

Estimation de la variance stochastique multivariée avec un filtre gaussien basé sur la méthode de Laplace

Ivan GORYNIN, Emmanuel MONFRINI, Wojciech PIECZYNSKI

SAMOVAR, Telecom SudParis, CNRS, Université Paris-Saclay
9 rue Charles Fourier, 91000 Evry, France
ivan.gorynin@telecom-sudparis.eu

Résumé – Les modèles de variance stochastique (MVS) traitent la variance de la variable observée comme un processus aléatoire caché. L’estimation de ce processus est un défi car ces modèles ne sont pas linéaires. Notre étude porte sur l’estimation de la variance stochastique en grande dimension avec une nouvelle approche basée sur la méthode de Laplace. Nous considérons un MVS et nous comparons, en termes de temps de calcul et de qualité d’estimation, la méthode proposée avec le filtrage particulaire (FP), qui est une méthode de référence. Nous trouvons qu’en grande dimension, la nouvelle approche fournit un meilleur ratio qualité/durée de traitement que celui du FP. Cette méthode peut donc constituer une alternative au FP pour des problèmes de filtrage statistique en grande dimension dans les modèles non linéaires.

Abstract – In the stochastic variance models, the variance of the observed variable is assumed to be a hidden stochastic process. Estimation of the state process in these models is challenging due to their nonlinearity. Our study is about estimating the stochastic variance in high dimension by using a new filtering technique based on the Laplace approximation. We compare the new method with the particle filter (PF) in terms of the estimation accuracy and processing time in the framework of a stochastic variance model. We find that in high dimension, the new method has a better ratio of accuracy per processing time than the PF. We conclude that this method can be preferred to the PF in the case of high-dimensional filtering in non-linear models.

1 Introduction

Dans les modèles de variance stochastique [1, 2] on considère que les matrices de covariance d’un processus observé sont des réalisations d’un autre processus non observé. Considérons l’équation générale suivante :

$$\mathbb{E}[\mathbf{Y}_n \mathbf{Y}_n^\top | \mathbf{X}_n] - \mathbb{E}[\mathbf{Y}_n | \mathbf{X}_n] \mathbb{E}[\mathbf{Y}_n | \mathbf{X}_n]^\top = \boldsymbol{\Sigma}_n(\mathbf{X}_n),$$

où $n \in \mathbb{N}^*$ est l’indice de temps, $\mathbf{X}_n \in \mathbb{R}^p$ et $\mathbf{Y}_n \in \mathbb{R}^q$ sont des variables cachées et observées respectivement et $\boldsymbol{\Sigma}_n$ est une fonction qui associe à \mathbf{X}_n une matrice de covariance. On se place souvent dans le cas des modèles de Markov cachés, c’est-à-dire que $(\mathbf{X}_n)_{n>0}$ est une chaîne de Markov, les variables $(\mathbf{Y}_n)_{n>0}$ sont indépendantes conditionnellement à $(\mathbf{X}_n)_{n>0}$ et chaque variable de $(\mathbf{Y}_n)_{n>0}$ ne dépend que de la seule variable de $(\mathbf{X}_n)_{n>0}$ ayant l’indice correspondant.

En général, ces modèles ne sont pas linéaires et l’estimation de $(\mathbf{X}_n)_{n>0}$ peut se faire soit par linéarisation du système si $q = 1$ comme dans [3] soit en utilisant des méthodes de Monte-Carlo séquentielles *cf.* [4].

Néanmoins, malgré le fait que les méthodes de Monte-Carlo soient asymptotiquement optimales, en pratique elles peuvent ne pas être les plus efficaces du point de vue du rapport qualité/temps de calcul. Parmi les alternatives possibles, on trouve le concept du filtre gaussien basé sur

la méthode de Laplace, connu comme Laplace-Kalman Filter, récemment introduit dans [5]. Les filtres gaussiens appliqués aux problèmes non linéaires sont exposés dans [6].

