Préparation à l'agrégation
Epreuve de modélisation

Séries chronologiques I

On va étudier les méthodes de Box et Jenkins qui sont des méthodes très puissantes qui exploitent au maximum le fait que l'évolution de la série temporelle étudiée est considérée comme **une** des réalisations d'un processus stochastique doué d'une structure assez forte. En effet, la structure une fois mise en évidence permet de prédire avec plus d'assurance la suite de la série. La contrepartie est la nécessité d'une assez longue durée d'observations pour que la prévision soit fiable. Les auteurs préconisent de 5 à 6 **périodes** dans le cas de phénomènes périodiques, et sinon un **minimum** de 30 observations.

Ces méthodes donnent d'excellents résultats pour les prévisions à court terme des séries macro-économiques, en particulier pour les indices de production industrielle.

Elles sont fondées sur l'hypothèse que chaque observation dépend assez fortement des observations précédentes. En gros, cette dépendance du passé remplace la multiplicité d'observations en statistique qui permet d'estimer les paramètres en appliquant la loi des grands nombres.

1 Définitions

On considère des suites de variables aléatoires indexées dans $\mathbb{Z}$ et à valeurs dans $\mathbb{C}$, à savoir:

$$\forall n \in \mathbb{Z},\ X_n \text{ est une variable aléatoire } : (\Omega, \mathcal{F}, \mathbb{P}) \to (\mathbb{C}, \mathcal{B}) \ .$$

On cherche à modéliser l'application $n \mapsto X_n$ avec une partie tendancielle, une composante saisonnière et l'erreur de mesure.

Définition 1 *: On appelle* **série chronologique stationnaire** *(en abrégé SCS) une suite de variables aléatoires $(X_n, n \in \mathbb{Z})$ à valeurs complexes, centrées, de carré intégrable, telles que*

$$cov(X_n, X_{n-k}) = \mathbb{E}[X_n \overline{X_{n-k}}]$$

ne dépende que de k. On note $\gamma(k)$ cette valeur.

La fonction γ s'appelle la **fonction d'autocovariance***.*

On définit aussi la **fonction d'autocorrélation** $\rho(k) = \frac{\gamma(k)}{\gamma(0)}$.

2 Exemples

Définition 2 *On appelle* **bruit blanc** *une suite (ε_n) de variables aléatoires centrées de variance un et orthogonales.*

On a donc pour un bruit blanc:

$$\begin{array}{rcl} \gamma(0) & = & 1 \\ \gamma(k) & = & 0 \quad \text{pour } k \neq 0 \ . \end{array}$$

et il est facile de vérifier que dans ce cas

$$\gamma(n) = \frac{1}{2\pi} \int_0^{2\pi} e^{in\lambda} d\lambda.$$

Ce processus "bruit blanc" sert à modéliser l'erreur de mesure.

Si on se donne une suite $(Z_j, j = 1, \ldots, p)$ de variables aléatoires centrées orthogonales et une suite $(\lambda_j, j = 1, \ldots, p)$ de nombres réels dans l'intervalle $[-\pi, \pi[$, alors la suite

$$X_n = \sum_{j=1}^{p} Z_j e^{in\lambda_j}, \quad n \in \mathbb{Z}$$

est une SCS appelée **somme d'harmoniques**.

Définition 3 *Soit ε un bruit blanc et une suite $(a_k, k \in \mathbb{Z}) \in l^2(\mathbb{Z})$. La suite*

$$X_n = \sum_{k \in \mathbb{Z}} a_k \varepsilon_{n-k}$$

est appelée une **moyenne mobile**. *En abrégé on dira* **MA**.

Proposition 1 *La fonction de covariance d'une moyenne mobile $X_n = \sum_{k \in \mathbb{Z}} a_k \varepsilon_{n-k}$ s'écrit:*

$$\gamma(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} |\sum_{k \in \mathbb{Z}} a_k e^{-ik\lambda}|^2 e^{in\lambda} d\lambda \ .$$

Définition 4 *Si les coefficients a_k sont nuls pour tous les $k < 0$, on dit que la moyenne mobile est* **unilatère**. *S'il existe seulement un nombre fini de coefficients a_k non nuls, soit $(a_0, \ldots, a_p)$, on dit que X est une* **moyenne mobile d'ordre** *p, soit en abrégé* **MA(p)**.

Un exemple particulièrement important est donné par la moyenne mobile unilatère suivante.

Proposition 2 *Soit $\alpha \in \mathbb{C}$ vérifiant $|\alpha| < 1$ et soit ε un bruit blanc. La moyenne mobile unilatère*

$$X_n = \sum_{j=0}^{\infty} \alpha^j \varepsilon_{n-j}$$

vérifie la relation de récurrence

$$X_n = \alpha X_{n-1} + \varepsilon_n$$

et sa fonction de covariance vaut

$$\gamma(n) = \frac{1}{2\pi} \int_{-\pi}^{\pi} \frac{e^{in\lambda}}{|1 - \alpha e^{-i\lambda}|^2} d\lambda \ .$$

Définition 5 *On dit que le processus X précédent est un processus* **autorégressif d'ordre** *1, ou* **AR(1)**.

