

# Oubli de la condition initiale dans un filtrage non linéaire gaussien et filtre de Kalman étendu (FKE)

Matio ROBINSON<sup>1</sup>, Yvon ANDRIANAHARISON<sup>2</sup>,  
Falimanana RANDIMBINDRAINIBE<sup>3</sup>  
Laboratoire d'Electricité, Signaux et Automatique  
Ecole Supérieure Polytechnique d'Antananarivo

Université d'ANTANANARIVO

BP 1500, Ankatso Antananarivo 101, MADAGASCAR

[1mat\\_robinson2000@yahoo.fr](mailto:mat_robinson2000@yahoo.fr), [2yvonsakura@yahoo.fr](mailto:yvonsakura@yahoo.fr), [3falimanana@mail.ru](mailto:falimanana@mail.ru)

*Résumé* : Dans cet article, on étudie les filtrages non linéaires gaussiens. Le filtre optimal est déterminé à partir d'une théorie probabiliste. Comme on est dans le cas gaussien, on peut aussi appliquer le filtrage de Kalman pour avoir des estimations des paramètres du filtre. Sous certaines hypothèses, cette solution théorique permet d'affirmer qu'on peut oublier la condition initiale, c'est à dire que quelque soit la loi initiale, on obtient toujours la même loi du filtre optimal. La méthode d'approximation de Monte Carlo des intégrales sert de base de simulation pour les calculs numériques en application.

*Mots-clefs* :-Filtrage non linéaire, Filtre de Kalman, théorème de Girsanov, Approximation de Monte Carlo, Chaîne de Markov.

*Abstract* : In this paper, we study the Gaussian nonlinear filtering. The optimal filter is determined from a probabilistic theory. As in the Gaussian case, the Kalman filtering could be applied for the estimation of the parameters of the filter. Under certain hypothesis, the theoretical solution lets say we can forget initial condition, i.e whatever the initial law, we always get the same law of optimal filter. Monte Carlo approximation method of integral is the basis of simulation for numerical calculations in applications.

*Keywords* :-Nonlinear filter, Kalman filter, Girsanov theorem, Monte Carlo approximation, Markov chain.

## 1 Introduction

Le filtrage [3], [4] est une opération qui consiste à estimer l'état d'un système dynamique à partir d'observations partielles et bruitées : on cherche à chaque instant à estimer la vraie valeur de l'état à partir des observations.

Dans le cas markovien [1], [4], une solution théorique est établie et dans le cas linéaire gaussien, le filtre de Kalman est le plus pratique dans cette estimation. Une extension de ce filtrage de Kalman [5], [6], [7] est considérée pour les filtres non linéaires.

Intuitivement, la solution du problème de filtrage dépend de l'état initial. La loi de cet état initial peut être oubliée dans le sens où si on s'est trompé de loi de l'état initial, à un certain moment, le filtre tend toujours vers le filtre optimal [8]; on dit dans ce cas qu'il y a stabilité du filtre.

Dans la poursuite de trajectoires, entre autres, la stabilité du filtre est une propriété importante pour une approximation plus vraisemblable.

Dans cet article, cette stabilité pour les filtres non linéaires gaussiens est introduite ainsi que pour les filtres de Kalman [6], [9] qui sont plus adaptés à ces filtres gaussiens.

Les méthodes d'approximations de Monte Carlo [1][3] permettent de faire des simulations pour vérifier effectivement ces théories.

## 2 Modélisation du problème

On considère une chaîne non observable  $X = (X_n)_{n \in \mathbb{N}}$  à valeurs dans  $\mathbb{R}^d$  régie par une formule récurrente perturbée par un bruit additif.

On dispose d'une suite d'observations  $Y_1, \dots, Y_n$  à valeurs dans  $\mathbb{R}^k$  perturbée par une suite  $V = (V_n)_{n \in \mathbb{N}}$  de bruits d'observations.

On a le modèle suivant pour tout  $n \in \mathbb{N}$

$$\begin{cases} Y_n = g(X_n) + V_n & (1) \\ X_{n+1} = h(X_n) + W_n & (2) \end{cases}$$

$h: \mathbb{R}^d \rightarrow \mathbb{R}^d$  une fonction mesurable de même pour  $g: \mathbb{R}^d \rightarrow \mathbb{R}^k$ .

Le filtre est la loi conditionnelle de  $X_n$ ; sachant  $\mathcal{Y}_n$ ,  $\mathcal{Y}_n$  est la tribu des observations  $Y_1, \dots, Y_n$ , et on note  $\mu_n(dx) = \mathbb{P}_{X_n}(X_n | \mathcal{Y}_n)$  ce filtre.

On établit une relation de récurrence pour le filtre avec un algorithme récursif qui permet le traitement dans l'immédiat avec la méthode classique de décomposition en deux étapes, en loi prédite  $\mu_n^-(dx) = \mathbb{P}_{X_n}(dx | \mathcal{Y}_{n-1})$  puis en loi corrigée  $\mu_n(dx) = \mathbb{P}_{X_n}(dx | \mathcal{Y}_n)$ .  $\mathcal{Y}_{n-1}$  qui est la tribu des observations  $Y_1, \dots, Y_{n-1}$ .

