Correction du devoir en temps libre

L'objet des deux exercices du devoir est l'étude de processus dits autorégressifs. Ces suites chronologiques ont de très nombreuses applications en démographie, biologie, climatologie, économétrie, finance etc..

Exercice 1 *Processus auto-regressif - cas stable*
On considère la suite de variables aléatoires réelles (X_n) définie par $X_0 = 0$ et pour $n \geq 0$:

$$X_{n+1} = \theta X_n + \varepsilon_{n+1},$$

où θ est un paramètre inconnu tel que $0 < |\theta| < 1$ et où (ε_n) est une suite de variables aléatoires réelles indépendantes et de même loi $\mathcal{N}(0,1)$. On dit que la suite (X_n) est un processus autorégressif d'ordre 1.

1. Montrer par récurrence que $X_n = \sum_{k=1}^n \theta^{n-k}\varepsilon_k$.
   Solution : On a bien $X_1 = \varepsilon_1$. Si l'on suppose que pour un entier n donné $X_n = \sum_{k=1}^n \theta^{n-k}\varepsilon_k$, alors par définition $X_{n+1} = \theta X_n + \varepsilon_{n+1} = \sum_{k=1}^n \theta^{n+1-k}\varepsilon_k + \varepsilon_{n+1} = \sum_{k=1}^{n+1} \theta^{n+1-k}\varepsilon_k$, d'où le résultat.
2. En déduire que, lorsque n tend vers l'infini :

$$X_n \xrightarrow{loi} \mathcal{N}\left(0, \frac{1}{1-\theta^2}\right).$$

   Solution : D'après la question précédente, et par indépendance des variables (ε_i), les variables X_n sont gaussiennes, précisément $X_n \sim \mathcal{N}(0, \sum_{k=0}^{n-1} \theta^{2k})$. Autrement dit, la fonction caractéristique de X_n s'écrit :

$$\phi_{X_n}(t) := \mathbb{E}\left[e^{itX_n}\right] = \exp\left(-\frac{t^2}{2} \times \sum_{k=0}^{n-1} \theta^{2k}\right).$$

   Pour tout $t \in \mathbb{R}$, on a alors

$$\lim_{n\to+\infty} \phi_{X_n}(t) = \exp\left(-\frac{t^2}{2} \times \sum_{k=0}^{+\infty} \theta^{2k}\right) = \exp\left(-\frac{t^2}{2} \times \frac{1}{1-\theta^2}\right).$$

   On reconnaît la fonction caractéristique d'une variable gaussienne centrée de variance $\frac{1}{1-\theta^2}$, d'où le résultat.
   On cherche à présent à deviner le paramètre inconnu θ à partir des observations. Plus précisément, on cherche à exhiber un estimateur consistant de θ, c'est-à-dire une fonction mesurable $\widehat{\theta}_n = \widehat{\theta}(X_1, \ldots, X_n)$ des données jusqu'au temps n, dont l'expression ne dépend pas de θ, et qui converge presque sûrement vers θ lorsque n tend vers l'infini. On considère ainsi la suite

$$\widehat{\theta}_n := \frac{\sum_{k=1}^n X_k X_{k-1}}{\sum_{k=1}^n X_{k-1}^2}.$$

3. On pose $M_n := \sum_{k=1}^n X_{k-1}\varepsilon_k$. Vérifier que (M_n) est une martingale de carré intégrable, calculer son compensateur prévisible $\langle M\rangle_n$ et montrer que l'on a la relation :

$$\widehat{\theta}_n - \theta = \frac{M_n}{\langle M\rangle_n}.$$

Solution : Les variables X_k et ε_k sont gaussiennes, elles admettent donc des moments de tout ordre. Par suite, les produits $X_{k-1}\varepsilon_k$ sont de carré intégrable, ainsi que les sommes partielles M_n. Par définition, le compensateur $\langle M\rangle_n$ de M_n vérifie

$$\begin{aligned}
\langle M\rangle_n &= \textstyle\sum_{k=1}^n \mathbb{E}\left[(M_k - M_{k-1})^2 \,|\mathcal{F}_{k-1}\right] \\
&= \textstyle\sum_{k=1}^n \mathbb{E}\left[X_{k-1}^2\varepsilon_k^2|\mathcal{F}_{n-1}\right] \\
&= \textstyle\sum_{k=1}^n X_{k-1}^2\mathbb{E}\left[\varepsilon_k^2|\mathcal{F}_{n-1}\right] \\
&= \textstyle\sum_{k=1}^n X_{k-1}^2\mathbb{E}\left[\varepsilon_k^2\right] \\
&= \textstyle\sum_{k=1}^n X_{k-1}^2 = \sum_{k=0}^{n-1} X_k^2.
\end{aligned}$$