L’originalité de notre étude est de traiter le problème de l’estimation de la variance stochastique en grande dimension par les filtres gaussiens fondés sur la méthode de Laplace. Nous trouvons que de tels filtres permettent d’avoir un meilleur ratio qualité/temps de calcul que celui du filtre particulaire. La méthode proposée est donc intéressante dans le cadre de son application au problème de la gestion du risque d’un portefeuille composé d’un grand nombre d’actifs.

Ce document est organisé en quatre parties. Dans la partie 2 nous rappelons les méthodes statistiques que nous avons utilisées. La partie 3 contient la présentation de l’algorithme de filtrage par la méthode de Laplace dans le contexte d’un modèle de variance stochastique à corrélation constante. La dernière partie est réservée aux expériences.

2 Méthodes

2.1 Notations

— $\mathbf{X} \sim p(\mathbf{x})$: variable aléatoire \mathbf{X} suit $p(\mathbf{x})$;

- $\mathbf{X}_{1:N}$: variables aléatoire $\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N$;
- $\mathcal{N}(\mathbf{m}, \Sigma)$: distribution gaussienne de moyenne \mathbf{m} et de covariance Σ ;
- $\mathbf{V}[i]$: i -ème composante du vecteur \mathbf{V} ;
- $\mathbf{M}[i, j]$: (i, j) -ème composante de la matrice \mathbf{M} ;
- $\mathbf{1}$: vecteur colonne vérifiant $\forall i, \mathbf{1}[i] = 1$;
- $\|\mathbf{x} - \mathbf{y}\|_{\mathbf{P}}$: la distance de Mahalanobis entre \mathbf{x} et \mathbf{y} induite par une matrice réelle symétrique définie positive \mathbf{P} , $\|\mathbf{x} - \mathbf{y}\|_{\mathbf{P}} \stackrel{def}{=} \sqrt{(\mathbf{x} - \mathbf{y})^\top \mathbf{P} (\mathbf{x} - \mathbf{y})}$;
- $\det \mathbf{M}$: le déterminant d'une matrice \mathbf{M} ;
- $\mathcal{D}(\mathbf{V})$: matrice diagonale qui vérifie $\forall i, \mathcal{D}(\mathbf{V})[i, i] = \mathbf{V}[i]$;
- $\nabla f(\mathbf{x})$ gradient d'une fonction vectorielle f en \mathbf{x} ;
- $\nabla^2 f(\mathbf{x})$ hessienne d'une fonction vectorielle f en \mathbf{x} ;
- $\mathbf{A} \circ \mathbf{B}$: produit de Hadamard des deux matrices de mêmes dimensions : $\forall (i, j), \mathbf{A} \circ \mathbf{B}[i, j] \stackrel{def}{=} \mathbf{A}[i, j] \mathbf{B}[i, j]$.

2.2 Filtre particulaire

Considérons un modèle de Markov caché général suivant :

$$\mathbf{X}_{n+1} \sim p(\mathbf{x}_{n+1} | \mathbf{x}_n) ; \quad (1a)$$

$$\mathbf{Y}_n \sim p(\mathbf{y}_n | \mathbf{x}_n), \quad (1b)$$

où $n \in \mathbb{N}^*$, $\mathbf{X}_n \in \mathbb{R}^p$ et $\mathbf{Y}_n \in \mathbb{R}^q$ sont des variables cachées et observées respectivement, les distributions dans (1) sont données.

