

Modèle ARMA (données simulées)

On considère le processus ARMA(2,1) donné par

$$X_t - 1.3X_{t-1} + 0.5X_{t-2} = \epsilon_t + 0.5\epsilon_{t-1}$$

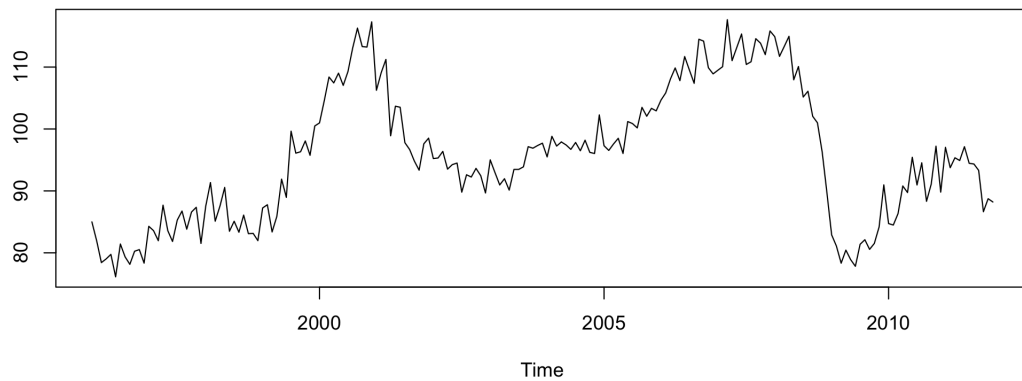
où $\{\epsilon_t\}$ est un bruit blanc gaussien centré réduit.

1. Quelles sont les propriétés de ce processus ?
2. Simuler puis tracer une trajectoire de longueur $T = 500$ de ce processus.
3. Estimer et tracer les fonctions d'autocorrélation et d'autocorrélation partielle du processus. Que peut-on en dire ?
4. Utiliser les 400 premières observations pour estimer les paramètres du modèle ARMA.
Comment procède-t-on si on ne connaît pas l'ordre du modèle pour l'inférer ? Proposer une méthode basée sur le critère BIC puis une autre basée sur des prédictions à un pas de temps.
5. Valider le modèle : étude des résidus, validation croisée.
6. Prédire les observations 401 à 410. Donner un intervalle de prédiction.

0.1 Modèle ARIMA (données réelles)

On considère l'évolution d'un indice de la production d'équipement électronique en Europe (ordinateur, optique, etc.) : `elecequip.Rdata`. La série est normalisée pour que l'indice soit égal à 5 en 2005.

1. Tracer les données. Et commenter le graphique.
2. Desaisonnaliser la série et tracer la série sans saison notée `eeadj` pour obtenir le graphique ci-dessous



3. La série `eeadj` admet-elle une série d'ordre 1 ?
4. Tracer la différence d'ordre 1 : `B(eeadj)` ainsi que ses fonction d'autocorrélation et d'auto corrélation partielles empiriques. Commenter les graphiques obtenus.
5. Proposer un modèle arima pour modéliser la série `eeadj`.
6. Trouver quels sont les modèles AR et MA sélectionnés par le critère AIC pour la série `diff(eeadj)`.
7. Ajuster un modèle `arima(3,1,0)` à la série `eeadj`. Le comparer aux modèles `arima(4,1,0)` et `arima(3,1,1)`.
8. Faire l'étude des résidus pour le modèle `arima(3,1,1)`.
9. Prédire les valeurs de l'indice pour l'année 2017.

Modèle SARIMA (données réelles)

On souhaite prévoir le nombre d'usagers de la SNCF pour les 12 mois de l'année 2001. Pour réaliser la prévision, on dispose des données mensuelles sur 11 années entre 1990 et 2000. Ces données sont disponibles dans le fichier suivant

www.math.sciences.univ-nantes.fr/~philippe/lecture/donnees-sncf-1990-2000.txt

1. Justifier le choix d'une modélisation SARIMA sur cette série à l'aide de quelques graphiques. En déduire une estimation des ordres de saisonnalité et d'intégration.
2. Ajuster un modèle SARIMA. Valider el modèle.
3. Calculer les prévisions mensuelles pour l'année 2001. On pourra comparer les résultats à la série disponible sous le lien suivant.

www.math.sciences.univ-nantes.fr/~philippe/lecture/donnees-sncf-2001.txt