Examen de Processus Stochastiques
Jeudi 6 Janvier 2011

$1^{ère}$ Partie : Chaîne de Markov

1) Les chaînes stationnaires $\pi = (\pi_0, \pi_1, \pi_2)$ associées à la matrice $\mathbf{P}$ vérifient

$$\left\{ \begin{array}{c} \pi = \pi \mathbf{P} \\ \sum_{k=0}^{2} \pi_k = 1 \end{array} \right. \Longleftrightarrow \left\{ \begin{array}{c} \pi_0 = p\pi_2 \\ \pi_1 = \pi_2 \\ \pi_0 + \pi_1 + \pi_2 = 1 \end{array} \right.$$

On en déduit que la chaîne de Markov admet une seule loi stationnaire définie par

$$\pi_0 = \frac{p}{p+2}, \pi_1 = \frac{1}{p+2}, \pi_2 = \frac{1}{p+2}$$

2) Des calculs élémentaires conduisent à

$$\mathbf{P}^2 = \begin{pmatrix} 0 & 0 & 1 \\ p & q & 0 \\ 0 & p & q \end{pmatrix} \text{ et } \mathbf{P}^3 = \begin{pmatrix} p & q & 0 \\ 0 & p & q \\ qp & q^2 & p \end{pmatrix}$$

Comme la deuxième colonne de $\mathbf{P}^3$ est constituée d'éléments strictement positifs, le théorème du cours s'applique. Donc, le chaîne de Markov admet une seule loi limite qui est aussi sa loi stationnaire.

3) On suppose $p = \frac{1}{4}$. On a alors $q = \frac{3}{4}$ et

$$\mathbf{P}^3 = \begin{pmatrix} \frac{1}{4} & \frac{3}{4} & 0 \\ 0 & \frac{1}{4} & \frac{3}{4} \\ \frac{3}{16} & \frac{9}{16} & \frac{1}{4} \end{pmatrix}$$

On peut donc appliquer le théorème du cours avec $n_0 = 3$ et

$$\delta = \frac{1}{4} \text{ pour la colonne 2, i.e. } s = 1$$

d'où

$$\|\mathbf{P}^n - \mathbf{P}^\infty\| \leq \left(1 - \frac{1}{4}\right)^{\frac{n}{3}-1} = \left(\frac{3}{4}\right)^{\frac{n}{3}-1}$$

Pour avoir $\|\mathbf{P}^n - \mathbf{P}^\infty\| \leq 10^{-2}$, il suffit donc de choisir n tel que

$$\left(\frac{3}{4}\right)^{\frac{n}{3}-1} \leq 10^{-2}$$

En prenant le logarithme décimal, on obtient

$$\left(\frac{n}{3}-1\right)\log_{10}\left(\frac{3}{4}\right)\leq -2$$

d'où

$$n\geq 3\left[1+\frac{-2}{\log_{10}\left(\frac{3}{4}\right)}\right]\simeq 51.02$$

Pour avoir $\|\mathbf{P}^n-\mathbf{P}^\infty\|\leq 10^{-2}$, il suffit donc de prendre

$$n>51$$

4) Pour $p=\frac{1}{4}$, les valeurs propres de la matrice $\mathbf{P}$ se déterminent facilement par résolution de

$$\begin{aligned}\begin{vmatrix}-\lambda & 1 & 0\\ 0 & -\lambda & 1\\ \frac{1}{4} & \frac{3}{4} & -\lambda\end{vmatrix} &= 0 \Longleftrightarrow -\lambda\left(\lambda^2-\frac{3}{4}\right)-\frac{1}{4}=0\\ &\Longleftrightarrow -\lambda^3+\frac{3}{4}\lambda-\frac{1}{4}=0\end{aligned}$$

D'après le cours, on sait que $\lambda=1$ est solution de cette équation. On en déduit

$$\begin{aligned}-\lambda^3+\frac{3}{4}\lambda-\frac{1}{4} &= (\lambda-1)\left(-\lambda^2-\lambda-\frac{1}{4}\right)\\ &= -(\lambda-1)\left(\lambda+\frac{1}{2}\right)^2\end{aligned}$$

La matrice $\mathbf{P}$ admet donc la valeur propre simple $\lambda=1$ et la valeur propre double $\lambda=-\frac{1}{2}$, ce qui permet d'appliquer le théorème. Ceci confirme le résultat obtenu à la question 2).
5) Pour $p=1$, que la matrice $\mathbf{P}$ s'écrit

$$\mathbf{P}=\begin{pmatrix}0&1&0\\0&0&1\\1&0&0\end{pmatrix}$$

d'où les puissances successives de $\mathbf{P}$

$$\mathbf{P}^2=\begin{pmatrix}0&0&1\\1&0&0\\0&1&0\end{pmatrix},\mathbf{P}^3=\begin{pmatrix}1&0&0\\0&1&0\\0&0&1\end{pmatrix}=I_3$$