Nous considérons les méthodes de filtrage particulaire qui sont des méthodes d'échantillonnage d'importance séquentielles avec rééchantillonnage (Sampling Importance Resampling). Ces méthodes sont couramment utilisées en pratique et consistent à répéter la procédure suivante, pour tout $n \geq 0$:

1. Pour $m = 1 : M$, tirer $\tilde{\mathbf{x}}_{n+1}^{(m)} \sim p(\mathbf{x}_{n+1} | \mathbf{x}_n^{(m)})$;
2. Pour $m = 1 : M$, calculer $\omega_{n+1}^{(m)} = p(\mathbf{y}_{n+1} | \tilde{\mathbf{x}}_{n+1}^{(m)})$;
3. Obtenir $\{\mathbf{x}_{n+1}^{(m)}\}_{1 \leq m \leq M}$ en tirant M particules de $\{\tilde{\mathbf{x}}_{n+1}^{(m)}\}_{1 \leq m \leq M}$ avec les probabilités proportionnelles à $\{\omega_{n+1}^{(m)}\}_{1 \leq m \leq M}$;
4. Approcher $\mathbb{E}[\mathbf{X}_{n+1} | \mathbf{Y}_{1:n+1}]$ par $\frac{1}{M} \sum_{m=1}^M \mathbf{x}_{n+1}^{(m)}$.

Dans l'algorithme présenté, $M > 0$ est le nombre de particules utilisées par le filtre particulaire.

2.3 Filtre gaussien

Nous remarquons que le FP approche la densité de filtrage $p(\mathbf{x}_n | \mathbf{y}_{1:n})$ par $\frac{1}{M} \sum_{m=1}^M \delta(\mathbf{x}_n = \mathbf{x}_n^{(m)})$. Un filtre gaussien approche $p(\mathbf{x}_n | \mathbf{y}_{1:n})$ par une distribution gaussienne

$$p(\mathbf{x}_n | \mathbf{y}_{1:n}) \approx \mathcal{N}(\hat{\mathbf{x}}_n, \mathbf{P}_n), \quad (2)$$

où $\hat{\mathbf{x}}_n$ et \mathbf{P}_n sont calculés récursivement pour tout $n \geq 0$.

L'approximation (2) est exacte dans le cas où le système (1) est linéaire gaussien et le filtre de Kalman permet de calculer la densité de filtrage.

2.4 Méthode de Laplace

Soit f une fonction $\mathbb{R}^d \rightarrow \mathbb{R}$ deux fois dérivable, convexe, qui admet un minimum en \mathbf{x}_0 . Le principe de la méthode de Laplace est d'admettre que seuls les points du voisinage de \mathbf{x}_0 contribuent de façon significative à l'intégrale

$$I = \int \exp(-f(\mathbf{x})) d\mathbf{x}. \quad (3)$$

Par application du théorème de Taylor, $f(\mathbf{x}) - f(\mathbf{x}_0)$ s'écrit au voisinage \mathbf{x}_0 comme suit :

$$\nabla f(\mathbf{x}_0)(\mathbf{x} - \mathbf{x}_0) + \frac{1}{2} \|\mathbf{x} - \mathbf{x}_0\|_{\nabla^2 f(\mathbf{x}_0)}^2 + o(\|\mathbf{x} - \mathbf{x}_0\|_{\nabla^2 f(\mathbf{x}_0)}^2),$$

où $\|\mathbf{x} - \mathbf{x}_0\|_{\nabla^2 f(\mathbf{x}_0)}^2 = (\mathbf{x} - \mathbf{x}_0)^\top \nabla^2 f(\mathbf{x}_0) (\mathbf{x} - \mathbf{x}_0)$. Sachant que \mathbf{x}_0 est le minimum de f , $\nabla f(\mathbf{x}_0) = \mathbf{0}$, donc

$$f(\mathbf{x}) \approx f(\mathbf{x}_0) + \frac{1}{2} \|\mathbf{x} - \mathbf{x}_0\|_{\nabla^2 f(\mathbf{x}_0)}^2, \quad (4)$$

et l'approximation de (3) correspondante est

$$I \approx \exp(-f(\mathbf{x}_0)) \int \exp\left(-\frac{1}{2} \|\mathbf{x} - \mathbf{x}_0\|_{\nabla^2 f(\mathbf{x}_0)}^2\right) d\mathbf{x} = \exp(-f(\mathbf{x}_0)) \frac{(2\pi)^{d/2}}{(\det \nabla^2 f(\mathbf{x}_0))^{1/2}}. \quad (5)$$