Dans le cas où la chaîne est markovienne [1], une théorie classique de calcul des probabilités permet de donner une solution théorique en établissant les relations de récurrence.

On suppose que l'on utilise pour le filtre une loi initiale de densité  $\bar{p}_0(x)$  alors que la véritable loi initiale est de densité  $p_0(x)$ .

Avec  $\bar{p}_0(x)$ , on met en œuvre un filtre  $\bar{p}_n^-(x)$  de la prédiction, et  $\bar{p}_n(x)$  de la correction, alors que le vrai filtre est  $p_n^-(x)$ , pour la prédiction, et  $p_n(x)$  pour la correction.

Le filtre oublie sa condition initiale [4] si :

$$\lim_{n \rightarrow \infty} -ps \int_{\mathbb{R}^d} |\bar{p}_n(x) - p_n(x)| dx = 0$$

### 2.1 Filtrage non linéaire

On considère le modèle de chaînes de Markov cachées homogènes [3] tel que l'état du système non observé  $X = (X_n)_{n \in \mathbb{N}}$ , dans l'espace mesurable  $(\mathbb{R}^d; \mathcal{B}(\mathbb{R}^d))$  est de loi initiale  $\mu_0$  et de probabilité de transition  $\pi$  telle que pour toute fonction mesurable bornée  $f$

$$\mathbb{E}[f(X_{n+1}) | X_n = y] = \int_{\mathbb{R}^d} f(x) \pi(y; dx)$$

$$\mathbb{E}f(X_0) = \int_{\mathbb{R}^d} f(x) \mu_0(dx)$$

L'observation est un processus  $Y = (Y_n)_{n \geq 1}$  tel que :

$Y_n = g(X_n) + V_n$  avec  $h: \mathbb{R}^d \rightarrow \mathbb{R}^k$ . On montre qu'il existe une mesure de probabilité  $\sigma_n$  telle que sous cette mesure,  $X_n$  est une chaîne de Markov et de probabilité de transition indépendante de  $\mathcal{Y}_n$  et que

$$\mu_n(dx) = \left( \int_{\mathbb{R}^d} \sigma_n(dy) \right)^{-1} \sigma_n(dx)$$

Cette mesure permet d'établir les relations de récurrence dans la prédiction-correction. Notons  $Q$ , la matrice de covariance de  $W_n$ .

Comme

$$\begin{aligned} \mathbb{E}[f(Y_n | X_n = x)] &= \mathbb{E}f(g(x) + V_n) \\ &= \int_{\mathbb{R}^d} f(g(x) + v) f_V(v) dv \end{aligned}$$

La densité de  $Y_n$  sachant  $X_n = x$  est  $f_V(y - g(x))$ . Notons la quantité

$$Z(x, y) = \exp \left[ \langle Q^{-1}g(x), y \rangle - \frac{1}{2} \left| Q^{-\frac{1}{2}}g(x) \right|^2 \right]$$

Pour un échantillon de taille  $n$ ,  $X_1, \dots, X_n$ , considérons la vraisemblance  $L_n = Z(X_1, Y_1) \times \dots \times Z(X_n, Y_n)$ . Pour tout  $n \geq 1$ , on montre aisément que  $E(L_n^{-1}) = 1$ , et d'après le théorème de Radon-Nikodym [4], on définit alors la probabilité  $\bar{P}_n$  sur  $(\Omega, \mathcal{F})$  par  $\frac{d\bar{P}_n}{dP} = L_n^{-1}$  et on construit la mesure aléatoire  $\sigma_n$  telle que

$$\int_{\mathbb{R}^d} f(x) \sigma_n(dx) = \bar{\mathbb{E}}_n[f(X_n) L_n | \mathcal{Y}_n]$$

$$\mathbb{E}[f(X_n) | \mathcal{Y}_n] \cdot \bar{\mathbb{E}}_n[L_n | \mathcal{Y}_n] = \bar{\mathbb{E}}_n[f(X_n) L_n | \mathcal{Y}_n]$$

Et on a la formule de *Kallianpur-Striebel*

$$\mathbb{E}[f(X_n) | \mathcal{Y}_n] = \frac{\bar{\mathbb{E}}_n[f(X_n) L_n | \mathcal{Y}_n]}{\bar{\mathbb{E}}_n[L_n | \mathcal{Y}_n]}$$

En d'autres termes,

$$\mu_n(dx) = \left( \int_{\mathbb{R}^d} \sigma_n(dy) \right)^{-1} \sigma_n(dx)$$

en appliquant le théorème de Girsanov [2]:

### Proposition

Sous  $\bar{\mathbb{P}}_n$ ,  $(X_1, \dots, X_n)$  est une chaîne de Markov de loi initiale  $\mu_0$  et de probabilité de transition  $\pi$  indépendante de la suite  $(Y_1, \dots, Y_n)$ .

