Statistiques des séries temporelles

Processus linéaires, MA, AR, ARMA

Partie I

Processus des innovations

Processus des innovations

- Soit X_t un processus stationnaire au second ordre centré.
- On définit son passé linéaire $\mathcal{H}_t^X = \overline{\mathrm{Vect}}(X_s, s \leq t)$.
- Le processus des innovations de X est défini par

$$\varepsilon_t = X_t - \mathrm{proj}(X_t | \mathcal{H}_{t-1}^X)$$

où la projection doit être comprise au sens L^2 (espace de Hilbert).
- Proposition : le processus des innovations est un bruit blanc faible.
- Sa variance σ^2 est appelée variance d'innovation de X.
- Si $\sigma^2 > 0$, X est appelé processus régulier.
- Si $\sigma^2 = 0$, X est appelé processus déterministe.

Processus des innovations partielles

- Le passé linéaire d'ordre p du processus X est défini par

$$\mathcal{H}_{t,p}^X = Vect(X_t \dots X_{t-p+1}).$$

- La prédiction d'ordre p est définie par

$$\mathrm{proj}(X_t | \mathcal{H}_{t-1,p}^X) = \sum_{k=1}^{p} \phi_{k,p} X_{t-k}$$

où $\phi_p = [\phi_{1,p} \dots \phi_{p,p}]^T$ est indépendant de t.
- Comme $\mathcal{H}_t^X = \overline{\cup_{p \geq 1} \mathcal{H}_{t,p}^X}$, on a

$$\lim_{p \to +\infty} \mathrm{proj}(X_t | \mathcal{H}_{t-1,p}^X) = \mathrm{proj}(X_t | \mathcal{H}_{t-1}^X).$$

Partie II

Processus linéaires

Théorème de filtrage des processus

Soit $(\psi_k)_{k\in\mathbb{Z}}$ une suite absolument sommable : $\sum_{k\in\mathbb{Z}}|\psi_k| < +\infty$.
Soit $(X_t)_{t\in\mathbb{Z}}$ un processus stationnaire au second ordre de moyenne $\mu_X = \mathbb{E}[X_t]$ et de fonction d'autocovariance $\gamma_X(h) = \mathrm{cov}(X_{t+h}, X_t)$.
Alors le processus $Y_t = \sum_{k\in\mathbb{Z}}\psi_k X_{t-k}$ est stationnaire au second ordre, de moyenne $\mu_Y = \mu_X \sum_{k\in\mathbb{Z}}\psi_k$, de fonction d'autocovariance

$$\gamma_Y(h) = \sum_{j\in\mathbb{Z}}\sum_{k\in\mathbb{Z}}\psi_j\overline{\psi_k}\gamma_X(h+k-j),$$

et de mesure spectrale

$$\nu_Y(d\lambda) = |\psi(e^{-i\lambda})|^2\nu_X(d\lambda),$$

où $\psi(e^{-i\lambda}) = \sum_{k\in\mathbb{Z}}\psi_k e^{-ik\lambda}$.

Processus linéaire

- Définition (processus linéaire)
  - $(X_t)_{t\in\mathbb{Z}}$ est un processus linéaire ssi il existe $Z_t \sim \mathrm{BB}(0,\sigma^2)$ et $(\psi_k)_{k\in\mathbb{Z}} \in l_1(\mathbb{Z})$ tels que $X_t = \mu + \sum_{k=-\infty}^{+\infty}\psi_k Z_{t-k}$ $\forall t\in\mathbb{Z}$, où $\mu\in\mathbb{C}$.
  - $(X_t)_{t\in\mathbb{Z}}$ est causal par rapport à $(Z_t)_{t\in\mathbb{Z}}$ ssi $\psi_k = 0$ $\forall k < 0$.
  - $(X_t)_{t\in\mathbb{Z}}$ est inversible par rapport à $(Z_t)_{t\in\mathbb{Z}}$ ssi il existe une suite $(\pi_k)_{k\geq 0} \in l_1(\mathbb{Z})$ telle que $Z_t = \sum_{k=0}^{+\infty}\pi_k(X_{t-k}-\mu)$ $\forall t\in\mathbb{Z}$
- Propriétés (théorème de filtrage des processus)
  - X_t est SSL de moyenne μ, de fonction d'autocovariance $\gamma_X(h) = \mathbb{E}((\overline{X}_t - \overline{\mu})(X_{t+h}-\mu)) = \sigma^2\sum_{k=-\infty}^{+\infty}\overline{\psi}_k\psi_{k+h}$, et de densité spectrale $f_X(\lambda) = \frac{\sigma^2}{2\pi}|\psi(e^{-i\lambda})|^2$, où $\psi(e^{-i\lambda}) = \sum_{k\in\mathbb{Z}}\psi_k e^{-ik\lambda}$.