En combinant (4) et (5), nous remarquons que

$$\frac{1}{I} \exp(-f(\mathbf{x})) \approx \frac{(\det \nabla^2 f(\mathbf{x}_0))^{1/2}}{(2\pi)^{d/2}} \exp\left(-\frac{1}{2} \|\mathbf{x} - \mathbf{x}_0\|_{\nabla^2 f(\mathbf{x}_0)}^2\right),$$

donc la méthode de Laplace peut être interprétée comme approximation de la densité proportionnelle à $\exp(-f(\mathbf{x}))$ par $\mathcal{N}(\mathbf{x}_0, (\nabla^2 f(\mathbf{x}_0))^{-1})$.

3 Filtrage dans un modèle de variance stochastique

3.1 Modèle de variance stochastique

Soit $d \in \mathbb{N}^*$ la dimension de l'espace d'états, $\mathbf{S} \in \mathbb{R}^{d \times d}$ une matrice de corrélation, c'est-à-dire une matrice réelle symétrique définie positive dont les coefficients diagonaux sont égaux à 1. Soit $(\mathbf{X}_n)_{n \geq 0}$ un processus gaussien markovien dans \mathbb{R}^d défini par :

$$\mathbf{X}_1 \sim \mathcal{N}(\mathbf{m}_0, \Sigma_0); \quad (6a)$$

$$\mathbf{X}_{n+1} \sim \mathcal{N}(\mathbf{F}\mathbf{X}_n + \mathbf{a}, \mathbf{Q}). \quad (6b)$$

Définissons, pour tout $n > 0$, le vecteur $\boldsymbol{\nu}(\mathbf{X}_n)$ et la matrice $\boldsymbol{\Sigma}(\mathbf{X}_n)$ par :

$$\boldsymbol{\nu}(\mathbf{X}_n)[i] = \exp\left(\frac{1}{2}\mathbf{X}_n[i]\right) \quad \forall 1 \leq i \leq d; \quad (7a)$$

$$\boldsymbol{\Sigma}(\mathbf{X}_n) = \mathcal{D}\left(\boldsymbol{\nu}(\mathbf{X}_n)\right)\mathbf{S}\mathcal{D}\left(\boldsymbol{\nu}(\mathbf{X}_n)\right). \quad (7b)$$

Le processus $(\mathbf{Y}_n)_{n>0}$ est défini par :

$$\mathbf{Y}_n \sim \mathcal{N}\left(\mathbf{0}, \boldsymbol{\Sigma}(\mathbf{X}_n)\right). \quad (8)$$

Le modèle (6)-(8) de $(\mathbf{X}_n, \mathbf{Y}_n)_{n>0}$ est considéré comme étant un modèle de Markov caché.

Pour tout $n > 0$ et pour tout $i, 1 \leq i \leq d$, la variance de $\mathbf{Y}_n[i]$ conditionnellement à \mathbf{X}_n est $\exp(\mathbf{X}_n[i])$ et la matrice de corrélation de \mathbf{Y}_n conditionnellement à \mathbf{X}_n est \mathbf{S} . C'est donc un modèle de variance stochastique à corrélation constante.

3.2 Filtrage basé sur la méthode de Laplace

Nous considérons l'approximation (2). Alors, la densité $p(\mathbf{x}_{n+1} | \mathbf{y}_{1:n})$ est gaussienne, car la densité de transition (6b) est gaussienne. Nous avons donc :

$$p(\mathbf{x}_{n+1} | \mathbf{y}_{1:n}) = \frac{\exp\left(-\frac{1}{2}\|\mathbf{x}_{n+1} - \widehat{\mathbf{x}}_{n+1}^-\|_{\mathbf{P}_{n+1|n}^{-1}}^2\right)}{(2\pi)^{d/2}\left(\det \mathbf{P}_{n+1|n}\right)^{1/2}} \quad (9)$$

avec

$$\begin{aligned} \widehat{\mathbf{x}}_{n+1}^- &= \mathbf{F}\widehat{\mathbf{x}}_n + \mathbf{a}; \\ \mathbf{P}_{n+1|n} &= \mathbf{F}\mathbf{P}_n\mathbf{F}^\top + \mathbf{Q}. \end{aligned}$$

De manière équivalente, (9) s'écrit :

$$\log p(\mathbf{x}_{n+1} | \mathbf{y}_{1:n}) = -\frac{1}{2}\|\mathbf{x}_{n+1} - \widehat{\mathbf{x}}_{n+1}^-\|_{\mathbf{P}_{n+1|n}^{-1}}^2 + const.$$

La distribution $p(\mathbf{y}_{n+1} | \mathbf{x}_{n+1})$ du modèle (6)-(8) est :

$$p(\mathbf{y}_{n+1} | \mathbf{x}_{n+1}) = \frac{\exp\left(-\frac{1}{2}\mathbf{y}_{n+1}^\top \boldsymbol{\Sigma}(\mathbf{x}_{n+1})^{-1} \mathbf{y}_{n+1}\right)}{(2\pi)^{d/2}\left(\det \boldsymbol{\Sigma}(\mathbf{x}_{n+1})\right)^{1/2}}.$$

Nous avons, d'après (7b) :

$$\boldsymbol{\Sigma}(\mathbf{x}_{n+1})^{-1} = \mathcal{D}\left(\boldsymbol{\nu}(\mathbf{x}_{n+1})\right)^{-1} \mathbf{S}^{-1} \mathcal{D}\left(\boldsymbol{\nu}(\mathbf{x}_{n+1})\right)^{-1}.$$

Soit le vecteur colonne $\mathbf{t}_{n+1}(\mathbf{x}_{n+1})$ dans \mathbb{R}^d défini par :

$$\mathbf{t}_{n+1}(\mathbf{x}_{n+1})[i] = \exp\left(-\frac{1}{2}\mathbf{x}_{n+1}[i]\right) \mathbf{y}_{n+1}[i] \quad \forall 1 \leq i \leq d,$$

alors nous avons :

$$\mathbf{y}_{n+1}^\top \boldsymbol{\Sigma}(\mathbf{x}_{n+1})^{-1} \mathbf{y}_{n+1} = \mathbf{t}_{n+1}(\mathbf{x}_{n+1})^\top \mathbf{S}^{-1} \mathbf{t}_{n+1}(\mathbf{x}_{n+1}).$$

Ensuite, nous avons à une constante additive près,

$$\begin{aligned} \log p(\mathbf{y}_{n+1} | \mathbf{x}_{n+1}) &= \\ &= -\log\left(\det \boldsymbol{\Sigma}(\mathbf{x}_{n+1})\right)^{1/2} - \frac{1}{2}\mathbf{t}_{n+1}(\mathbf{x}_{n+1})^\top \mathbf{S}^{-1} \mathbf{t}_{n+1}(\mathbf{x}_{n+1}). \end{aligned}$$

Nous pouvons simplifier davantage l'équation précédente en remarquant que :

$$\left(\det \boldsymbol{\Sigma}(\mathbf{x}_{n+1})\right)^{1/2} = (\det \mathbf{S})^{1/2} \det \mathcal{D}\left(\boldsymbol{\nu}(\mathbf{x}_{n+1})\right).$$

Nous avons alors :

$$\log\left(\det \boldsymbol{\Sigma}(\mathbf{x}_{n+1})\right)^{1/2} = \frac{1}{2}\mathbf{1}^\top \mathbf{x}_{n+1} + \log(\det \mathbf{S})^{1/2},$$

et finalement :

$$\begin{aligned} \log p(\mathbf{y}_{n+1} | \mathbf{x}_{n+1}) &= \\ &= -\frac{1}{2}\mathbf{1}^\top \mathbf{x}_{n+1} - \frac{1}{2}\mathbf{t}_{n+1}(\mathbf{x}_{n+1})^\top \mathbf{S}^{-1} \mathbf{t}_{n+1}(\mathbf{x}_{n+1}) + const. \end{aligned}$$

Nous considérons une approximation de $p(\mathbf{x}_{n+1} | \mathbf{y}_{1:n+1})$ par une distribution gaussienne en utilisant la méthode de Laplace. Sachant que :

$$p(\mathbf{x}_{n+1} | \mathbf{y}_{1:n+1}) \propto p(\mathbf{x}_{n+1} | \mathbf{y}_{1:n}) p(\mathbf{y}_{n+1} | \mathbf{x}_{n+1}),$$

nous avons :

$$\begin{aligned} \log p(\mathbf{x}_{n+1} | \mathbf{y}_{1:n+1}) &= \\ \log p(\mathbf{x}_{n+1} | \mathbf{y}_{1:n}) + \log p(\mathbf{y}_{n+1} | \mathbf{x}_{n+1}) + const &= \\ -f_{n+1}(\mathbf{x}_{n+1}) + const, \end{aligned}$$

avec la fonction $f_{n+1}(\mathbf{x}_{n+1})$ définie par :

$$\begin{aligned} f_{n+1}(\mathbf{x}_{n+1}) &= \frac{1}{2}\mathbf{1}^\top \mathbf{x}_{n+1} + \frac{1}{2}\mathbf{t}_{n+1}(\mathbf{x}_{n+1})^\top \mathbf{S}^{-1} \mathbf{t}_{n+1}(\mathbf{x}_{n+1}) + \\ &+ \frac{1}{2}\|\mathbf{x}_{n+1} - \widehat{\mathbf{x}}_{n+1}^-\|_{\mathbf{P}_{n+1|n}^{-1}}^2. \end{aligned}$$

Le gradient et l'hessienne de $f_{n+1}(\mathbf{x}_{n+1})$ sont respectivement

$$\begin{aligned} \nabla f_{n+1}(\mathbf{x}_{n+1}) &= \frac{1}{2}\mathbf{1} - \frac{1}{2}\mathcal{D}\left(\mathbf{t}_{n+1}(\mathbf{x}_{n+1})\right) \mathbf{S}^{-1} \mathbf{t}_{n+1}(\mathbf{x}_{n+1}) + \\ &+ \mathbf{P}_{n+1|n}^{-1}\left(\mathbf{x}_{n+1} - \widehat{\mathbf{x}}_{n+1}^-\right) \end{aligned}$$

et

$$\begin{aligned} \nabla^2 f_{n+1}(\mathbf{x}_{n+1}) &= \frac{1}{4}\mathcal{D}\left(\mathbf{t}_{n+1}(\mathbf{x}_{n+1})\right) \circ \left(\mathbf{S}^{-1} \mathbf{t}_{n+1}(\mathbf{x}_{n+1}) \mathbf{1}^\top\right) + \\ &+ \frac{1}{4}\mathcal{D}\left(\mathbf{t}_{n+1}(\mathbf{x}_{n+1})\right) \mathbf{S}^{-1} \mathcal{D}\left(\mathbf{t}_{n+1}(\mathbf{x}_{n+1})\right) + \mathbf{P}_{n+1|n}^{-1}. \end{aligned}$$

Par application de la méthode de Laplace, nous définissons $\widehat{\mathbf{x}}_{n+1}$ par une solution de $\nabla f_{n+1}(\mathbf{x}_{n+1}) = \mathbf{0}$ et nous avons $\mathbf{P}_{n+1} = \left(\nabla^2 f_{n+1}(\widehat{\mathbf{x}}_{n+1})\right)^{-1}$.

4 Expériences

Soit ρ et ϕ deux réels, nous considérons le MVS (6)-(8), où pour tout $n > 0$:

$$p(\mathbf{x}_n, \mathbf{x}_{n+1}) = \mathcal{N}(\mathbf{0}, \boldsymbol{\Gamma});$$

$$\boldsymbol{\Gamma} = \begin{bmatrix} 1 & \phi & \cdots & \phi \\ \phi & 1 & \cdots & \phi \\ \vdots & & \ddots & \phi \\ \phi & \phi & \cdots & 1 \end{bmatrix} \in \mathbb{R}^{2d \times 2d} \text{ et } \mathbf{S} = \begin{bmatrix} 1 & \rho & \cdots & \rho \\ \rho & 1 & \cdots & \rho \\ \vdots & & \ddots & \rho \\ \rho & \rho & \cdots & 1 \end{bmatrix}.$$

Les matrices $\mathbf{\Gamma}$ et \mathbf{S} sont définies positives lorsque $|\phi| < 1$ et $|\rho| < 1$. La matrice $\mathbf{\Gamma}$ peut être présentée par blocs de taille $d \times d$:

$$\mathbf{\Gamma} = \begin{bmatrix} \mathbf{\Gamma}_{1,1} & \mathbf{\Gamma}_{1,2} \\ \mathbf{\Gamma}_{2,1} & \mathbf{\Gamma}_{2,2} \end{bmatrix},$$

d'où nous avons $\Sigma_0 = \mathbf{\Gamma}_{1,1}$, $\mathbf{m}_0 = \mathbf{a} = \mathbf{0}$ et \mathbf{F} , \mathbf{Q} de (6) sont donnés par les formules classiques du conditionnement gaussien $\mathbf{F} = \mathbf{\Gamma}_{2,1}\mathbf{\Gamma}_{1,1}^{-1}$ et $\mathbf{Q} = \mathbf{\Gamma}_{2,2} - \mathbf{F}\mathbf{\Gamma}_{1,1}$. Nous considérons le cas où $\phi = 0.5$ et $\rho = 0.5$.

Nous appliquons la démarche suivante. Nous simulons une réalisation du MVS de taille $N = 1000$, puis nous estimons la trajectoire cachée $\mathbf{X}_{1:N}$ par le filtre basé sur la méthode de Laplace (noté FML) décrit dans la section 3.2 et par le filtrage particulière FP(M) détaillé dans la section 2.2. Les critères de comparaison entre FML et FP(M) sont, d'une part, l'erreur quadratique moyenne (EQM) :

$$\text{EQM} = \frac{1}{d} \frac{1}{N} \sum_{n=1}^N (\mathbf{x}_n - \hat{\mathbf{x}}_n)^\top (\mathbf{x}_n - \hat{\mathbf{x}}_n), \quad (10)$$

et d'autre part, le temps de calcul moyen (noté Temps) nécessaire pour traiter une trajectoire $\mathbf{Y}_{1:N}$. Ce temps est mesuré entre l'instant où le filtre en question reçoit toute la trajectoire $\mathbf{Y}_{1:N}$ en entrée et l'instant où il retourne la séquence filtrée $\hat{\mathbf{x}}_{1:N}$.

La Figure 1 montre l'évolution de l'EQM des filtres considérés en fonction de la dimension d pour différentes valeurs de M .

