

Chacune des réponses aux questions doit être appuyée par des démonstrations, des analyses, des tableaux ou des graphes supportant vos affirmations. N'hésitez pas à produire des figures ou des tableaux supplémentaires pour appuyer votre diagnostic.

Ce travail doit être réalisé en équipe de 3 ou 4 étudiants. Je m'attends à ce que tous les membres de l'équipe s'impliquent dans chacune des composantes du travail, autant dans le codage, que dans les analyses, les dérivations mathématiques et la rédaction du rapport. Vos serez questionné à ce sujet lors des entrevues.

Date de remise le 30 novembre 2021.

1 Heston model

Le document Heston_93.pdf décrit comment implémenter un filtre à particule pour estimer le modèle à volatilité stochastique de Heston. Le code Matlab Script_simulation.m et les fonctions associées permettent d'estimer le modèle en données simulées.

1.1 Environnement simulé

Un filtre dépend toujours de l'échantillon utilisé afin de filtrer les variables latentes et estimer les paramètres du modèle. Dans le code qui vous est fourni, nous étudions deux cas: (a) le modèle est estimé en utilisant les rendements quotidiens de l'actif $y_{1:T} = \{y_t\}_{t=1}^T$ et (b) en utilisant les rendements quotidiens ainsi que les variances réalisées $\{y_t, RV_t\}_{t=1}^T$.

Puisque nous sommes dans un environnement simulé, nous connaissons la valeur de la variable latente (la variance instantanée) tous les jours: $\{V_t\}_{t=1}^T$. Nous connaissons aussi les vraies valeurs $\Theta = (\mu, \kappa, \theta, \sigma, \rho, \eta)$ des paramètres. Il est donc possible de comparer $\{V_t\}_{t=1}^T$ à la variance instantanée filtrée $\{\hat{V}_t\}_{t=1}^T$. Cette variance filtrée dépend du choix de l'échantillon (avec ou sans RV) et de la valeur des paramètres (les vrais versus ceux estimés par la méthode du maximum de vraisemblance).

Question 1. Pour chacun des 4 cas possibles, tracez un histogramme des erreurs d'estimation $\varepsilon_t = \hat{V}_t - V_t$. Obtenez-vous un estimateur sans biais de la variance instantanée? Quelle approche vous semble la plus précise?

Question 2. Le programme Script_simulation.m produit deux figures présentant 5 panels. Il s'agit de la fonction de log-vraisemblance (plus précisément, le négatif de la log-vraisemblance). Pour chacune des 5 dimensions (une dimension par paramètre). La première figure est le résultat de la log-vraisemblance lorsque seuls les rendements quotidiens sont utilisés pour l'estimation. La deuxième figure ajoute l'information par rapport aux variances réalisées. Quel est l'impact de la convexité de la fonction de log-vraisemblance sur l'estimation des paramètres et leur précision? Y a-t-il des paramètres qui vous semblent plus difficiles à estimer que d'autres?

2 Environnement en données réelles

Le fichier Data.xlsx contient la date, le rendement quotidien du S&P500 ainsi que la variance réalisée quotidienne. Le but est de comparer les résultats de l'estimation en utilisant deux échantillons différents. L'échantillon A ne contient que les rendements du S&P500. L'échantillon B contient deux séries chronologiques: celle de S&P500 et celle de la variance réalisée.

Question 3. Estimer le modèle de Heston avec l'échantillon A.

- (a) Présentez un tableau donnant l'estimation de chacun des paramètres. Interprétez vos résultats.
- (b) Tracez un graphe présentant l'évolution de la variance instantanée filtrée ainsi que son intervalle de

confiance de niveau 95%. Ajoutez la courbe de variance réalisée annualisée et la comparer à la variance instantanée filtrée. Interprétez les différences et les similitudes.

Question 4. Reprenez la question 3 en utilisant l'échantillon B.

Question 5. Pour cette question, le but est de comparer les résultats de l'estimation du modèle de Heston selon l'échantillon utilisé. (a) Comparez l'estimation des paramètres. Interprétez vos résultats. (b) Comparez les variances filtrées entre elles ainsi qu'avec la variance réalisée annualisée. Quelles sont vos conclusions? (c) Comparez les longueurs des intervalles de confiance autour des variances filtrées. Quelle approche vous semble la plus précise? (d) Il y a des périodes pour lesquelles le modèle semble avoir de la difficulté à s'ajuster aux données. Quelles sont ces périodes? Est-ce que ce sont les mêmes périodes pour les deux échantillons? (e) Le niveau de la variance filtrée avec l'échantillon A est presque tout le temps supérieur au niveau de la variance filtrée avec l'échantillon B. Cependant, en données simulées, ce phénomène n'est pas présent. Expliquez pourquoi. (f) Comment faudrait-il modifier le modèle pour avoir un meilleur ajustement aux données?