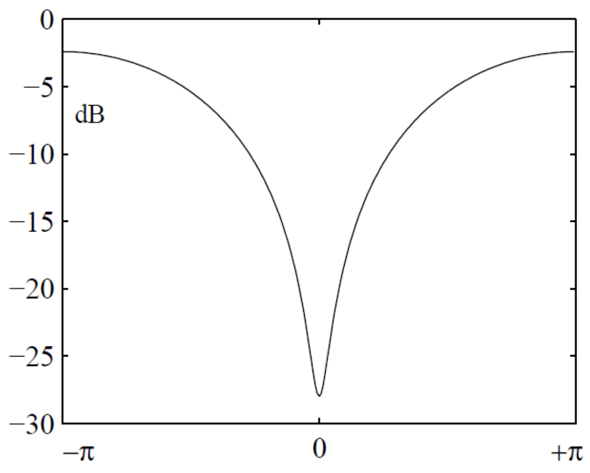
Partie III

Processus à moyenne ajustée

Processus MA(q)

- Définition
  - Le processus $(X_t)_{t\in\mathbb{Z}}$ est à moyenne ajustée d'ordre q (ou MA(q)) ssi $X_t = \sum_{k=0}^{q} \theta_k Z_{t-k}$ où $Z_t \sim \mathrm{BB}(0,\sigma^2)$, $\theta_i \in \mathbb{C}$ et $\theta_0 = 1$.
- Propriétés (théorème de filtrage des processus)
  - X_t est SSL de moyenne 0, de fonction d'autocovariance $\gamma_X(h) = \sigma^2 \sum_{k=0}^{q-h} \overline{\theta}_k \theta_{k+h}$ pour $0 \leq h \leq q$ et $\gamma_X(h) = 0$ si $h > q$, et de densité spectrale $f_X(\lambda) = \frac{\sigma^2}{2\pi}\left|\sum_{k=0}^{q} \theta_k e^{-ik\lambda}\right|^2$.

Processus MA(q)



Densité spectrale (en dB) d'un processus MA-1, pour $\sigma = 1$ et $\theta = -0.9$.

Caractérisation d'un processus MA(q)

- Théorème (Caractérisation d'un processus MA(q)).
  - Soit $(X_t)_{t\in\mathbb{Z}}$ un processus centré SSL de fonction d'autocovariance $\gamma(h)$, et soit $q \geq 1$. Alors les deux assertions suivantes sont équivalentes :
    - X_t est un processus MA d'ordre minimal q ;
    - $\gamma(q) \neq 0$ et $\gamma(h) = 0$ $\forall h \geq q+1$.
  - Preuve : $\mathcal{H}_t^X = \mathcal{H}_{t-q-1}^X \overset{\perp}{\oplus} \mathrm{Vect}(\varepsilon_{t-q} \ldots \varepsilon_t)$ et $X_t \perp \mathcal{H}_{t-q-1}$
- Corollaire
  - La somme de deux processus MA(q) décorrélés est un processus MA(q).