Sous  $\bar{\mathbb{P}}_n$ ,  $(X_1, \dots, X_n)$  et  $(Y_1, \dots, Y_n)$  sont donc indépendantes et on peut alors définir sous une loi  $\bar{\mathbb{P}}$ , d'espérance  $\bar{\mathbb{E}}$  de la façon suivante :

$$\begin{aligned} \bar{\mathbb{E}}_n[\phi(X_1, \dots, X_n, Y_1, \dots, Y_n) | Y_1, \dots, Y_n] &= \\ = \bar{\mathbb{E}}[\phi(X_1, \dots, X_n, Y_1, \dots, Y_n)] \end{aligned}$$

et on obtient les relations de récurrence sur les mesures  $\sigma_n$  et  $\mu_n$

$$\sigma_n(dx) = Z(x, Y_n) \int_{\mathbb{R}^d} \pi(x'; dx) \sigma_{n-1}(dx')$$

Avec la loi de  $X_0$ ,  $\sigma_0(dx)$  ou  $\mu_0(dx)$

$$\mu_n(dx) = C_n Z(x, Y_n) \int_{\mathbb{R}^d} \pi(x'; dx) \mu_{n-1}(dx')$$

Où  $C_n$  est une constante de normalisation telle que :

$$C_n^{-1} = \int_{\mathbb{R}^d} Z(x, Y_n) \int_{\mathbb{R}^d} \pi(x'; dx) \mu_{n-1}(dx')$$

permettant d'établir des équations de récurrence pour le filtre optimal. Ce qui entraîne la décomposition en deux étapes classiques :

### Etape de prédiction

$$\mu_n^-(dx) = \int_{\mathbb{R}^d} \pi(x'; dx) \mu_{n-1}(dx')$$

### Etape de correction

$\mu_n(dx) = C_n Z(x, Y_n) \mu_n^-(dx)$  avec

$$Z(x, Y_n) = \exp \left[ \langle Q^{-1}g(x), Y_n \rangle - \frac{1}{2} \left| Q^{-\frac{1}{2}}g(x) \right|^2 \right]$$

Pour cette solution théorique, le bruit d'observation est gaussien. Si de plus le bruit de mesure est aussi gaussien, il est préférable d'appliquer le filtrage de Kalman pour déterminer le filtre optimal. Seules les estimations de sa moyenne  $\hat{X}_n$  et de sa matrice

de covariance  $R_n$  sont nécessaires à la détermination de cette loi.

## 2.2 Filtre de Kalman étendu

Considérons le système non linéaire gaussien [7] tel que:

$$\begin{cases} Y_n = h_n(X_n) + V_n \\ X_{n+1} = f_n(X_n) + g_n(X_n)W_n \end{cases}$$

$V_n$  représente le bruit de mesure, bruit blanc gaussien de matrice de covariance  $Q_n^V$ .

L'état  $X_n$  du système dynamique (2) n'est pas observé et son équation d'évolution est perturbée par un bruit blanc gaussien de matrice de covariance  $Q_n^W$  et, de condition initiale  $X_0$  gaussienne  $\mathcal{N}(\bar{X}_0; Q_0)$ .

Les bruits  $V_n$  et  $W_n$  et la condition initiale  $X_0$  sont mutuellement indépendants.

Le critère d'optimalité est de minimiser la variance de l'erreur d'estimation donc de déterminer la loi conditionnelle de  $X_n$  sachant  $\mathcal{Y}_n$ .

Le filtre de Kalman [5] est adapté pour les filtres linéaires gaussiens. Dans le cas non linéaire, dans l'extension du filtre de Kalman, on linéarise par une approximation fonctionnelle le système.

Soit une suite déterministe  $x_n$  dans  $\mathbb{R}^d$ , solution approchée du système, appelée **trajectoire nominale**. Les fonctions  $f_n$  et  $g_n$  sont supposées dérivables. On linéarise  $f_n$  et  $g_n$  autour de  $x_n$  avec les développements habituels et on fait les approximations :

$$f_n(x) \simeq f_n(x_n) + \vec{\nabla} f_n(x_n)(x - x_n)$$

$$g_n(x) \simeq g_n(x_n)$$

Il en est de même pour  $h_n$  autour de  $\bar{x}_n$ ,

$h_n(x) \simeq h_n(x_n) + \vec{\nabla} h_n(x_n)(x - \bar{x}_n)$ . On obtient le système linéarisé suivant:

$$\begin{cases} Y_n = H_n(X_n - \bar{x}_n) + \bar{h}_n + V_n \\ X_{n+1} = F_n(X_n - \bar{x}_n) + \bar{f}_n + G_n W_n \end{cases}$$

Où  $f_n(\bar{x}_n) = \bar{f}_n, F_n = \vec{\nabla} f_n(\bar{x}_n), G_n = g_n(\bar{x}_n)$

$H_n = \vec{\nabla} h_n(\bar{x}_n)$  et  $h_n(\bar{x}_n) = \bar{h}_n$ .

