Examen de Processus Stochastiques
Jeudi 14 Janvier 2009

$1^{ère}$ Partie : Chaîne de Markov

On considère une chaîne de Markov à deux états 0 et 1 dont la matrice de transition est définie par

$$\mathbf{P} = \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix}$$

où a et b appartiennent tous deux à l'intervalle $]0,1[$.
1) Les chaînes stationnaires $\pi = (\pi_0, \pi_1)$ associées à la matrice $\mathbf{P}$ vérifient

$$\left\{ \begin{array}{c} \pi = \pi \mathbf{P} \\ \sum_{k=0}^{1} \pi_k = 1 \end{array} \right. \iff \left\{ \begin{array}{c} \pi_0 = (1-a)\pi_0 + b\pi_1 \\ \pi_1 = a\pi_0 + (1-b)\pi_1 \\ \pi_0 + \pi_1 = 1 \end{array} \right.$$

On en déduit

$$\left\{ \begin{array}{c} a\pi_0 = b\pi_1 \\ \pi_0 + \pi_1 = 1 \end{array} \right. \iff \left\{ \begin{array}{c} \pi_0 = 1 - \pi_1 \\ a(1-\pi_1) = b\pi_1 \end{array} \right.$$

Comme $a + b > 0$, on en déduit qu'il existe une seule solution stationnaire définie par

$$\pi_1 = \frac{a}{a+b}, \pi_0 = \frac{b}{a+b}$$

Puisque a et b sont tous deux des éléments de l'intervalle $]0,1[$, les deux colonnes de $\mathbf{P}$ sont constitués d'éléments non nuls et donc le théorème du cours nous permet d'affirmer que la loi stationnaire de la chaîne de Markov est aussi sa loi limite.
2) Supposons que

$$\mathbf{P}^n = \frac{1}{a+b} \begin{pmatrix} b + a\,(1-a-b)^n & a - a\,(1-a-b)^n \\ b - b\,(1-a-b)^n & a + b\,(1-a-b)^n \end{pmatrix}$$

alors

$$\begin{aligned} \mathbf{P}^{n+1} &= \mathbf{P}\mathbf{P}^n \\ &= \frac{1}{a+b} \begin{pmatrix} 1-a & a \\ b & 1-b \end{pmatrix} \begin{pmatrix} b + a\,(1-a-b)^n & a - a\,(1-a-b)^n \\ b - b\,(1-a-b)^n & a + b\,(1-a-b)^n \end{pmatrix} \\ &= \frac{1}{a+b} \begin{pmatrix} b + a\,(1-a-b)^{n+1} & a - a\,(1-a-b)^{n+1} \\ b - b\,(1-a-b)^{n+1} & a + b\,(1-a-b)^{n+1} \end{pmatrix} \end{aligned}$$

De plus, pour $n=0$, on a

$$\begin{aligned}
\frac{1}{a+b}\begin{pmatrix} b+a\left(1-a-b\right)^n & a-a\left(1-a-b\right)^n \\ b-b\left(1-a-b\right)^n & a+b\left(1-a-b\right)^n \end{pmatrix} &= \frac{1}{a+b}\begin{pmatrix} b+a & a-a \\ b-b & a+b \end{pmatrix} \\
&= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}
\end{aligned}$$

Donc, nous avons démontré par récurrence que pour tout $n \in \mathbb{N}$

$$\mathbf{P}^n = \frac{1}{a+b}\begin{pmatrix} b+a\left(1-a-b\right)^n & a-a\left(1-a-b\right)^n \\ b-b\left(1-a-b\right)^n & a+b\left(1-a-b\right)^n \end{pmatrix}.$$

Si on note $\mathbf{q}\left(n\right) = \left(P\left[x_n=0\right], P\left[x_n=1\right]\right)$ où x_n est l'état de la chaîne de Markov à l'instant n, on a

$$\mathbf{q}\left(n\right) = \mathbf{q}\left(0\right)\mathbf{P}^n,$$

donc

$$\lim_{n\to\infty}\mathbf{q}\left(n\right) = \mathbf{q}\left(0\right)\lim_{n\to\infty}\mathbf{P}^n$$

Mais, puisque $a \in \left]0,1\right[$ et $b \in \left]0,1\right[$, on a

$$0 < a+b < 2 \Longleftrightarrow -1 < 1-a-b < 1,$$

d'où

$$\lim_{n\to\infty}\mathbf{P}^n = \frac{1}{a+b}\begin{pmatrix} b & a \\ b & a \end{pmatrix}.$$

Si on pose $\mathbf{q}\left(0\right) = \left(u, 1-u\right)$, on obtient

$$\begin{aligned}
\lim_{n\to\infty}\mathbf{q}\left(n\right) &= \mathbf{q}\left(0\right)\lim_{n\to\infty}\mathbf{P}^n \\
&= \frac{\left(u,1-u\right)}{a+b}\begin{pmatrix} b & a \\ b & a \end{pmatrix} \\
&= \left(\frac{b}{a+b}, \frac{a}{a+b}\right).
\end{aligned}$$

On retrouve donc la loi limite obtenue à la question précédente.
3) Pour $a=\frac{1}{2}$ et $b=\frac{1}{3}$, on a

$$\mathbf{P} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \end{pmatrix}$$

On peut donc appliquer le théorème du cours avec $n_0 = 1$ et

$$\begin{aligned}
\delta &= \frac{1}{2} \text{ pour la colonne 2, i.e. } s=1 \\
\delta &= \frac{1}{3} \text{ pour les colonnes 1 et 2, i.e. } s=2
\end{aligned}$$

d'où

$$\begin{aligned}
\left\|\mathbf{P}^n - \mathbf{P}^\infty\right\| &\leq \left(1-\frac{1}{2}\right)^{n-1} = \left(\frac{1}{2}\right)^{n-1} \\
&\text{et} \\
\left\|\mathbf{P}^n - \mathbf{P}^\infty\right\| &\leq \left(1-\frac{2}{3}\right)^{n-1} = \left(\frac{1}{3}\right)^{n-1}
\end{aligned}$$

Le meilleur majorant est donc

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

4) Pour $a = b = 1$, on a

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

donc

$$\begin{aligned} \mathbf{q}(1) &= \mathbf{q}(0)\,\mathbf{P} \\ &= (u, 1-u)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= (1-u, u). \end{aligned}$$

De même

$$\begin{aligned} \mathbf{q}(2) &= \mathbf{q}(1)\,\mathbf{P} \\ &= (1-u, u)\begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \\ &= (u, 1-u) \end{aligned}$$

On en déduit pour tout $n \in \mathbb{N}$

$$\begin{aligned} \mathbf{q}(2n) &= \mathbf{q}(0) = (u, 1-u) \\ \mathbf{q}(2n+1) &= (1-u, u). \end{aligned}$$

Comme $u \neq \frac{1}{2}$, la suite des valeurs de $\mathbf{q}(n)$ est une suite alternée dont les valeurs sont successivement $(u, 1-u)$ et $(1-u, u)$.
5) Pour $a = 0$ et $b \in]0, 1[$, on a

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

La chaîne possède un unique état absorbant qui est le premier état $e_1 = 0$. On suppose $\mathbf{q}(0) = (0, 1)$ de façon à ce que l'état initial de la chaîne soit $x_0 = 1$. Le temps d'absorption T est alors tel que pour tout $n \geq 1$

$$\begin{aligned} P[T = n] &= P[x_n = e_1, x_{n-1} \neq e_1] \\ &= P[x_n = e_1, x_{n-1} = e_2] \\ &= P[x_n = e_1 \,|x_{n-1} = e_2]\, P[x_{n-1} = e_2] \\ &= bP[x_{n-1} = e_2] \end{aligned}$$

Mais

$$\mathbf{P}^n = \begin{pmatrix} 1 & 0 \\ 1-(1-b)^n & (1-b)^n \end{pmatrix}$$

donc

$$\begin{aligned} \mathbf{q}(n) &= \mathbf{q}(0)\,\mathbf{P}^n \\ &= (0, 1)\begin{pmatrix} 1 & 0 \\ 1-(1-b)^n & (1-b)^n \end{pmatrix} \\ &= (1-(1-b)^n, (1-b)^n) \end{aligned}$$

On en déduit $P[x_{n-1}=e_2]=(1-b)^{n-1}$ et donc

$$P[T=n]=b(1-b)^{n-1},n\geq 1$$

La moyenne et la variance de T vérifient alors

$$\begin{aligned} E[T] &= b\sum_{n=1}^{\infty} n(1-b)^{n-1} \\ E\left[T^2\right] &= b\sum_{n=1}^{\infty} n^2(1-b)^{n-1} \end{aligned}$$

En utilisant les relations suivantes (valables pour $x\in\left]-1,+1\right[$)

$$\begin{aligned} \sum_{n=1}^{\infty} nx^{n-1} &= \frac{1}{(1-x)^2} \\ \sum_{n=1}^{\infty} n^2x^{n-1} &= \sum_{n=1}^{\infty}[n(n-1)+n]\,x^{n-1}=\frac{2x}{(1-x)^3}+\frac{1}{(1-x)^2} \end{aligned}$$

on en déduit

$$\begin{aligned} E[T] &= b\times\frac{1}{b^2}=\frac{1}{b} \\ E\left[T^2\right] &= b\left[\frac{2(1-b)}{b^3}+\frac{1}{b^2}\right]=\frac{2}{b^2}-\frac{1}{b} \\ \operatorname{var}[T] &= E\left[T^2\right]-E[T]^2=\frac{1}{b^2}-\frac{1}{b} \end{aligned}$$

$\boxed{\mathbf{2}^{\grave{e}me}\ \textbf{Partie : Suite stationnaire}}$

On considère le système d'équations récurrentes

$$\begin{aligned} y_t &= by_{t-1}+ax_t+\varepsilon_t, \qquad t\in\mathbb{Z} \\ x_t &= cx_{t-1}+e_t, \qquad t\in\mathbb{Z} \end{aligned}$$

où $\{\varepsilon_t,t\in\mathbb{Z}\}$ et $\{e_t,t\in\mathbb{Z}\}$ sont des bruits blancs (de moyennes nulles) et de variance respectives σ_ε^2 et σ_e^2. On supposera de plus que les suites $\{\varepsilon_t,t\in\mathbb{Z}\}$ et $\{e_t,t\in\mathbb{Z}\}$ vérifient la propriété de décorrélation

$$E[\varepsilon_t e_u]=0, \qquad \forall t,\forall u$$

1) Comme $\{\varepsilon_t,t\in\mathbb{Z}\}$ et $\{e_t,t\in\mathbb{Z}\}$ sont des bruits blancs, les autocorrélations $r_\omega(h)=E[\omega_t\omega_{t+h}]$ de la suite

$$\omega_t=ae_t+\varepsilon_t-c\varepsilon_{t-1}, \qquad t\in\mathbb{Z}$$

sont nulles pour tout h tel que $|h|\geq 2$. De plus, des calculs élémentaires conduisent aux résultats suivants

$$\begin{aligned} r_\omega(0) &= a^2\sigma_e^2+\left(1+c^2\right)\sigma_\varepsilon^2 \\ r_\omega(1) &= r_\omega(-1)=-c\sigma_\varepsilon^2. \end{aligned}$$

2) Les autocorrélation déterminées à la question précédentes sont celles d'une suite MA(1). Pour déterminer θ, il suffit de calculer les autocorrélations de la suite

$$\omega_t = u_t - \theta u_{t-1}, \qquad t \in \mathbb{Z}$$

On obtient alors

$$\begin{aligned} r_\omega(0) &= \left(1+\theta^2\right)\sigma_u^2 \\ r_\omega(1) &= r_\omega(-1) = -\theta\sigma_u^2 \end{aligned}$$

d'où

$$\left\{ \begin{array}{c} a^2\sigma_e^2 + (1+c^2)\,\sigma_\varepsilon^2 = (1+\theta^2)\,\sigma_u^2 \\ -c\sigma_\varepsilon^2 = -\theta\sigma_u^2 \end{array} \right.$$

La seconde équation permet d'obtenir

$$\sigma_u^2 = \frac{c}{\theta}\sigma_\varepsilon^2$$

En remplaçant σ_u^2 dans la première équation, on montre que θ vérifie l'équation du second degré suivante

$$c\sigma_\varepsilon^2\theta^2 - \left[a^2\sigma_e^2 + \left(1+c^2\right)\sigma_\varepsilon^2\right]\theta + c\sigma_\varepsilon^2 = 0.$$

3) En prenant la transformée en Z du système d'équations

$$\begin{aligned} y_t &= by_{t-1} + ax_t + \varepsilon_t, \qquad t \in \mathbb{Z} \\ x_t &= cx_{t-1} + e_t, \qquad t \in \mathbb{Z} \end{aligned}$$

on obtient

$$\begin{aligned} Y(z)(1-bz^{-1}) &= aX(z) + \epsilon(z) \\ X(z)(1-cz^{-1}) &= E(z) \end{aligned}$$

d'où

$$\begin{aligned} Y(z) &= \frac{aX(z)}{1-bz^{-1}} + \frac{\epsilon(z)}{1-bz^{-1}} \\ X(z) &= \frac{E(z)}{1-cz^{-1}} \end{aligned}$$

et finalement

$$\begin{aligned} Y(z) &= \frac{aE(z) + (1-cz^{-1})\,\epsilon(z)}{(1-bz^{-1})\,(1-cz^{-1})} \\ &= \frac{\omega(z)}{(1-bz^{-1})\,(1-cz^{-1})} \\ &= \frac{(1-\theta z^{-1})\,U(z)}{(1-bz^{-1})\,(1-cz^{-1})} \end{aligned}$$

qui est la TZ d'une suite ARMA(2,1) d'entrée u_t et de fonction de transfert

$$H(z) = \frac{1-\theta z^{-1}}{(1-bz^{-1})\,(1-cz^{-1})}$$

L'équation de récurrence d'une telle suite s'obtient en prenant la TZ inverse de

$$\left(1-bz^{-1}\right)\left(1-cz^{-1}\right)Y(z)=\left(1-\theta z^{-1}\right)U(z)$$

On obtient

$$y_t-(b+c)\,y_{t-1}+bcy_{t-2}=u_t-\theta u_{t-1},\qquad t\in\mathbb{Z}$$

4) La densité spectrale de puissance de la suite y_t s'obtient à l'aide de la relation de Wiener Lee

$$\begin{aligned} s_y(f) &= s_u(f)\,|H(f)|^2 \\ &= \sigma_u^2\frac{\left|1-\theta e^{-j2\pi f}\right|^2}{\left|1-be^{-j2\pi f}\right|^2\left|1-ce^{-j2\pi f}\right|^2} \end{aligned}$$

5) On suppose $b<1,\theta<1$ et $c>1$. On sait que l'innovation i_t peut être obtenue par filtrage linéaire de y_t. Si on appelle $\beta(z)$ la fonction de transfert de ce filtre, on a

$$s_i(f)=1=|\beta(f)|^2\,s_y(f)$$

d'où

$$|\beta(f)|^2=\frac{1}{s_y(f)}$$

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

$$\frac{1}{s_y(f)}=\frac{c^2}{\sigma_u^2}\frac{\left(1-be^{-j2\pi f}\right)\left(1-be^{j2\pi f}\right)\left(1-\frac{1}{c}e^{-j2\pi f}\right)\left(1-\frac{1}{c}e^{j2\pi f}\right)}{\left(1-\theta e^{-j2\pi f}\right)\left(1-\theta e^{j2\pi f}\right)}$$

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

$$\begin{aligned} \beta(z) &= \frac{c}{\sigma_u}\frac{\left(1-bz^{-1}\right)\left(1-\frac{1}{c}z^{-1}\right)}{1-\theta z^{-1}} \\ &= \frac{\left(1-bz^{-1}\right)\left(c-z^{-1}\right)}{\sigma_u\left(1-\theta z^{-1}\right)} \\ &= \frac{1}{\sigma_u}\frac{c-(bc+1)\,z^{-1}+bz^{-2}}{1-\theta z^{-1}} \\ &= \frac{1}{\sigma_u}\left[K-\frac{b}{\theta}z^{-1}+\frac{c-K}{1-\theta z^{-1}}\right] \end{aligned}$$

avec

$$K=\frac{1+bc}{\theta}-\frac{b}{\theta^2}$$

c'est-à-dire

$$I(z)=\frac{1}{\sigma_u}\left[K-\frac{b}{\theta}z^{-1}\right]Y(z)+\frac{c-K}{\sigma_u}\left(\sum_{n=0}^{\infty}\theta^n z^{-n}\right)Y(z)$$

