Devoir 1

David Benatia

Économétrie des séries temporelles

MATH60837, M.Sc.

Instructions:

1. *Pour réaliser ce devoir, formez une équipe comprenant au maximum trois membres.*
2. *Utilisez le logiciel que vous préférez.*
3. *Fournissez un document pdf détaillé exporté depuis un notebook Python ou R, ou un Livescript MATLAB (par exemple ce document). Le code doit être propre et facile à comprendre, les figures doivent avoir des titres, les axes doivent avoir des étiquettes, les graphiques doivent avoir des légendes. Des explications, des justifications et des interprétations sont requises.*
4. *Le travail doit être terminé pour le 2 mars 2023. Soumettre électroniquement dans ZoneCours (remise de Travaux).*
5. *Une seule remise par équipe suffit.*
6. *Aucun retard ne sera toléré.*
7. *Toute forme de plagiat est inacceptable.*
8. *Téléchargez le fichier accessible à l'adresse https://www.dropbox.com/s/vfq5cfz9dibzobv/data_W2023.csv?dl=0. Il contient des données sur le taux de change bilatéral entre le dollar américain et le dollar canadien.*

Questions

Partie 1. Analyse préliminaire (20 points)

Définissons la série $y_t = log(100\times USDCAD_t) pour le reste de ce travail.

1.a) Chargez les données et ne gardez que les observations jusqu'à décembre 2019. Tracez la série brute y_t avec des étiquettes sur chaque axe et un titre 'log(100 USDCAD=x)'. La dimension temporelle doit apparaître sur l'axe des abscisses.

Interprétation :

1.b) Supposons qu'il existe une tendance temporelle déterministe (polynôme d'ordre 2) : $\tau_t = \alpha + \beta t + \gamma t^2$. Appliquez la transformation requise pour stationnariser la série logarithmique (le logarithme du taux de change) et tracez la série résultante.

Interprétation :

1.c) Supposons qu'il existe une tendance temporelle stochastique, c'est-à-dire une marche aléatoire. Appliquez la transformation requise pour stationnariser la série logarithmique (le logarithme du taux de change) et tracez la série résultante.

Interprétation :

1.d) Analysez les fonctions d'autocorrélation de l'échantillon pour évaluer si les séries sont stationnaires. Quelle série choisiriez-vous pour estimer un modèle de série chronologique ?

Interprétation :

1.e) A partir de maintenant, utilisez exclusivement la série qui est stationnaire. Effectuez les tests de Ljung-Box avec 1 à 18 lags pour vérifier si la série est un bruit blanc et concluez.

Interprétation :

Partie 2. Choix du modèle (30 points)

Considérons les 8 modèles suivants :

AR(1): $y_t = \delta + \phi y_{t-1} + u_t$

AR(2): $y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + u_t$

AR(3): $y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + u_t$

AR(4): $y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \phi_4 y_{t-4} + u_t$

ARMA(1,1): $y_t = \delta + \phi_1 y_{t-1} + u_t - \theta_1 u_{t-1}$

ARMA(2,2): $y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2}$

ARMA(3,3): $y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + u_t - \theta_1 u_{t-1} - \theta_3 u_{t-2} - \theta_2 u_{t-3}$

ARMA(4,4): $y_t = \delta + \phi_1 y_{t-1} + \phi_2 y_{t-2} + \phi_3 y_{t-3} + \phi_4 y_{t-4} + u_t - \theta_1 u_{t-1} - \theta_2 u_{t-2} - \theta_3 u_{t-3} - \theta_4 u_{t-4}$

2.a) Estimez les 8 modèles par maximum de vraisemblance, présentez les résultats de l'estimation et vérifiez si les conditions de stationnarité sont satisfaites.

Interprétation :

2.b) Effectuez des tests de rapport de vraisemblance pour justifier la sélection de deux modèles. Sélectionnez d'abord le meilleur parmi AR(1), AR(2), AR(3) et AR(4). Sélectionnez ensuite le meilleur modèle parmi ARMA(1,1), ARMA(2,2), ARMA(3,3) et ARMA(4,4). Utilisez le BIC pour justifier la sélection du meilleur de ces deux modèles.