On en déduit

$$\mathbf{P}^{3n}=I_3,\mathbf{P}^{3n+1}=\mathbf{P}\text{ et }\mathbf{P}^{3n+2}=\mathbf{P}^2$$

On voit donc que les puissances $\mathbf{P}$ prennent les valeurs successives

$$I_3,\mathbf{P},\mathbf{P}^2,I_3,\mathbf{P},\mathbf{P}^2,...$$

On a donc une suite alternée qui ne converge pas.

6) Les lois stationnaires de la chaîne de Markov de matrice de transition

$$\mathbf{Q} = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & 0 & \dots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & 0 & 1 \\ p & q & 0 & \dots & \dots & 0 \end{pmatrix}$$

sont définies par

$$\begin{cases} \pi = \pi\mathbf{Q} \\ \sum_{k=0}^{m-1} \pi_k = 1 \end{cases} \Longleftrightarrow \begin{cases} \pi_0 = p\pi_{m-1} \\ \pi_1 = \pi_0 + q\pi_{m-1} \\ \pi_2 = \pi_1 \\ \pi_3 = \pi_2 \\ \vdots \\ \pi_{m-1} = \pi_{m-2} \\ \sum_{k=0}^{m-1} \pi_k = 1 \end{cases}$$

On en déduit

$$\begin{cases} \pi_1 = \pi_2 = \dots = \pi_{m-1} \\ \pi_0 = p\pi_{m-1} \\ \sum_{k=0}^{m-1} \pi_k = 1 \end{cases}$$

La dernière relation donne alors

$$\pi_{m-1} = \frac{1}{p+m-1}$$

d'où

$$\pi = \left(\frac{p}{p+m-1}, \frac{1}{p+m-1}, ..., \frac{1}{p+m-1} \right)$$

Pour montrer que cette chaîne de Markov admet une unique loi limite, on peut chercher une puissance de P qui admet une colonne d'éléments tous strictement positifs. D'après le théorème du cours, on en déduit alors que l'unique loi stationnaire est aussi la loi limite.

$\mathbf{2}^{ème}$ Partie : Suite stationnaire

On considère la suite stationnaire AR définie par l'équation récurrente

$$y_t = c + \frac{7}{12} y_{t-1} - \frac{1}{12} y_{t-2} + e_t, \qquad t \in \mathbb{Z}$$

où $\{e_t, t \in \mathbb{Z}\}$ est un bruit blanc (de moyenne nulle) σ_e^2.
1) En remplaçant y_t par $x_t + m$, on obtient

$$x_t + m = c + \frac{7}{12} x_{t-1} + \frac{7}{12} m - \frac{1}{12} x_{t-2} - \frac{1}{12} m$$

Pour avoir l'équation récurrente demandée, il suffit donc d'avoir

$$m = c + \frac{7}{12} m - \frac{1}{12} m \text{ d'où } m = 2c$$

2) A partir de l'équation

$$x_t = \frac{7}{12}x_{t-1} - \frac{1}{12}x_{t-2} + e_t$$

on obtient

$$X(z)\left(1 - \frac{7}{12}z^{-1} + \frac{1}{12}z^{-2}\right) = E(z)$$

qui montre que la suite $\{x_t, t \in \mathbb{Z}\}$ peut être obtenue par un filtre linéaire d'entrée $\{e_t, t \in \mathbb{Z}\}$ et de transmittance

$$H(z) = \frac{X(z)}{E(z)} = \frac{1}{1 - \frac{7}{12}z^{-1} + \frac{1}{12}z^{-2}}$$

Des calculs élémentaires permettent d'obtenir

$$\frac{1}{1 - \frac{7}{12}z^{-1} + \frac{1}{12}z^{-2}} = \frac{1}{\left(1 - \frac{1}{3}z^{-1}\right)\left(1 - \frac{1}{4}z^{-1}\right)}$$

La fonction de transfert $H(z)$ admet donc deux pôles $z = \frac{1}{3}$ et $z = \frac{1}{4}$ situés à l'intérieur du cercle unité. La relation reliant x_t et e_t est donc causale.