On introduit le gain en (performance)/(temps de calcul) du FML par rapport au FP(M) :

$$G_r[\text{FML}/\text{FP}(M)] = \frac{\text{EQM}[\text{FP}(M)]}{\text{EQM}[\text{FML}]} \frac{\text{Temps}[\text{FP}(M)]}{\text{Temps}[\text{FML}]}. \quad (11)$$

Sachant que lorsque le nombre de particules M augmente, le temps de calcul du FP(M) augmente et son EQM diminue, nous considérons deux stratégies de choix de M :

$$M_1 = \min\{M \in \mathbb{N}^* | \text{EQM}[\text{FP}(M)] \leq \text{EQM}[\text{FML}]\}; \quad (12)$$

$$M_2 = \max\{M \in \mathbb{N}^* | \text{Temps}[\text{FP}(M)] \leq \text{Temps}[\text{FML}]\}. \quad (13)$$

Les résultats sont donnés dans le Tableau 1. Ils indiquent qu'en dimension $d \geq 25$, le FML est plus efficace que le FP(M) du point de vue du rapport (performance)/(temps de calcul). La Figure 1 illustre que l'EQM du FP(M) croît avec la dimension pour tout M fixé, tandis que l'utilisation du FML permet une décroissance monotone de l'EQM.

5 Remerciements

Les auteurs remercient les rapporteurs pour la pertinence de leurs remarques.

FIGURE 1 – EQM (10) des estimateurs FML et FP(M) en fonction de la dimension d pour différents M .

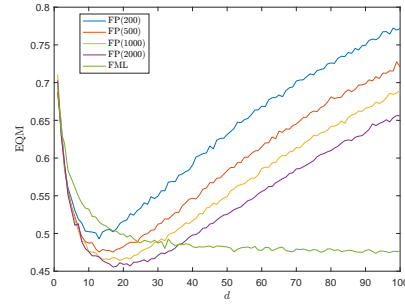


TABLE 1 – $G_r[\text{FML}/\text{FP}(M)]$ (11) en fonction de la dimension d , pour M_1 et M_2 .

	$d = 10$	$d = 25$	$d = 50$	$d = 99$
M_1 – choisi cf. (12)	75	500	10^4	$2 \cdot 10^5$
M_2 – choisi cf. (13)	350	385	400	700
EQM[FP(M_1)]	0.53	0.49	0.48	0.47
EQM[FP(M_2)]	0.49	0.50	0.59	0.70
EQM[FML]	0.53	0.49	0.48	0.47
Temps[FP(M_1)]	0.08 s.	0.70 s.	20.3 s.	600 s.
Temps[FP(M_2)]	0.20 s.	0.50 s.	0.90 s.	2.70 s.
Temps[FML]	0.20 s.	0.50 s.	0.90 s.	2.70 s.
$G_r[\text{FML}/\text{FP}(M_1)]$	0.60	1.40	22.6	222
$G_r[\text{FML}/\text{FP}(M_2)]$	0.92	1.02	1.04	1.49

Références

- [1] A. Harvey, E. Ruiz, and N. Shephard, “Multivariate stochastic variance models,” *The Review of Economic Studies*, vol. 61, no. 2, pp. 247–264, 1994.
- [2] E. Jacquier, N. G. Polson, and P. E. Rossi, “Bayesian analysis of stochastic volatility models,” *Journal of Business & Economic Statistics*, vol. 20, no. 1, pp. 69–87, 2002.
- [3] E. Ruiz, “Quasi-maximum likelihood estimation of stochastic volatility models,” *Journal of econometrics*, vol. 63, no. 1, pp. 289–306, 1994.
- [4] A. Doucet and A. M. Johansen, “A tutorial on particle filtering and smoothing : Fifteen years later,” *Handbook of nonlinear filtering*, vol. 12, no. 656-704, p. 3, 2009.
- [5] P. B. Quang, C. Musso, and F. Le Gland, “The Kalman Laplace filter : A new deterministic algorithm for nonlinear bayesian filtering,” in *2015 18th International Conference on Information Fusion (Fusion)*, Washington, DC, USA, July 2015, pp. 1566–1573.
- [6] K. Ito and K. Xiong, “Gaussian filters for nonlinear filtering problems,” *Automatic Control, IEEE Transactions on*, vol. 45, no. 5, pp. 910–927, 2000.