

Forecasting "High" and "Low" of financial time series by Particle systems and Kalman filters

S. ANANDARAJAH, S. BOUTIGNY, F. DI FABIO, N. LATRACHE

Introduction

L'article *Forecasting "High" and "Low" of financial time series by Particle systems and Kalman filters* écrit par S. Dablemon, S. Van Bellegem et M. Verleysen et publié en 2007 a pour but de présenter des méthodes qui peuvent être utiles pour la prédiction de chroniques. En particulier, ils s'intéressent à la prédiction du premier "stop-loss time" évalué à partir de "tick data" et d'un modèle de trading. Le premier "stop-loss time" est défini comme le moment t où la série dépasse pour la première fois un seuil défini par le trader. L'originalité de cette recherche est représentée par le fait que les auteurs ne s'appuient plus sur des modèles financiers théoriques (comme les modèles à volatilité stochastique). En effet, pour leurs estimations, ils font appel à une représentation à espace d'état non paramétrique avec des fonctions non linéaires *Radial Basis Function*. Les fonctions utilisées sont définies de la façon suivante:

$$f(x_k) = \sum_{i=1}^I \lambda_i^f \exp \left[-\frac{1}{\sigma_i^f} (x_k - c_i^f)^t (x_k - c_i^f) \right]$$
$$h(x_k) = \sum_{i=1}^I \lambda_i^h \exp \left[-\frac{1}{\sigma_i^h} (x_k - c_i^h)^t (x_k - c_i^h) \right]$$

Les estimations sont faites à l'aide du filtre à particule, une méthode de filtrage permettant d'abandonner le cadre gaussien et de mettre à jours de façon récursive la distribution à posteriori à l'aide d'un échantillon de points pondérés. Avant de présenter cette méthode, les auteurs parcourent d'autres méthodes de filtrage (à savoir LKF, EKF, UKF) afin d'en présenter les limites et justifier ainsi leurs choix. Tout au long de l'article le système à espace d'état est défini de la façon suivante:

$$\begin{cases} x_{k+1} = f_k(x_k, u_k, v_k; w) \\ y_k = h_k(x_k, v_k; w) \end{cases}$$

où f (fonction d'état) et h (fonction de mesure) sont des fonctions linéaires pour le filtre LKF et non linéaires pour EKF, UKF et SPPF.

Linear Kalman Filter

Le filtre de Kalman linéaire est une procédure récursive pour calculer les estimations optimales du vecteur d'état. Il fonctionne en deux temps: la prévision et la mise à jour. La première étape de

prédiction consiste à prédire en t la valeur de la variable d'état en $t + 1$ ainsi que sa matrice de variance-covariance. La seconde étape de mise à jour consiste à corriger les valeurs prédites sur la base des valeurs observées et de la matrice de gain de Kalman. La matrice de gain de Kalman attribut un poids à l'innovation, représentée par l'écart entre l'observation et la valeur prédite de l'observation. En d'autres termes, on filtre la valeur de la variables d'état en $t + 1$ à l'aide de l'information connue jusqu'à $t + 1$.

Le filtre de Kalman permet donc d'estimer de façon récursive la distribution de la variable d'état x (sous hypothèse de normalité jointe des résidus). Quand l'hypothèse de linéarité est enlevée, le LKF fournit une estimation optimale parmi les estimateurs linéaires mais pas en général. Etant donnée la forte non linéarité que l'on peut observer dans les chroniques financières, il convient d'utiliser d'autres filtres capables de traiter la non linéarité des fonctions.

Extended Kalman Filter

Le filtre de Kalman exige des équations linéaires. Lorsque ce n'est plus le cas, il convient d'utiliser le filtre de Kalman étendu avec une densité à posteriori approximée par une distribution gaussienne. La méthode de filtrage est, comme dans le cas précédent, divisé en deux temps: prévision et mise-à-jour. Etant donnée l'utilisation de fonctions non linéaires, le filtre EKF prévoit la linéarisation des équation d'état et de mesure afin de calculer la matrice de variance covariance prévu du vecteur d'état. Pour cela, nous calculons ses matrices jacobiennes soit les matrices de dérivées partielles par rapport aux variables $x = \hat{x}_{k|k-1}$ et $v = \hat{v}$. Le filtre se déroule alors de la façon suivante:

Prévision:

$$\begin{aligned}\hat{x}_{k|k-1} &= f(\hat{x}_{k-1|k-1}, \bar{v}, u_k) \\ P_{k|k-1} &= F_k P_{k-1|k-1} F_k^t + G_k R G_k^t\end{aligned}$$