Donc, l'innovation i_t s'écrit

$$i_t=\frac{c}{\sigma_u}y_t+\left[\frac{c\,(\theta-b)-1}{\sigma_u}\right]y_{t-1}+\left(\frac{c-K}{\sigma_u}\right)\sum_{n>1}^{\infty}\theta^n y_{t-n}$$

La décomposition de Wold de y_t s'obtient en utilisant la relation inverse

$$\begin{aligned} Y(z) &= \frac{1}{\beta(z)} I(z) \\ &= \frac{\sigma_u \left(1 - \theta z^{-1}\right)}{\left(1 - b z^{-1}\right)\left(c - z^{-1}\right)} I(z) \end{aligned}$$

En décomposant en éléments simples cette fraction rationnelle, on obtient

$$\begin{aligned} \frac{1}{\beta(z)} &= \frac{\sigma_u (\theta - b)}{1 - bc} \frac{1}{1 - b z^{-1}} + \frac{\sigma_u (1 - \theta c)}{1 - bc} \frac{1}{c - z^{-1}} \\ &= \frac{\sigma_u}{1 - bc} \left[(\theta - b) \sum_{n=0}^{\infty} b^n z^{-n} + \left(\frac{1}{c} - \theta \right) \sum_{n=0}^{\infty} \left(\frac{1}{c} \right)^n z^{-n} \right] \end{aligned}$$

On pose

$$\alpha_n(b, c, \theta) = \frac{\sigma_u}{1 - bc} \left[(\theta - b) b^n + \left(\frac{1}{c} - \theta \right) \left(\frac{1}{c} \right)^n \right]$$

et on obtient alors

$$\alpha(z) = \frac{1}{\beta(z)} = \sum_{n=0}^{\infty} \alpha_n(b, c, \theta) z^{-n}$$

d'où

$$y_t = \sum_{n=0}^{\infty} \alpha_n(b, c, \theta) i_{t-n}$$

Pour simplifier, on notera $\alpha_n = \alpha_n(b, c, \theta)$ dans la dernière question.
6) La meilleure prédiction linéaire de y_t sur l'espace engendré par son passé $y_{t-1}, y_{t-2}, \ldots$ s'obtient à l'aide de la décomposition de Wold

$$\widehat{y}_t = \sum_{n=1}^{\infty} \alpha_n i_{t-n}$$

L'erreur d'approximation est

$$\sigma^2 = E\left[\left| \widehat{y}_t - y_t \right|^2 \right] = \alpha_0^2$$

La loi de $(y_t, \widehat{y}_t)$ est une loi normale de moyenne $(0,0)$ et de matrice de covariance$_0$

$$\Sigma = \begin{pmatrix} \operatorname{var}[y_t] & \operatorname{cov}[y_t \widehat{y}_t] \\ \operatorname{cov}[y_t \widehat{y}_t] & \operatorname{var}[\widehat{y}_t] \end{pmatrix}$$

Mais en utilisant l'orthogonalité entre $y_t - \widehat{y}_t$ et $\widehat{y}_t$, on a

$$E\left[(y_t - \widehat{y}_t) \widehat{y}_t \right] = 0 \Longrightarrow \operatorname{cov}[y_t \widehat{y}_t] = E[y_t \widehat{y}_t] = \operatorname{var}[\widehat{y}_t]$$

d'où

$$\Sigma = \begin{pmatrix} \operatorname{var}[y_t] & \operatorname{var}[\widehat{y}_t] \\ \operatorname{var}[\widehat{y}_t] & \operatorname{var}[\widehat{y}_t] \end{pmatrix}$$

Les variances de y_t et $\widehat{y}_t$ se calculent aisément à partir de leurs expressions en fonctions des innovations

$$\begin{aligned} \operatorname{var}[y_t] &= \sum_{n=0}^{\infty} \alpha_n^2 \\ \operatorname{var}[\widehat{y}_t] &= \sum_{n=1}^{\infty} \alpha_n^2 \end{aligned}$$