Par conséquent, en utilisant le fait que $X_k = \theta X_{k-1} + \varepsilon_k$, on a bien :

$$\widehat{\theta}_n = \frac{\sum_{k=1}^n X_k X_{k-1}}{\sum_{k=1}^n X_{k-1}^2} = \theta + \frac{\sum_{k=1}^n X_{k-1}\varepsilon_k}{\sum_{k=1}^n X_{k-1}^2} = \theta + \frac{M_n}{\langle M\rangle_n}.$$

4. Montrer en utilisant l'inégalité de Cauchy-Schwarz que

$$X_n^2 \leq \frac{1}{1-|\theta|}\sum_{k=1}^n |\theta|^{n-k}\varepsilon_k^2, \quad \text{et} \quad \sum_{k=1}^n X_k^2 \leq \left(\frac{1}{1-|\theta|}\right)^2 \sum_{k=1}^n \varepsilon_k^2.$$

Solution : On rappelle que $X_n = \sum_{k=1}^n \theta^{n-k}\varepsilon_k = \sum_{k=1}^n |\theta|^{\frac{n-k}{2}} \times |\theta|^{\frac{n-k}{2}}\varepsilon_k$. D'après l'inégalité de Cauchy-Schwarz, on a donc

$$X_n^2 \leq \left(\sum_{k=1}^n |\theta|^{n-k}\right) \times \left(\sum_{k=1}^n |\theta|^{n-k}\varepsilon_k^2\right) \leq \frac{1}{1-|\theta|}\sum_{k=1}^n |\theta|^{n-k}\varepsilon_k^2.$$

Il sommant, et en inversant les deux sommes, on en déduit

$$\sum_{k=1}^n X_k^2 \leq \left(\frac{1}{1-|\theta|}\right)^2 \sum_{k=1}^n \varepsilon_k^2.$$

5. En déduire que presque sûrement $\langle M\rangle_n = O(n)$ et $M_n = o(n)$ lorsque n tend vers l'infini.
Solution : Les variables ε_k^2 étant intégrables, indépendantes et identiquement distribuées, d'après la loi forte des grands nombres $n^{-1}\sum_{k=1}^n \varepsilon_k^2$ tend presque sûrement vers 1 lorsque n tend vers l'infini, et d'après la dernière inégalité, on a $\langle M\rangle_n = O(n)$. On sait que la martingale M_n converge si et seulement si le compensateur $\langle M\rangle_n$ converge. Si c'est le cas, on a naturellement $M_n = O(1) = o(n)$. Si le crochet $\langle M\rangle_n$ diverge, la loi des grands nombres pour les martingales de carré intégrable assure que $M_n/\langle M\rangle_n$ tend vers zéro, et l'on a également $M_n = o(\langle M\rangle_n) = o(O(n)) = o(n)$.

6. En remarquant que $(1-\theta^2)\langle M\rangle_{n+1} = -\theta^2 X_n^2 + 2\theta M_n + \sum_{k=1}^n \varepsilon_k^2$, montrer que

$$\lim_{n\to+\infty} \frac{\langle M\rangle_n}{n} = \frac{1}{1-\theta^2},$$

et en déduire que $\widehat{\theta}_n$ est bien un estimateur consistant de θ.

Solution : D'après la relation $(1-\theta^2)\langle M\rangle_{n+1} = -\theta^2 X_n^2 + 2\theta M_n + \sum_{k=1}^n \varepsilon_k^2$ et la question précédente, il suffit de montrer que X_n^2/n tend presque sûrement vers zéro lorsque n tend vers l'infini. Fixons $\eta > 0$, on a alors d'après l'inégalité de Chebychev

$$\begin{aligned}
\mathbb{P}\left(\frac{X_n^2}{n} > \eta\right) &= \mathbb{P}(X_n^2 > n\eta) = \mathbb{P}\left(\mathcal{N}\left(0, \frac{1-\theta^{2n}}{1-\theta^2}\right)^2 > n\eta\right) \\
&= \mathbb{P}\left(\mathcal{N}(0,1)^2 > n\eta\left(\frac{1-\theta^2}{1-\theta^{2n}}\right)\right) \\
&\leq \frac{\mathbb{E}\left[\mathcal{N}(0,1)^4\right]}{n^2\eta^2\left(\frac{1-\theta^2}{1-\theta^{2n}}\right)^2} \leq \frac{cste}{n^2}
\end{aligned}$$

Le terme de droite est le terme général d'une série convergente. D'après le lemme de Borel-Cantelli, avec probabilité un, l'évènement $\{X_n^2/n > \eta\}$ se produit un nombre fini de fois, i.e avec probabilité un, à partir d'un certain rang, on a $X_n^2/n \leq \eta$. Ceci étant vrai pour tout rationnel η, on conclut que presque sûrement, la suite X_n^2/n tend vers zéro lorsque n tend vers l'infini. On a donc montré que le compensateur $\langle M\rangle_n$ tend presque sûrement vers l'infini avec n ; d'après la loi des grands nombres pour les martingales de carré intégrable, on peut alors conclure que $M_n/\langle M\rangle_n$ tend vers zéro presque sûrement, en par conséquent $\widehat{\theta}_n$ est bien un estimateur consistant de θ.