Mise à jour:

$$\begin{aligned}\Sigma_{k|k-1} &= H_k P_{k|k-1} H_k^t + D_k R D_k^t \\ K_k &= P_{k|k-1} H_k^t \Sigma_{k|k-1}^{-1} \\ e_k &= y_k - h_k(\hat{x}_{k|k-1}, \bar{n}) \\ \hat{x}_{k|k} &= \hat{x}_{k|k-1} + K_k e_k \\ P_{k|k} &= P_{k|k-1} - K_k H_k P_{k|k-1}\end{aligned}$$

où F_k , G_k et H_k sont des matrices jacobiennes (par exemple $F_k = \nabla f(x, \bar{v}, u_k)|_x$). Après linéarisation, les équations du filtre de Kalman linéaire peuvent être utilisées dans le cadre du filtre de Kalman étendu. Cependant, la stabilité du filtre de Kalman étendu est difficile à garantir. En effet, ce filtre n'est valable que lorsque la densité postérieure $p(x_k|y_{1:k})$ est bien approximée par une distribution gaussienne. Cela suppose que l'erreur d'estimation initiale et tous les bruits soient suffisamment petits et que la non-linéarité du modèle ne soit pas importante, le calcul de la matrice jacobienne étant compliqué pour certains modèles. De plus, l'EKF fournit des approximations de premier ordre des termes optimaux qui peuvent introduire des erreurs. Ces limites conduisent à des performances sous-optimales et parfois à des divergences du filtre. Ainsi, par la suite nous allons voir le filtre de Kalman non-parfumé (UKF) qui propose une autre méthode de prise en charge des équations non-linéaires.

Unscented Kalman Filter

L'unscented Kalman filter est une technique de filtrage nous donnant la possibilité d'estimer le vecteur d'état sans passer par la case de la linéarisation. Cela permet de résoudre les problèmes d'estimation et de filtrage sous-optimal engendrés par le filtre EKF.

Supposons d'avoir des points distribués selon une loi gaussienne que l'on veut mapper à l'aide d'une fonction non linéaire f ; étant donné la non linéarité de la fonction, nous ne pouvons pas conclure que les points mappés sont normalement distribués. L'UKF fait alors appelle à la technique de l'unscented transformation: cette technique nous dit comment choisir les points (que l'on appelle sigma points) pour que les points mappés aient une distribution qui approxime au mieux une loi gaussienne.

L'unscented Kalman filter se déroule en 4 étapes:

1. On calcule les sigma points:

$$\begin{aligned}\chi^0 &= \mu \\ \chi^i &= \mu + \gamma(\sqrt{\Sigma})_i \quad \forall i = 1, \dots, L \\ \chi^i &= \mu - \gamma(\sqrt{\Sigma})_i \quad \forall i = L + 1, \dots, 2L\end{aligned}$$

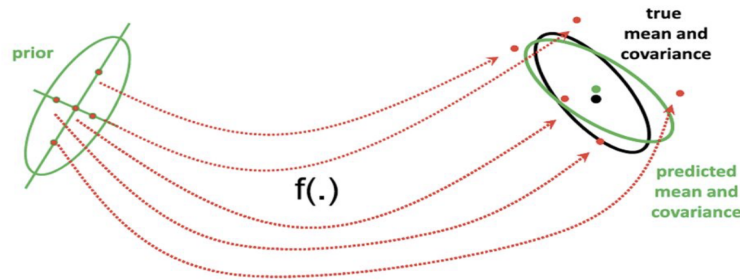
avec $\gamma = \sqrt{L + \lambda}$ où L est la dimension du vecteur et λ est un paramètre d'échelle qui nous indique combien on s'éloigne de la moyenne. La matrice racine carrée $\sqrt{\Sigma}$ peut être obtenue à l'aide de la décomposition de Cholesky ou de la diagonalisation.

2. On calcule des poids:

$$\begin{aligned}w_m^0 &= \frac{\lambda}{L + \lambda} \text{ Poids pour le calcul de la moyenne} \\ w_c^0 &= \frac{\lambda}{L + \lambda} + (1 - \alpha^2 + \beta) \text{ Poids pour le calcul de la matrice de covariance} \\ w_m^i &= \frac{\lambda}{2(L + \lambda)} \quad \forall i = 1, \dots, 2L \text{ Poids pour le calcul de la moyenne} \\ w_c^i &= \frac{\lambda}{2(L + \lambda)} \quad \forall i = 1, \dots, 2L \text{ Poids pour le calcul de la matrice de covariance}\end{aligned}$$

avec $\alpha \in (0, 1]$ et β peuvent être utiles pour contrôler les moments d'ordre supérieurs ($\beta = 2$ est le choix optimal pour une distribution gaussienne).