On note e_n la fonction $\lambda \mapsto e^{in\lambda}$ définie sur $\Pi = [-\pi, \pi[$.

Théorème 1 *A toute SCS on peut associer une unique mesure bornée sur Π telle que $\mu(e_n) = \gamma(n)$, $\forall n \in \mathbb{Z}$.*

Définition 6 *Cette mesure μ s'appelle la* **mesure spectrale** *de la SCS. Si elle est absolument continue par rapport à la mesure de Lebesgue, la densité $\frac{d\mu}{d\lambda}$ s'appelle la* **densité spectrale** *de la SCS.*

Si l'autocovariance $\gamma \in l^2(\mathbb{Z})$, on obtient par inversion de Fourier

$$f(\lambda) = \frac{1}{2\pi} \sum_{n \in \mathbb{Z}} e^{-in\lambda} \gamma(n)$$

dans l'espace $L^2(\Pi, d\lambda)$.

3 Opérateur retard, équations ARMA

Soit X une SCS.

Définition 7 *Le sous-espace vectoriel fermé engendré par l'ensemble $\{X_p, p \in \mathbb{Z}, p \leq n\}$ dans L^2 est noté H_n^X et s'appelle le* **passé linéaire** *de X. Le sous-espace fermé $H_{+\infty}^X = H^X$ engendré par toutes les variables $\{X_p, p \in \mathbb{Z}\}$ s'appelle l'***enveloppe linéaire** *de X.*

Pout tout $n \in \mathbb{Z}$, on note P_n^X la projection orthogonale de H^X sur H_n^X.

Proposition 3 *Il existe une unique isométrie, notée S^X, de H^X sur lui-même telle que pour tout $n \in \mathbb{Z}$ on ait $S^X(X_n) = X_{n-1}$. Cette isométrie s'appelle* **l'opérateur retard**.

Théorème 2 *S'il existe une suite $(d_n, n \in \mathbb{N})$ de $l^2(\mathbb{N})$ et un bruit blanc ε tels que*

$$X_n = \sum_{p \geq 0} d_p \varepsilon_{n-p} .$$

on peut choisir ε en sorte que les passés linéaires de X et ε coïncident et alors ce bruit blanc et la suite (d_n) associée sont uniques à un coefficient multiplicatif près.

Définition 8 *Ce bruit blanc s'appelle le* **bruit blanc d'innovation** *de X.*

Dans ce cas, la projection sur le passé est extrêmement simple: pour tout $m \leq n$,

$$P_m^X(X_n) = \sum_{p \geq n-m} d_p \varepsilon_{n-p} .$$

Définition 9 *Soit X une SCS et ε un bruit blanc, P et Q deux polynômes à coefficients dans $\mathbb{C}$. On dit que X est solution de l'***équation ARMA***(P,Q) si ce processus vérifie pour tout n de $\mathbb{Z}$*

$$P(S^X)(X_n) = Q(S^\varepsilon)(\varepsilon_n) .$$

C'est-à-dire qu'il existe des coefficients complexes $(a_0, \ldots, a_p)$ et $(b_0, \ldots, b_q)$ tels que pour tout n de $\mathbb{Z}$

$$\sum_{i=0}^{p} a_i X_{n-i} = \sum_{j=0}^{q} b_j \varepsilon_{n-j} .$$

Si $p = 0$, X est une **MA(q)**. Si $q = 0$, X est un **AR(p)**. Dans le cas général, on dit que X est un **ARMA(p,q)**.

4 Filtres linéaires

On peut filtrer une série X (pas forcément stationnaire), en lui appliquant un opérateur polynôme de l'opérateur retard S. Il s'agit par exemple d'éliminer un bruit, une tendance ou une composante saisonnière pour se ramener à une SCS.

4.1 Quelques exemples

- Pour éliminer une tendance polynomiale de la série $Y_n = f(n) + X_n$ où f est un polynôme de degré p et X une SCS, on applique le filtre $\Delta = I - S$. Alors Δf est un polynôme de degré $p - 1$ et ΔX reste une SCS. On dit que l'on a différencié la série. Si l'on applique Δ^{p+1} alors $\Delta^{p+1} f = 0$ et $\Delta^{p+1} Y = \Delta^{p+1} X$ est une SCS. Dans la pratique, on différencie la série jusqu'à obtenir une SCS.

- Pour éliminer une tendance périodique (désaisonnaliser) d'une série du type

$$Y_n = f(e^{2i\pi n/T}) + X_n$$

où X est une SCS et f s'appelle la composante saisonnière, on effectue une moyenne mobile; on peut en effet remarquer que

$$\forall n \ \ (AY)_n \equiv \frac{1}{T}\sum_{i=0}^{T-1} Y_{n-i} = a + (AX)_n$$

et donc ΔAY est une SCS. Ou bien on applique $B = I - S^T$ qui désaisonnalise aussi la série.

- Les moyennes mobiles finies sont aussi des filtres:

$$Y_n = \sum_{i=0}^{q} a_i X_{n-i} = (\sum_{i=0}^{q} a_i S^i)(X_n)$$

est une SCS si au départ X en est une.