En notant,  $\hat{X}_n = X_n - \bar{x}_n$ ,  $\hat{f}_n = f_n - \bar{x}_{n+1}$  et  $\hat{W}_n = G_n W_n$ , on a un nouveau système linéaire:

$$\begin{cases} Y_n = H_n \hat{X}_n + \bar{h}_n + V_n \\ X_{n+1} = F_n \hat{X}_n + \bar{f}_n + \hat{W}_n \end{cases}$$

Les bruits  $\hat{W}_n, V_n$  et  $\hat{X}_0 = X_0 - \bar{x}_0$  sont mutuellement indépendants. La condition initiale est gaussienne  $\mathcal{N}(\bar{X}_0 - \bar{x}_0; Q_0)$ .

$\hat{W}_n$ , la perturbation aléatoire est un bruit blanc gaussien de matrice de covariance  $Q_n^W$ .

Dans le cas gaussien, seules la moyenne  $\mathcal{Y}_{n-1}$  et la matrice de covariance  $R_n = \mathbb{E}(\bar{X}_n - \hat{X}_n)(\bar{X}_n - \hat{X}_n)^t$  sont nécessaires à la définition de cette loi.

Pour la prévision,

$$\bar{X}_n^- = \mathbb{E}(\hat{X}_n | \mathcal{Y}_{n-1}) \text{ et}$$

$$R_n^- = \mathbb{E}[(\bar{X}_n - \hat{X}_n^-)(\bar{X}_n - \hat{X}_n^-)^t]$$

On construit le processus d'innovation défini par l'information apportée par  $Y_n$  par rapport aux observations passées  $\mathcal{Y}_{n-1}$ :

$$I_n = Y_n - \hat{Y}_n^- \text{ où } \hat{Y}_n^- = \mathbb{E}(Y_n | \mathcal{Y}_{n-1}).$$

L'innovation est donc:

$$I_n = Y_n - (H_n \hat{Y}_n^- + \bar{h}_n)$$

Le processus  $I_n$  est un processus gaussien à valeurs dans  $\mathbb{R}^k$ , en particulier  $I_n$  est un vecteur aléatoire gaussien de moyenne nulle et de matrice de covariance  $Q_n^I = H_n R_n^- H_n^t + Q_n^V$  indépendant de  $\mathcal{Y}_{n-1}$ . Et on a le théorème:

### Théorème de Kalman-Bucy [9]

Si la matrice de covariance  $Q_n^V$  est inversible pour tout  $n \in \mathbb{N}$ , alors les processus  $\hat{X}$  et  $R_n$  sont définies par les équations suivantes:

Prédiction

$$\begin{cases} \hat{X}_n^- = F_n \hat{X}_{n-1} + \bar{f}_n \\ R_n^- = F_n R_{n-1}^- F_n^t + Q_n^W \end{cases}$$

Correction

$$\begin{cases} \hat{X}_n = \hat{X}_n^- + K_n [Y_n - (H_n \hat{X}_n^- + \bar{h}_n)] \\ R_n = [I - K_n H_n] R_n^- \end{cases}$$

où le gain de Kalman est la matrice

$$K_n = R_n^- H_n^t [H_n R_n^- H_n^t + Q_n^V]^{-1} = R_n^- H_n^t [Q_n^I]^{-1}$$

avec les initialisations:

$$\hat{X}_0^- = \tilde{X}_0 - \bar{x}_0 = \mathbb{E}X_0 - \bar{x}_0, R_0^- = Q_0^X$$

Le gain matriciel de Kalman  $K_n$  est à déterminer de façon à fournir une estimation

$\hat{X}_n$  optimale au sens de l'erreur quadratique moyenne. On définit ainsi l'erreur à posteriori:

$e_n = X_n - \hat{X}_n$  et la matrice de covariance

$p_n = \mathbb{E}[e_n e_n^t]$ . On peut alors définir l'erreur à priori

$e_n^- = X_n^- - \hat{X}_n^-$ , ainsi que la matrice de covariance de l'erreur à priori:

$$p_n^- = \mathbb{E}[e_n^- (e_n^-)^t].$$

La solution théorique et la solution par approximation par filtre de Kalman dépendent de la condition initiale. On peut montrer que, tout en restant dans le cas gaussien, qu'avec une condition initiale erronée, on peut avoir le même résultat qu'avec le véritable filtre. Il y a donc une stabilité du filtre par rapport à la mesure initiale.

### 2.3 Oubli de la condition initiale

On suppose donc que l'on utilise pour le filtre une densité  $\bar{p}_0$  alors que la véritable loi initiale est  $p_0$ . Avec  $\bar{p}_0$ , on met en œuvre un filtre,  $\bar{p}_n^-$  la prédiction et  $\bar{p}_n$ , la correction, alors que le vrai filtre est  $p_n^-$ , la prédiction et  $p_n$  la correction.