3. On mappe les sigma points à l'aide de la fonction non linéaire:



4. On calcule la moyenne et la matrice de covariance:

$$\begin{aligned}\tilde{\mu} &= \sum_{i=0}^{2L} w_m^i f(\chi^i) \\ \tilde{\Sigma} &= \sum_{i=0}^{2L} w_c^i (f(\chi^i) - \tilde{\mu})(f(\chi^i) - \tilde{\mu})^t\end{aligned}$$

Particle Filter

LKF, EKF et UKF sont des techniques de filtrage qui supposent une distribution à posteriori gaussienne. Cela pourrait engendrer des problèmes d'optimalité lorsque le système est non linéaire et non gaussien. Le filtre à particule nous permet d'estimer le vecteur d'état sans faire des hypothèses spécifiques sur la distribution. L'idée est de représenter la distribution à posteriori à l'aide d'un échantillon de points et des poids associés en utilisant deux techniques: le *Sequential Importance Sampling* et le *Resampling*. Supposons de vouloir tirer des échantillons d'une distribution dont on ne connaît pas sa forme analytique (que l'on appelle target). Ce que le SIS nous dit est que pour définir la distribution target on peut tirer des points d'une distribution connue (que l'on appelle proposal) auxquels on associe des poids $w = \frac{target}{proposal}$. Les poids peuvent être calculés de façon explicite de la façon suivante:

$$w_k^i = w_{k-1}^i \frac{p(y_k | x_k^i) p(x_k^i | x_{k-1}^i)}{\pi(x_k^i | x_{0:k-1}^i, y_{1:k})}$$

où $p(y_k | x_k^i)$ représente la fonction de vraisemblance, $p(x_k^i | x_{k-1}^i)$ représente la distribution de transition à priori et $\pi(x_k^i | x_{0:k-1}^i, y_{1:k})$ le proposal. Le *Resampling* nous permet, ensuite, de donner encore plus d'importance au point ayant un grand poids.

Le filtre à particule est une technique de filtrage qui implémente un filtre Bayésien récursif avec des simulations Monte Carlo. Il se déroule en 4 étapes:

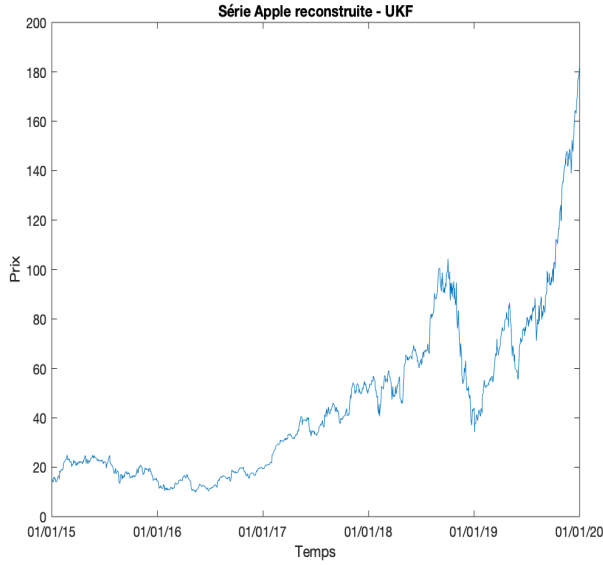
1. Initialisation: on tire N particule de la distribution à priori $p(x_0)$.
2. SIS: on applique l'UKF à chaque particule calculée au point précédent. A l'aide de la moyenne et la matrice de covariance on tire un autre échantillon et on estime la distribution de transition à priori, la fonction de vraisemblance et le proposal. On calcule enfin les poids associés à chaque $\hat{x}_i \forall i = 1, \dots, N$.
3. Resampling: on supprime les points dont le poids est bas et on double ceux dont le poids est élevé.
4. Estimation: on estime le paramètre d'état à l'aide des poids et des points simulés.

Applications numériques

Nous proposons deux applications numériques pour tester les algorithmes de filtrage EKF, UKF et SPPF. Nous avons utilisé la série de prix quotidiens de Apple du 02/01/2015 au 31/12/19.

Dans la première application, nous utilisons un système SSM avec des fonctions RBF à un seul layer caché afin de filtrer la variable latente. Nous avons ensuite utilisé les valeurs filtrées pour reconstruire

la trajectoire du vecteur d'observation. On obtient les deux graphiques suivants:



Dans la deuxième application, nous avons essayé d'extraire la volatilité de la série à l'aide du modèle d'Heston. Nous avons considéré le système SSM suivant:

$$dV_t = k(\theta - V_t)dt + \eta\rho(d\ln S_t - (\mu - \frac{1}{2}V_t)dt) + \eta\sqrt{V_t}\sqrt{1 - \rho^2}dB_t$$

$$d\ln S_t = (\mu - \frac{1}{2}V_t)dt + \sqrt{V_t}dW_t$$

On obtient les deux graphiques suivants:

