 120, 346-368 (2001).

[3] S. Haykin, P. Yee, E. Derbez, "Optimum Nonlinear Filtering".IEEE Trans. signal process, vol.45,No11 (1997).

[4] R.E. Kalman, "A new approach to linear filtering and prediction theory". Trans. of the ASME, Jour. of Basic Eng.,82,34-45 (1960).

[5] J.T. Lo,"Synthetic Approach to Optimal Filtering".IEEE Trans. neur. net.,vol.5,No5 (1994)

[6] N. Oudjane, "Stabilité et approximations particulaires en filtrage non linéaires, application au pistage". PhD thesis, Université de Rennes I (2000).

[7] P.V. Yee, S. Haykin, *Regularised radial basis function networks, Theory and Application*. Wiley, New York (2001)

[8] E. Youndjé, "Propriétés de convergence de l'estimateur à noyau de la densité conditionnelle". Publication URA CNRS D1378 (1993)