Interprétation :

2.c) Évaluez l'hypothèse du bruit blanc pour chacun des deux modèles, et pour le modèle AR(1). Que pouvons-nous conclure ?

Interprétation :

Partie 3. Réponse dynamique et prévision (20 points)

3.a) Pour les deux modèles sélectionnés, évaluez la réponse dynamique pour un horizon de 10 périodes suite à un choc positif de taille $\sigma = 1.15$ survenant à la première période de l'horizon. Tracez les deux fonctions de réponse impulsionnelle sur la même figure et commentez.

Interprétation :

3.b) Un choc négatif de taille $\sigma = 1.15$ se produit pendant 3 périodes consécutives de l'horizon (t, t+1, t+2). Tracez les deux fonctions de réponse impulsionnelle pour les deux modèles sélectionnés, montrant la réponse dynamique pour un horizon de 20 périodes après ces chocs, et expliquez ce qui se passe.

Interprétation :

3.c) Prévision : Divisez l'échantillon en un échantillon d'entraînement (training sample) et un échantillon de validation (holdout sample). L'échantillon de validation devrait être composé des 34 dernières observations.

- Ré-estimez les deux modèles sélectionnés en utilisant uniquement l'échantillon d'entraînement.
- Pour chacun des deux modèles, calculez les prévisions à un pas en avant : $E_t(USDCAD_{t+1})$ pour les modèles AR(4) et ARMA(4,4). Ces prévisions conditionnelles ne sont pas les mêmes !
- Tracez ces prévisions avec la série réelle (différence première) sur le même graphique pour la période couverte par l'échantillon de validation.
- Tracez la série réelle (données brutes en niveaux) contre les prévisions à un pas d'avance et la prévision naïve : $E_t(USDCAD_{t+1}) = USDCAD_t$.

Interprétation :

3.d) Indiquez l'erreur quadratique moyenne pour chaque modèle en utilisant uniquement l'échantillon de 34 observations pour évaluer la performance de la prévision. Comparez ces statistiques à l'estimateur naïf suivant : $E_t(y_{t+1}) = y_t$ pour h=1 seulement. Quel modèle devrions-nous utiliser dans chaque cas ?

Interprétation :

Partie 4. Améliorations supplémentaires par la modélisation de la variance conditionnelle (30 points)

4.a) Tracez les fonctions d'autocorrélation et d'autocorrélation partielle de y_t^2 . Que pouvez-vous conclure ?

Interprétation :

4.b) Estimez les versions ARCH(1) et GARCH(1,1) des deux modèles sélectionnés par maximum de vraisemblance en utilisant seulement l'échantillon d'entraînement, et rapportez les estimations.

Interprétation :

4.c) Effectuez des tests de rapport de vraisemblance pour sélectionner un modèle entre AR(4), AR(4) ARCH(1), et AR(4) GARCH(1,1). Répétez cette procédure pour sélectionner un modèle entre ARMA(4,4), ARMA(4,4) ARCH(1), et ARMA(4,4) GARCH(1,1).

Interprétation :

4.d) Effectuez un test de spécification sur les résidus pour AR(4) ARCH(1) et AR(4) GARCH(1,1). Quel est le meilleur modèle pour la variance conditionnelle ?

Interprétation :

4.e) Pour ces deux nouveaux modèles et les deux modèles étudiés dans la partie 3, tracez les fonctions de réponse impulsionnelle du second cas décrit dans la partie 3. Dans quelle mesure sont-elles différentes ?

Interprétation :

4.f) Comme dans la partie 3, tracez les prévisions à un pas d'avance $E_t(y_{t+1})$ pour la période couverte par l'échantillon de validation (holdout sample). Comparez-les avec celles obtenues dans la partie 3.

Interprétation :

4.g) Comparez les erreurs quadratiques moyennes de la prévision à un pas (pour la série transformée !) pour les deux modèles sélectionnés dans la partie 3, les deux nouveaux modèles, la prévision naïve $E_t(y_{t+1}) = y_t$ **et la prévision paresseuse** $E_t(y_{t+1}) = 0$ **pour tous les t. Cette dernière suppose que la série en niveaux ne changera pas. Que pouvez-vous conclure ? La modélisation de la covariance conditionnelle était-elle utile ?**

Interprétation :