Le filtre oublie sa condition initiale si

$$\lim_{n \rightarrow \infty} \text{ps} \left| \int_{\mathbb{R}^d} f(x) p_n(x) dx - \int_{\mathbb{R}^d} f(x) \bar{p}_n(x) dx \right| = 0$$

Pour toute  $f \in C_b(\mathbb{R}^d)$  ou encore,

$$\lim_{n \rightarrow \infty} \text{ps} \int_{\mathbb{R}^d} |\bar{p}_n(x) - p_n(x)| dx = 0$$

Appliquons le cas où  $d = 1$ . Le cas général se déduit assez facilement. Avec le système de filtrage (1) et (2)

$$\begin{cases} Y_n = g(X_n) + V_n & V_n \sim \mathcal{N}(0; \sigma_V) \\ X_{n+1} = h(X_n) + W_n & W_n \sim \mathcal{N}(0; \sigma_W) \end{cases}$$

on a les lois prédites et corrigées admettant les densités:  $p_n^-$  avec  $\mu_n^-(dx) = p_n^-(x)$

et  $p_n$  avec  $\mu_n(dx) = p_n(x)$ .

Et comme  $\pi(x'; dx) = f_W(x - g(x')) dx$ ,

$$p_n^-(x) = \int_{\mathbb{R}^d} f_W(x - g(x')) p_{n-1}^-(x') dx'$$

$$p_n(x) = C_n Z(x, Y_n) \cdot p_n^-(x)$$

La constante de normalisation est  $C_n$  telle que :  $C_n^{-1}$  est égal à

$$\int_{\mathbb{R}^d} \int_{\mathbb{R}^d} Z(y + g(x'), Y_n) f_W(y) p_{n-1}^-(x') dx'$$

$$Z(x, Y_n) = \exp \left[ \frac{h(x) Y_n}{\sigma_V^2} - \frac{h(x)^2}{2\sigma_W^2} \right]$$

Analytiquement,

$$|p_n(x) - \bar{p}_n(x)| = |Z(x, Y_n)| [C_n p_n^- - C_n \bar{p}_n^-]$$

La majoration de cette quantité n'est pas évidente avec les constantes de normalisation.

Il est nécessaire de faire le calcul numériquement à partir d'une méthode de Monte Carlo [1] consistant à approximer numériquement les intégrales sous la forme

$$\int_{\mathbb{R}^d} \varphi(x) p(x) dx \simeq \frac{1}{N} \sum_{i=1}^N \varphi(X_i)$$

$X_i$  sont simulées suivant la loi de densité  $p(x)$ . Cette approximation numérique est justifiée par la loi des grands nombres et le théorème central limite. Avec le principe d'échantillonnage d'importance, il est parfois utile de modifier l'intégrande :

$$\int_{\mathbb{R}^d} \varphi(x) p(x) dx = \int_{\mathbb{R}^d} \varphi(x) \frac{p(x)}{q(x)} q(x) dx$$

où le support de  $p(x)$  est contenu dans celui de  $q(x)$  et les  $X_i$  seront simulés suivant la densité  $q(x)$ . Le choix de  $q(x)$  est uniquement pour une raison pratique non pas pour une réduction de variance d'une meilleure convergence. Avec

cette approximation de Monte Carlo et avec  $N$  assez grand,

$$p_n^-(x) \approx \frac{1}{N} \sum_{i=1}^N f_W(x - g(X_{i,n-1}))$$

$$C_n^{-1}(x) \approx \frac{1}{N} \sum_{i=1}^N Z(W_i + g(X_{i,n-1}), Y_n)$$

$W_i$  sont simulées suivant la loi de densité  $f_W$  et  $X_{i,n-1}$  sont simulées suivant la loi de densité  $p_{n-1}$ .

Dans les simulations, les densités sont définies par des points espacés régulièrement. On définit ainsi leurs fonctions de répartition par des points espacés régulièrement en utilisant la proposition suivante:

Si  $F$  est une fonction de répartition et  $U \sim \mathcal{U}[0,1]$ , alors  $X = F^{-1}(U)$  admet  $F$  comme fonction de répartition.

Simuler une valeur d'une variable aléatoire de fonction de répartition  $F$ , revient à simuler une valeur de la variable aléatoire  $U$  suivant la loi uniforme  $\mathcal{U}[0,1]$ ,

et utiliser  $X = \arg \max[F^{-1}(U)]$ .

On a les applications suivantes:

### 3 Applications

#### 3.1 Filtrés optimaux et trajectoires

**Algo filtre** ( $h, g$  : fonctions,  $N, N_1$  : entiers,  $f_V, f_W, q_0, p_0$  densités  $Y_1, \dots, Y_n$  observations) :  
 $q$  filtre prédit,  $p$  filtre corrigé

**Début**

Initialisation de  $q$  à  $q_0$

Initialisation de  $p$  à  $p_0$

Pour  $n=1$  à  $N$  faire

$$Z(x, Y_n) = \exp \left[ \frac{h(x)Y_n}{\sigma_V^2} - \frac{h(x)^2}{2\sigma_W^2} \right]$$