3) On sait que l'innovation u_t peut être obtenue par filtrage linéaire de x_t. Si on appelle $\beta(z)$ la fonction de transfert de ce filtre, on a

$$s_u(f) = 1 = |\beta(f)|^2\, s_x(f)$$

d'où

$$|\beta(f)|^2 = \frac{1}{s_x(f)} = \frac{\left|\left(1 - \frac{1}{3}e^{-j2\pi f}\right)\left(1 - \frac{1}{4}e^{-j2\pi f}\right)\right|^2}{\sigma_e^2}$$

On peut décomposer $1/s_x(f)$ comme suit

$$\frac{1}{s_x(f)} = \frac{1}{\sigma_e^2}\left(1 - \frac{1}{3}e^{-j2\pi f}\right)\left(1 - \frac{1}{3}e^{j2\pi f}\right)\left(1 - \frac{1}{4}e^{-j2\pi f}\right)\left(1 - \frac{1}{4}e^{j2\pi f}\right)$$

En choisissant les zéros situés à l'intérieur du cercle unité on obtient

$$\beta(f) = \frac{1}{\sigma_e}\left(1 - \frac{1}{3}e^{-j2\pi f}\right)\left(1 - \frac{1}{4}e^{-j2\pi f}\right)$$

d'où

$$\beta(z) = \frac{1}{\sigma_e}\left(1 - \frac{7}{12}z^{-1} + \frac{1}{12}z^{-2}\right)$$

et donc

$$u_t = \frac{1}{\sigma_e}\left(x_t - \frac{7}{12}x_{t-1} + \frac{1}{12}x_{t-2}\right) = \frac{e_t}{\sigma_e}$$

On retrouve le résultat classique selon lequel dans le système causal, l'innovation est égale au bruit blanc d'entrée normalisé. La décomposition de Wold de x_t peut se déterminer facilement à l'aide de

$$\begin{aligned}
X(z) &= \frac{1}{\beta(z)}U(z) \\
&= \frac{\sigma_e}{\left(1 - \frac{1}{3}z^{-1}\right)\left(1 - \frac{1}{4}z^{-1}\right)}U(z) \\
&= \sigma_e\left[\frac{A}{1 - \frac{1}{3}z^{-1}} + \frac{B}{1 - \frac{1}{4}z^{-1}}\right]U(z)
\end{aligned}$$

Des calculs élémentaires permettent d'obtenir

$$A = 4 \text{ et } B = -3$$

d'où

$$X(z) = \sigma_e \left[4 \sum_{n=0}^{\infty} \left(\frac{1}{3}\right)^n z^{-n} - 3 \sum_{n=0}^{\infty} \left(\frac{1}{4}\right)^n z^{-n} \right] U(z)$$

La décomposition de Wold de x_t est donc

$$x_t = \sigma_e \sum_{n=0}^{\infty} \left[4 \left(\frac{1}{3}\right)^n - 3 \left(\frac{1}{4}\right)^n \right] u_{t-n}$$

On en déduit

$$y_t = x_t + m = 2c + \sum_{n=0}^{\infty} \left[4 \left(\frac{1}{3}\right)^n - 3 \left(\frac{1}{4}\right)^n \right] e_{t-n}$$

4) En utilisant la décomposition de Wold, on obtient

$$x_{t+k} = \sigma_e \sum_{n=0}^{\infty} \left[4 \left(\frac{1}{3}\right)^n - 3 \left(\frac{1}{4}\right)^n \right] u_{t+k-n}$$

d'où

$$\widehat{x}_{t+k} = \sum_{n=k}^{\infty} \alpha_n e_{t+k-n} \text{ avec } \alpha_n = 4 \left(\frac{1}{3}\right)^n - 3 \left(\frac{1}{4}\right)^n$$

L'erreur d'approximation est

$$\widehat{x}_{t+k} - x_{t+k} = \sum_{n=0}^{k-1} \alpha_n e_{t+k-n}$$

d'où en utilisant l'orthogonalité des entrées e_{t+k-n}

$$\sigma_k^2 = E\left[|\widehat{x}_{t+k} - x_{t+k}|^2 \right] = \sigma_e^2 \sum_{n=0}^{k-1} \alpha_n^2$$

La meilleure prédiction linéaire de y_{t+k} notée $\widehat{y}_{t+k}$ sur l'espace engendré par son passé $y_t, y_{t-1}, ...$ est

$$\widehat{y}_{t+k} = \widehat{x}_{t+k} + m$$

d'où l'erreur d'approximation

$$E\left[|\widehat{y}_{t+k} - y_{t+k}|^2 \right] = \sigma_k^2$$

$3^{ème}$ Partie : Méthodes de Simulation

1) La vraisemblance associée à l'estimation des paramètres a et b est définie par

$$\begin{aligned} f(x_1,...,x_n|\,\boldsymbol{\theta}) &= \prod_{i=1}^{n} f(x_i|\,\boldsymbol{\theta}) \\ &= \left[\frac{b^a}{\Gamma(a)}\right]^n \left(\prod_{i=1}^{n} x_i^{a-1}\right) \exp\left(-b\sum_{i=1}^{n} x_i\right) \end{aligned}$$