Partie IV

Processus autorégressifs

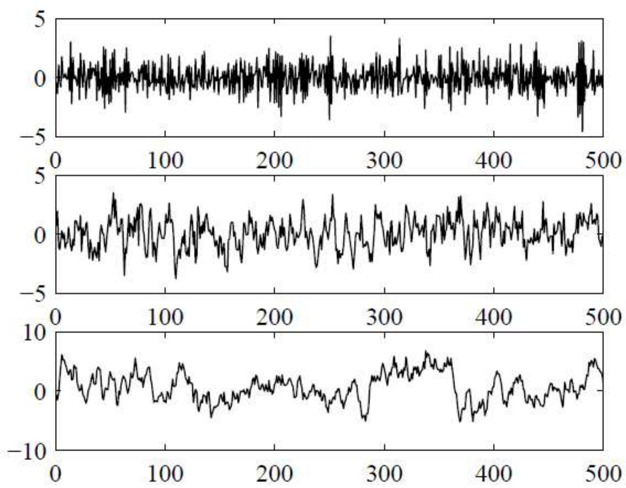
Processus AR(p)

- Définition
  - Le processus $(X_t)_{t\in\mathbb{Z}}$ est autorégressif d'ordre p (ou AR(p)) ssi il est SSL et solution de l'équation $X_t = Z_t + \sum_{k=1}^{p} \phi_k X_{t-k}$ où $Z_t \sim \mathrm{BB}(0,\sigma^2)$, $\phi_k \in \mathbb{C}$.
- L'existence et l'unicité d'une solution SSL est une question délicate, qui ne se posait pas pour les processus MA.

Processus AR(1), cas causal

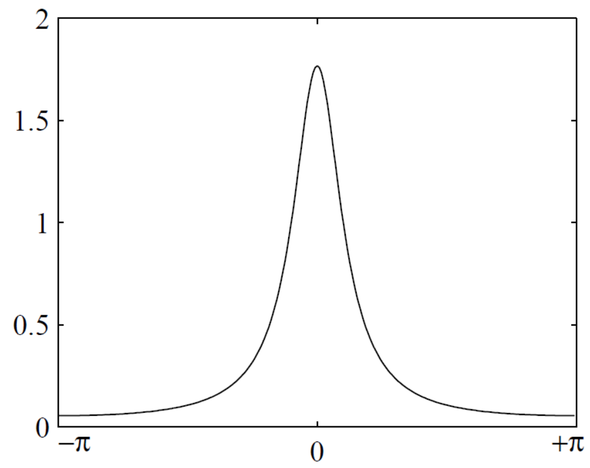
- On applique la récurrence $X_t = Z_t + \phi_1 X_{t-1}$ avec $|\phi_1| < 1$
- $X_t = \sum_{k=0}^{+\infty} \phi_1^k Z_{t-k}$ (convergence dans $L^2(\Omega,\mathcal{A},\mathbb{P})$ et p.s.)
- $X_t = \sum_{k=0}^{+\infty} \psi_k Z_{t-k}$ où $\psi(z) = \frac{1}{1-\phi_1 z} = \sum_{k=0}^{+\infty} \psi_k z^k \Rightarrow \psi_k = \phi_1^k \; \forall k \geq 0$
- Propriétés (théorème de filtrage des processus)
  - X_t est SSL de moyenne 0, de fonction d'autocovariance $\gamma_X(h) = \sigma^2 \sum_{k=0}^{+\infty} \overline{\phi_1^k} \phi_1^{k+h} = \sigma^2 \frac{\phi_1^h}{1-|\phi_1|^2}$ si $h \geq 0$, et de densité spectrale $f_X(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{k=0}^{+\infty} \phi_1^k e^{-ik\lambda} \right|^2 = \frac{\sigma^2}{2\pi} \frac{1}{|1-\phi_1 e^{-i\lambda}|^2}$.

Processus AR(1), cas causal



Trajectoires de longueur 500 d'un processus AR(1)) gaussien. Courbe du haut : $\phi_1 = -0.7$. Courbe du milieu : $\phi_1 = 0.5$. Courbe du bas : $\phi_1 = 0.9$

Processus AR(1), cas causal



Densité spectrale d'un processus AR(1), pour $\sigma = 1$ et $\phi_1 = 0.7$.