Pour  $i=1$  à  $N_1$  faire

Simuler  $W_i$  suivant  $p_W$

Simuler  $X_{i,n-1}$  suivant  $p_{n-1}$

$$C_n^{-1} = \sum_{i=1}^{N_1} \frac{1}{N_1} Z(W_i + g(X_{i,n-1}), Y_n)$$

$$p_n^-(x) = \sum_{i=1}^{N_1} \frac{1}{N_1} f_W(x - g(X_{i,n-1}))$$

fin pour  $i$

$$q_n(x) = p_n^-(x)$$

$$p_n(x) = C_n Z(x, Y_n) q_n(x)$$

Fin pour  $n$

$$q(x) = p_n^-(x)$$

$$p(x) = p_n(x)$$

Fin

Considérons  $g(x) = x^2$  et  $h(x) = x$ , on a le système :  $\begin{cases} Y_n = X_n^2 + V_n \\ X_{n+1} = X_n + W_n \end{cases}$

Prenons

$$\sigma_0 = 1, \sigma_w = 2, \sigma_v = 1, N = 100, N_1 = 50.$$

Les bruits de mesure et d'observation suivent  $\mathcal{M}(0,2)$  et  $\mathcal{M}(0,1)$  respectivement.

La trajectoire du signal observé est proportionnel à l'écart entre le signal mesuré et celui de la trajectoire du filtre corrigé (Figure 1).

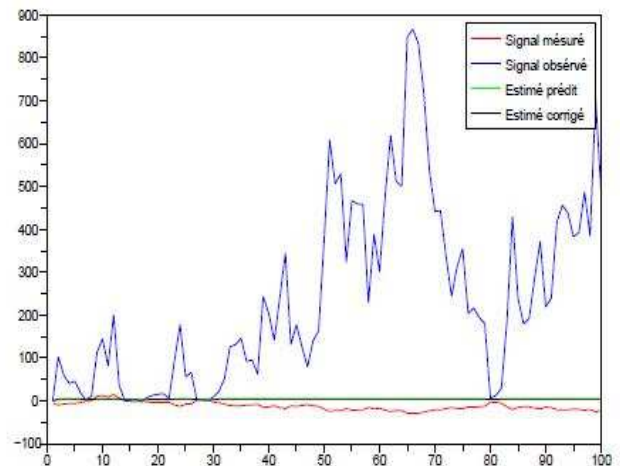


Fig1 : Trajectoires, N=100

Cette remarque est totalement différente pour un autre modèle défini par :

$$\begin{cases} Y_n = 50. \sin(X_n) + 50. \cos(X_n) + V_n \\ X_{n+1} = X_n + W_n \end{cases}$$

Les trajectoires des filtres prédit et corrigé sont représentées à la figure 2.

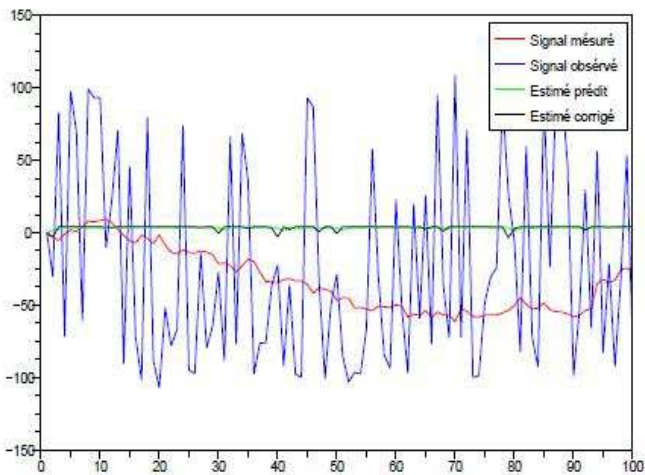


Fig2 : Filtre sinusoïdal

### 3.2 Filtres de Kalman

**Algo Kalman** ( $Q_n^V, Q_n^W$  matrices de covariance des bruits,  $f_n, h_n$ , fonctions

$N$ : entier

$Y_1, \dots, Y_n$  des observations)

$\hat{X}_n^-$  filtre prédit,  $\hat{X}_n$  filtre corrigé

**Début**

Initialisation de  $\hat{X}_0^-$  à  $\hat{X}_0 - \bar{x}_0$

Initialisation de  $R_0^-$  à  $Q_0^-$

Pour  $n=1$  à  $N$  faire

$$F_n = f'_n(\hat{X}_{n-1})$$

$$\tilde{f}_n = f_n(\hat{X}_{n-1})$$

$$G_n = g_n(\hat{X}_{n-1})$$

$$H_n = h'_n(\hat{X}_n^-)$$

$$\hat{X}_n^- = F_n \hat{X}_{n-1} + \tilde{f}_n$$

$$R_n^- = F_n R_{n-1} F_n^t + Q_n^W$$

$$Q_n^I = H_n R_{n-1} H_n^t + Q_n^V$$

$$K_n = R_n^- H_n^t Q_n^I$$

$$\hat{X}_n = \hat{X}_n^- + K_n [Y_n - (H_n \hat{X}_n^- + \tilde{h}_n)]$$

$$R_n = [I - K_n H_n] R_n^-$$

Fin pour  $n$

**Fin**

Les trajectoires par filtrage de Kalman sont représentées à la figure 3.

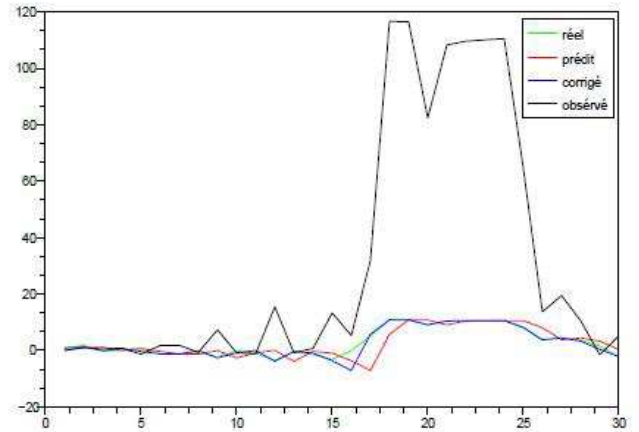


Figure 3 : Trajectoires par filtrage de Kalman

En pratique, l'algorithme du filtre de Kalman est plus adapté pour les filtres optimaux gaussiens. Le développement de  $h_n$  se fait autour de  $\hat{X}_n^-$  et celui de  $f_n$  autour de  $\hat{X}_{n-1}$ .

Les densités gaussiennes sont représentées à la figure 4.

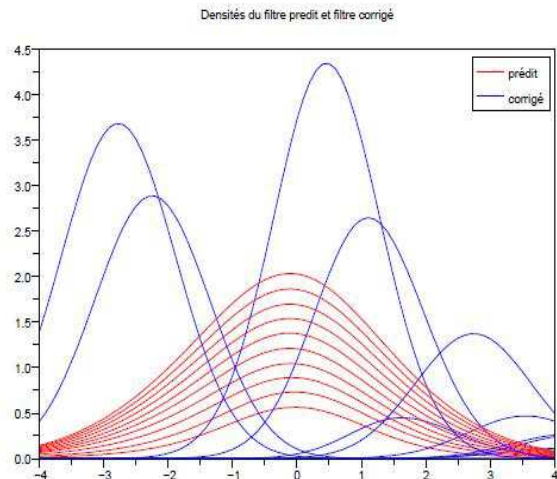


Figure 4 : Filtres prédits et corrigés

On montre à la figure 5 les densités des filtres optimaux de variances différentes.

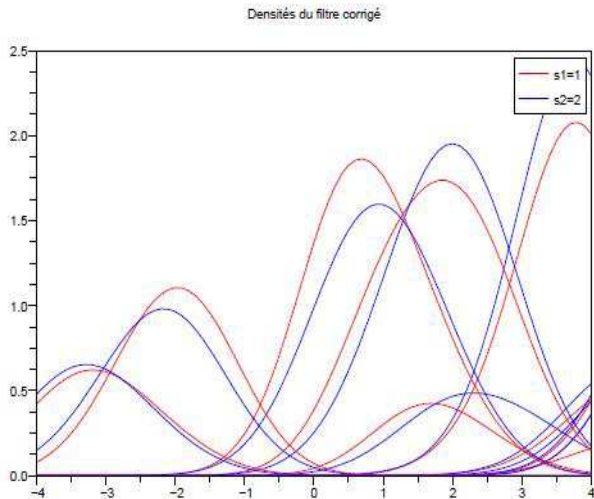


Figure 5 : Densité des filtres optimaux

L'algorithme suivant calcule les intégrales de l'erreur.

**Algo erreur** ( $p_0, \bar{p}_0$ , lois initiales ;  $N_1$  et  $N_2$  entiers,  $g$  densité)

$$\epsilon \text{ erreur}$$

$$\epsilon(1) = 0$$

Pour  $n=1$  à  $N$  faire

Calculer  $p_n(x)$  et  $\bar{p}_n(x)$

$$\alpha_n(x) = |p_n(x) - \bar{p}_n(x)|$$

$$\epsilon(n) = \int_{\mathbb{R}} \alpha_n(x) dx$$

Fin pour  $n$

La simulation numérique a donné les courbes des intégrales de l'erreur à la figure 6.

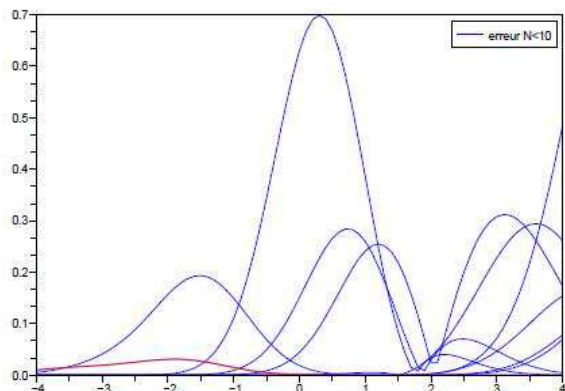


Figure 6 : Intégrales de l'erreur

## 4 Interprétations

Les signaux observés et mesurés, représentés à la figure 1 sont déterminés par le système :

$$\begin{cases} Y_n = X_n^2 + V_n, & V_n \sim \mathcal{N}(0,1) \\ X_{n+1} = X_n + W_n, & W_n \sim \mathcal{N}(0,2) \end{cases}$$

Les filtres prédits et corrigés s'approchent du signal observé. La connaissance de l'équation d'état est fondamentale pour avoir ces filtres.