On en déduit la loi a posteriori du vecteur paramètre $\boldsymbol{\theta} = (a, b)$

$$\begin{aligned} f(\boldsymbol{\theta}|\,x_1,...,x_n) &\propto f(x_1,...,x_n|\,\boldsymbol{\theta})f(a)f(b|\,\alpha,\beta) \\ &\propto \frac{b^{an+\alpha-1}}{[\Gamma(a)]^n}\left(\prod_{i=1}^{n} x_i^{a-1}\right) \exp\left[-b\left(\beta+\sum_{i=1}^{n} x_i\right)\right] \mathbb{I}_{[0,1]}(a)\,\mathbb{I}_{\mathbb{R}^+}(b) \end{aligned}$$

2) Les lois conditionnelles de $f(a|\,b,x_1,...,x_n)$ et $f(b|\,a,x_1,...,x_n)$ vérifient

$$\begin{aligned} f(a|\,b,x_1,...,x_n) &\propto \frac{\left[b^n\prod_{i=1}^{n} x_i\right]^a}{[\Gamma(a)]^n}\mathbb{I}_{[0,1]}(a) \\ f(b|\,a,x_1,...,x_n) &\propto b^{an+\alpha-1}\exp\left[-b\left(\beta+\sum_{i=1}^{n} x_i\right)\right]\mathbb{I}_{\mathbb{R}^+}(b) \end{aligned}$$

On remarque que la loi conditionnelle de $b|\,a,x_1,...,x_n$ est une loi gamma de paramètres $an+\alpha$ et $\beta+\sum_{i=1}^{n} x_i$, c'est-à-dire une loi

$$\Gamma\left(an+\alpha, \beta+\sum_{i=1}^{n} x_i\right)$$

3) Le principe de l'échantillonneur de Gibbs est de simuler des variables distribuées suivant les lois conditionnelles de la loi d'intérêt. Dans le cas présent, il peut être décrit de la façon suivante

Initialisation

Générer a et b par exemple suivant leur lois a priori et noter $a^{(0)}$ et $b^{(0)}$ les valeurs obtenues

Itérations

Pour $k = 1,...,K$, effectuer les simulations suivantes

- Générer suivant $f(a|\,b^{(k-1)},x_1,...,x_n)$
- Générer suivant $f(b|\,a^{(k)},x_1,...,x_n)$

  La deuxième simulation est simple puisque la loi de $b|\,a,x_1,...,x_n)$ est une loi gamma. Il est moins simple de simuler suivant la loi conditionnelle de $a|\,b,x_1,...,x_n$. On décide d'utiliser une étape de Metropolis-Hastings pour cette simulation avec une loi de proposition uniforme sur l'intervalle $[0,1]$. Le principe de cette étape de Metropolis-Hastings est à l'itération k de générer un candidat $z^{(k)}$ suivant la loi uniforme sur l'intervalle $[0,1]$ et d'accepter ce candidat avec la probabilité

$$\min\left[1, \frac{f(z^k|\,b^{(k-1)},x_1,...,x_n)}{f(a^{(k-1)}|\,b^{(k-1)},x_1,...,x_n)}\right]$$

4) Après avoir simulé K vecteurs notés $\boldsymbol{\theta}^{(k)}, k = 1, ..., K$ distribués suivant $f(\boldsymbol{\theta}|\, x_1, ..., x_n)$, on peut estimer les paramètres a et b à l'aide des estimateurs MAP ou MMSE. Si la période de chauffage de la chaîne de Markov est de longueur k_0 et qu'on décide d'utiliser les estimateurs MMSE, on obtiendra

$$\begin{aligned}
\widehat{a}_{\text{MMSE}} &= \frac{1}{K-k_0} \sum_{k=k_0+1}^{K} a^{(k)} \\
\widehat{b}_{\text{MMSE}} &= \frac{1}{K-k_0} \sum_{k=k_0+1}^{K} b^{(k)}
\end{aligned}$$

La période de chauffage (burn-in) de l'échantillonneur de Gibbs est la période qu'il faut attendre pour que les vecteurs simulés $\boldsymbol{\theta}^{(k)} = \left[a^{(k)}, b^{(k)}\right]$ soient distribués suivant la loi cible.