Processus AR(1), cas anti-causal

- On applique la récurrence $X_t = -\phi_1^{-1} Z_{t+1} + \phi_1^{-1} X_{t+1}$ avec $|\phi_1| > 1$
- $X_t = -\sum_{k=1}^{+\infty} \phi_1^{-k} Z_{t+k}$ (convergence dans $L^2(\Omega, \mathcal{A}, \mathbb{P})$ et p.s.)
- $X_t = \sum_{j=-\infty}^{-1} \psi_j Z_{t-j}$ où $\psi(z) = \frac{1}{1-\phi_1 z} = \frac{-(\phi_1 z)^{-1}}{1-(\phi_1 z)^{-1}} = \sum_{k=1}^{+\infty} -\phi_1^{-k} z^{-k} \Rightarrow$ $\psi_j = -\phi_1^j$ pour $j < 0$
- Propriétés (théorème de filtrage des processus)
  - X_t est SSL de moyenne 0, de fonction d'autocovariance $\gamma_X(h) = \sigma^2 \sum_{k=-\infty}^{-1} \overline{\phi_1}^{(k-h)} \phi_1^k = \sigma^2 \frac{\overline{\phi_1}^{-h}}{|\phi_1|^2 - 1}$ si $h \geq 0$, et de densité spectrale $f_X(\lambda) = \frac{\sigma^2}{2\pi} \left| \sum_{k=0}^{+\infty} \phi_1^k e^{-ik\lambda} \right|^2 = \frac{\sigma^2}{2\pi} \frac{1}{|1-\phi_1 e^{-i\lambda}|^2}$.

Processus AR(1), cas général

- Si $|\phi_1| < 1$, $X_t = \sum_{k=0}^{+\infty} \phi_1^k Z_{t-k}$
- Si $|\phi_1| > 1$, $X_t = -\sum_{k=1}^{+\infty} \phi_1^{-k} Z_{t+k}$
- Propriétés (théorème de filtrage des processus)
  - Si $|\phi_1| \neq 1$, X_t est SSL de moyenne 0 et de densité spectrale $f_X(\lambda) = \frac{\sigma^2}{2\pi} \frac{1}{|1-\phi_1 e^{-i\lambda}|^2}$.
- Si $|\phi_1| = 1$, l'équation de récurrence n'admet pas de solution SSL

Partie V

Processus ARMA

Processus ARMA(p, q)

- Théorème (Existence et unicité des processus ARMA(p, q))
  - Soit l'équation récurrente : $X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$, où $Z_t \sim \mathrm{BB}(0, \sigma^2)$ et $\phi_j, \theta_j \in \mathbb{C}$.
  - On pose $\phi(z) = 1 - \phi_1 z - \ldots - \phi_p z^p$ et $\theta(z) = 1 + \theta_1 z + \ldots + \theta_q z^q$.
  - On suppose que $\phi(z)$ et $\theta(z)$ n'ont pas de zéros communs.
  - Alors l'équation admet une solution SSL ssi $\phi(z) \neq 0$ $\forall |z| = 1$.
  - Cette solution est unique et a pour expression $X_t = \sum_{k=-\infty}^{+\infty} \psi_k Z_{t-k}$, où les ψ_k sont donnés par les coefficients du développement $\frac{\theta(z)}{\phi(z)} = \sum_{k=-\infty}^{+\infty} \psi_k z^k$, convergeant dans la couronne $\{z \in \mathbb{C}, \delta_1 < |z| < \delta_2\}$, où $\delta_1 < 1$ et $\delta_2 > 1$ sont définis par $\delta_1 = \max\{z \in \mathbb{C}, |z| < 1, \phi(z) = 0\}$ et $\delta_2 = \min\{z \in \mathbb{C}; |z| > 1; \phi(z) = 0\}$.

Densité spectrale d'un processus ARMA

- Théorème (Densité spectrale d'un processus ARMA(p,q)).
  - Soit (X_t) un processus ARMA(p,q), i.e. la solution stationnaire de l'équation $X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$, où $\theta(z)$ et $\phi(z)$ sont des polynômes de degré q et p n'ayant pas de zéros communs et $\phi(z) \neq 0\ \forall |z| = 1$. Alors (X_t) possède une densité spectrale qui a pour expression :

$$f(\lambda) = \frac{\sigma^2}{2\pi} \frac{\left|1 + \sum_{k=1}^{q} \theta_k e^{-ik\lambda}\right|^2}{\left|1 - \sum_{k=1}^{p} \phi_k e^{-ik\lambda}\right|^2}$$

Représentations d'un ARMA(p,q)