L'équation du signal observé est modifiée à la figure 2

Les filtres prédits et corrigés sont assez écartés par rapport à la trajectoire du signal réel.

Finalement, la solution théorique permet de

justifier l'existence de la solution, mais le filtre

de Kalman est plus adapté selon le résultat de

simulation représentée à la figure 3.

Dans la solution théorique, on a effectivement des densités gaussiennes (figure 4). Le filtre prédit est resté centré en 0. Par contre avec l'innovation apportée par  $Y_n$ , la moyenne du filtre corrigé est variable. Cette approximation est plus intéressante avec un horizon assez lointain.

En considérant deux lois initiales de différentes variances,  $p_0(x) = \frac{1}{\sigma_0 \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\sigma_0^2}\right)$  et  $\bar{p}_0(x) = \frac{1}{\bar{\sigma}_0 \sqrt{2\pi}} \exp\left(-\frac{x^2}{2\bar{\sigma}_0^2}\right)$ , à l'horizon  $n$ , on a la différence :

$$\epsilon_n = \int_{\mathbb{R}^d} |p_n(x) - \bar{p}_n(x)| dx$$

Les filtres optimaux se différencient de peu (figure 5). Cette différence diminue considéra-



blement et presque sûrement elle va s'annuler, (figure 6). L'intégrale est de plus en plus aplatie. On a bien un oubli de la condition initiale.

Le filtre de Kalman a pour but d'analyser l'état d'un processus inobservé en général markovien, à partir d'un processus observable non nécessairement linéaire sous conditions de bruits gaussiens [7].

Une autre application du filtre de Kalman se trouve en finance à propos des estimations et prévision de la volatilité stochastique et du rapport cours-bénéfice [6].

## 5 Conclusion

Cet article développe deux méthodes différentes dans la détermination du filtre optimal dans le filtrage non linéaire, l'une avec la construction d'une mesure de distribution avec une théorie probabiliste et l'autre avec le filtre de Kalman étendu. Un des intérêts du filtre de Kalman est de pouvoir estimer  $X_0$ , la valeur initiale, dans le problème balistique.

Le résultat montre qu'avec une loi initiale erronée, on obtient les mêmes résultats qu'avec le vrai filtre ; c'est à dire que le filtre obtenu avec la loi initiale erronée est capable d'oublier sa condition initiale. Ce résultat peut être étendu pour les cas des systèmes à bruit d'observation uniforme.

L'hypothèse de la non-linéarité du modèle est très délicate dans le sens où la multi modalité du filtre peut apparaître et dans ce cas le filtre de Kalman est inadapté surtout lorsque le système est fortement non linéaire, le filtre de Kalman étendu peut diverger.

Dans la simulation, le filtre de Kalman [6] est limité à des systèmes à bruits gaussiens.

Le cas général peut être envisagé en appliquant un autre algorithme comme le filtre particulaire qui est plus efficace mais exige plus d'hypothèses.

## Références

- [1] Dominique BAKRI. Martingales et Chaînes de Markov. Laboratoire de Statistique et Probabilités. Université Paul Sabatier, Toulouse 2003.
- [2] M.JEANBLANC, T.SIMON. Eléments de calculs stochastiques. IRBID Sept 2006.
- [3] E.PARDOUX. Processus de Markov et applications: Algorithmes, Réseaux, Génomes et Finance. Cours polycopié. Université de PROVENCE 2006.
- [4] E.PARDOUX. Filtrage non linéaire et équations aux dérivées partielles stochastiques associées. In P.L Hennequin, editor, Ecole d'été de Probabilités de Saint-Flour XIX-1989, volume 1464 of Lecture Notes in Mathematics, pages 67-163. Springer Verlag, Berlin, 1991.
- [5] D.ALAZARD. Introduction au filtrage de Kalman, polycopié de cours, SUPAERO, 2005.
- [6] F-E. RACICOT, R.THEORET. Quelques applications du filtre de Kalman en Finance: Estimation et Prévision de la volatilité Stochastique et du rapport cours-bénéfices. Université de Montréal, 2005.
- [7] F. LAFARGE. Introduction au filtrage de Kalman, théorie du filtre de Kalman discret et applications, cours polycopié ENSG 2009.
- [8] Frederic ROTELLA. Filtrage Optimal. Ecole Nationale d'Ingénieurs de Tarbes 2003.
- [9] Michel VALADIER. Sur le filtre de Kalman-Bucy en temps continu. Revue française d'automatique, informatique, recherche opérationnelle. Mathématique, tome 6, n\_2, p 9-22, 1972.