- Une suite (Y_n) de variables aléatoires de carré intégrable est une série chronologique **ARIMA**(p,d,q) si la suite $(I-S)^d Y(n)$ est un ARMA(p,q).

- Une suite (Y_n) est un **ARMA à effets saisonniers** s'il existe un entier positif T tel que la suite $(I - S^T)Y$ soit un ARMA.

4.2 Fonctions réponse

Soit μ^X la mesure spectrale de la SCS X. On définit sur $vect(e_n)$ l'application linéaire

$$Z^X(\sum_{j\in I} a_j e_j) = \sum_{j\in I} a_j X_j$$

qui se prolonge en une isométrie de $L^2(\mu^X)$ sur H^X.

Théorème 3 *Soit X une SCS de mesure spectrale μ^X. Si $h \in L^2(\mu^X)$, la suite $Y_n = Z^X(he_n)$ est une SCS de mesure spectrale $\mu^Y = |h|^2\mu^X$.*

On dit que Y est l'image de X par le filtre de **fonction réponse** ou de **gain** h, ou encore que Y est le **filtré** de X par h.

Proposition 4 (Composition des filtres) *Soit X une SCS de mesure spectrale μ^X et $h \in L^2(\mu^X)$. On pose $\nu = |h|^2\mu^X$ et soit $g \in L^2(\nu)$. Si Y est le filtré de X par h, alors le filtré de X par hg et le filtré de Y par g coïncident.*

Proposition 5 *Soit Y le filtré de X par h. On a*

- $H^Y \subset H^X$
- $H^Y = H^X \ \Leftrightarrow \ |h| > 0 \ \mu^X\text{−}p.p.$
- *dans ce dernier cas, X est filtré de Y par $1/h$.*

Un filtre tel que pour tout $n \in \mathbb{Z}$ on ait $H_n^Y \subset H_n^X$ est dit **adapté** ou **réalisable**.

Proposition 6 *Soit $g = \sum_{k\geq 0} c_k e_{-k}$ avec $c = (c_k) \in l^2(\mathbb{N})$ et $c \neq 0$. Alors $|g| > 0$ p.p.*

5 Processus ARMA

Définition 10 *On dit qu'une SCS est à* **spectre rationnel** *si elle admet une densité spectrale de la forme*

$$\frac{1}{2\pi}\left|\frac{Q}{P}\right|^2(e^{-i\lambda}),$$

où P et Q sont deux polynômes réels et P n'a pas de racine de module un.

Proposition 7 *Soit X une SCS à spectre rationnel. Il existe à un coefficient multiplicatif près un unique couple irréductible (P, Q) tel que les zéros de P soient de module > 1, les zéros de Q de module ≥ 1 et la densité spectrale de X soit de la forme précédente.*

Cette dernière représentation s'appelle la représentation de **Féjer-Riesz canonique**.

Théorème 4 *Soit X une SCS à spectre rationnel et à densité spectrale strictement positive Si (P, Q) est une représentation de Féjer-Riesz canonique associée à sa densité spectrale, il existe un bruit blanc d'innovation ε tel que l'on ait la relation ARMA canonique*

$$P(S^X)(X_n) = Q(S^\varepsilon)(\varepsilon_n)\ .$$

Si X est un processus à spectre rationnel ayant (P, Q) pour représentation de Féjer-Riesz canonique avec degré(P)=p et degré(Q)=q, on dit que X est un **ARMA de type minimal (p,q)**.

6 Prévision des processus ARMA

Soit X une série chronologique à spectre rationnel de représentation de Féjer-Riesz canonique (P, Q). On s'intéresse à la prédiction à un pas $\hat{X}_{n+1} = P_n^X(X_{n+1})$ de X_{n+1}.

Théorème 5 *La suite $\hat{X}_{n+1}$ est une SCS filtrée de X par*

$$g(\lambda) = e^{i\lambda}[1 - \frac{Q}{P}(0)\frac{P}{Q}(e^{-i\lambda})]$$

et l'erreur de prédiction à un pas $\|X_{n+1} - \hat{X}_{n+1}\|_2$ vaut $|\frac{Q}{P}(0)|$.

En particulier si Q n'a pas non plus de zéro de module un, alors P/Q est holomorphe dans un voisinage du disque unité et le calcul de son développement en série entière nous donne l'expression de $\hat{X}_{n+1}$ en fonction des $(X_{n-k}, k \geq 0)$. Le cas autorégressif pur est tout à fait simple.

Proposition 8 *Soit X une SCS à spectre rationnel de représentation canonique $(P, 1)$ avec $P(z) = \sum_{j=0}^p a_j z^j$. Alors la prédiction de X_{n+1} à un pas vaut*

$$\hat{X}_{n+1} = -\frac{1}{a_0}\sum_{j=1}^p a_j X_{n+1-j}$$

et l'erreur de prédiction vaut $1/|a_0|$.

Exercices

Démontrer:

- la proposition 1;
- la proposition 2;
- la proposition 3;
- le théorème 3;
- la proposition 4;
- la proposition 5;
- la proposition 7;
- le théorème 4;
- le théorème 5;
- la proposition 8.