- Soit X_t un processus ARMA(p,q) solution de $X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$.
- Alors X_t admet une représentation linéaire $X_t = \sum_{k=-\infty}^{+\infty} \psi_k Z_{t-k}$ pour une suite $\psi_k \in l^1(\mathbb{Z})$ bien choisie.
- On dit que la représentation ARMA(p,q) est
  - causale si le filtre $\psi(z)$ est causal ($\phi(z) \neq 0\ \forall |z| \leq 1$)
  - inversible si le filtre $\psi(z)$ est inversible et si son inverse est causal ($\theta(z) \neq 0\ \forall |z| \leq 1$)
  - canonique si elle est à la fois causale et inversible

Calcul des covariances d'un ARMA causal

- Première méthode
  - Utiliser l'expression $\gamma_X(h) = \sigma^2 \sum_{k=0}^{+\infty} \overline{\psi_k} \psi_{k+h}$ où (ψ_k) se détermine de façon récurrente à partir de $\psi(z)\phi(z) = \theta(z)$, par identification du terme en z^k. Pour les premiers termes on trouve :

$$\begin{aligned} \psi_0 &= 1 \\ \psi_1 &= \theta_1 + \psi_0 \phi_1 \\ \psi_2 &= \theta_2 + \psi_0 \phi_2 + \psi_1 \phi_1 \end{aligned}$$

- Deuxième méthode
  - Utiliser une formule de récurrence, vérifiée par la fonction d'autocovariance d'un processus ARMA(p,q), qui s'obtient en multipliant par $\overline{X}_{t-k}$ les deux membres de $X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$ et en prenant l'espérance.

Représentation canonique

- Théorème (représentation canonique)
  - Soit X_t un processus ARMA(p,q) solution de $X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$.
  - On suppose que $\phi(z) \neq 0$ et $\theta(z) \neq 0\ \forall |z| = 1$
  - Alors X_t admet une représentation canonique $X_t - \phi'_1 X_{t-1} - \ldots - \phi'_p X_{t-p} = Z'_t + \theta'_1 Z'_{t-1} + \ldots + \theta'_q Z'_{t-q}$.
- Théorème (innovations d'un processus ARMA)
  - Soit X_t un processus ARMA(p,q) dont la représentation canonique est $X_t - \phi_1 X_{t-1} - \ldots - \phi_p X_{t-p} = Z_t + \theta_1 Z_{t-1} + \ldots + \theta_q Z_{t-q}$.
  - Alors Z_t est le processus des innovations de X_t.
  - Preuve : $\varepsilon_t = X_t - \mathrm{proj}(X_t | \mathcal{H}^X_{t-1})$. Or $\theta(z)$ causal $\Rightarrow Z_t \in \mathcal{H}^X_t$, et $\phi(z)$ causal $\Rightarrow X_t \in \mathcal{H}^Z_t$, d'où $Z_t \perp \mathcal{H}^X_{t-1}$, dont on déduit $\mathrm{proj}(X_t | \mathcal{H}^X_{t-1})$.

Caractérisation d'un processus AR(p)

- Définition (Fonction d'autocorrélation partielle).
  - Soit $(X_t)_{t\in\mathbb{Z}}$ un processus centré SSL.
  - Les coefficients de prédiction $\phi_h = [\phi_{1,h}, \ldots, \phi_{h,h}]^T$ sont définis par l'égalité $\mathrm{proj}(X_t|\mathcal{H}^X_{t-1,h}) = \sum_{k=1}^h \phi_{k,h} X_{t-k}$.
  - Alors la suite $(u(h))_{h\geq 1}$, où $u(h) = \phi_{h,h}$, est appelée fonction d'autocorrélation partielle de X_t.
- Théorème (Caractérisation d'un processus AR(p)).
  - Soit $(X_t)_{t\in\mathbb{Z}}$ un processus centré SSL de fonction d'autocorrélation partielle $(u(h))_{h\geq 1}$, et soit $p \geq 1$. Alors les deux assertions suivantes sont équivalentes :
    - X_t est un processus AR d'ordre minimal p ;
    - $u(p) \neq 0$ et $u(h) = 0 \; \forall h \geq p+1$.