Exercice 2 *Processus auto-regressif - cas explosif*

On considère maintenant un processus autorégressif à coefficients aléatoires, i.e. une suite (X_n) telle que $X_0 = 0$ et pour $n \geq 0$

$$X_{n+1} = \theta_{n+1} X_n + \varepsilon_{n+1},$$

où

- la suite (θ_n) est une suite de variables aléatoires indépendantes et identiquement distribuées avec $\mathbb{E}[\theta_n] = \theta$ inconnu et $\mathbb{E}[\theta_n^2] = \tau^2 > \theta^2$,
- la suite (ε_n) est une suite de variables indépendantes et identiquement distribuées avec $\mathbb{E}[\varepsilon_n] = 0$ et $\mathbb{E}[\varepsilon_n^2] = 1$,
- les suites (θ_n) et (ε_n) sont indépendantes.

On note $\mathcal{F}_n = \sigma(X_1, X_2, \ldots, X_n)$ la filtration naturelle associée au modèle et on se place dans le cas dit explosif où $\tau^2 > 1$.

1. Montrer que pour tout $n \geq 1$, on a $\mathbb{E}[\theta_{n+1}\varepsilon_{n+1}] = 0$,

$$\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \theta X_n, \quad \mathbb{E}[X_{n+1}^2|\mathcal{F}_n] = \tau^2 X_n^2 + 1.$$

Solution : Les variables θ_n est ε_n sont de carré intégrable, donc d'après l'inégalité de Cauchy-Schwarz $\theta_{n+1}\varepsilon_{n+1}$ est intégrable. Par indépendance, on a alors $\mathbb{E}[\theta_{n+1}\varepsilon_{n+1}] = \mathbb{E}[\theta_{n+1}]\mathbb{E}[\varepsilon_{n+1}]$ et comme ε_{n+1} est centrée $\mathbb{E}[\theta_{n+1}\varepsilon_{n+1}] = 0$. Ensuite, comme X_n est $\mathcal{F}_n-$mesurable, on a

$$\begin{aligned}\mathbb{E}[X_{n+1}|\mathcal{F}_n] = \mathbb{E}[\theta_{n+1}X_n + \varepsilon_{n+1}|\mathcal{F}_n] \quad &= X_n \times \mathbb{E}[\theta_{n+1}|\mathcal{F}_n] + \mathbb{E}[\varepsilon_{n+1}|\mathcal{F}_n] \\ &= X_n \times \theta + 0,\end{aligned}$$

$$\begin{aligned}\mathbb{E}[X_{n+1}^2|\mathcal{F}_n] \quad &= \mathbb{E}[\theta_{n+1}^2 X_n^2 + \varepsilon_{n+1}^2 + 2X_n\theta_{n+1}\varepsilon_{n+1}|\mathcal{F}_n] \\ &= X_n^2 \times \mathbb{E}[\theta_{n+1}^2|\mathcal{F}_n] + \mathbb{E}[\varepsilon_{n+1}^2|\mathcal{F}_n] + 2X_n\mathbb{E}[\theta_{n+1}\varepsilon_{n+1}|\mathcal{F}_n] \\ &= X_n^2 \times \mathbb{E}[\theta_{n+1}^2] + \mathbb{E}[\varepsilon_{n+1}^2] + 2X_n\mathbb{E}[\theta_{n+1}\varepsilon_{n+1}] \\ &= X_n^2 \times \tau^2 + 1 + 0.\end{aligned}$$

2. Pour $n \geq 1$, on pose $Y_n = X_n^2/\tau^{2n}$. Montrer que la suite (Y_n) est une sous-martingale bornée dans $\mathbb{L}^1$.
   Solution : La suite Y_n est clairement adaptée à la filtration $\mathcal{F}_n$. D'après la dernière formule, les variables X_n^2 sont intégrables (récurrence immédiate), et l'on a

$$\mathbb{E}[Y_{n+1}|\mathcal{F}_n] = \frac{1}{\tau^{2n+2}}\mathbb{E}[X_{n+1}^2|\mathcal{F}_n] = \frac{X_n^2}{\tau^{2n}} + \frac{1}{\tau^{2n}} \geq \frac{X_n^2}{\tau^{2n}} = Y_n,$$

   autrement dit, la suite (Y_n) est une sous-martingale. Par ailleurs le calcul précédent montre que $\mathbb{E}[Y_{n+1}] = \mathbb{E}[Y_n] + \tau^{-2n}$ donc $\mathbb{E}[|Y_n|] = \mathbb{E}[Y_n] = \sum_{k=0}^{n-1}\tau^{-2k} = \frac{1-\tau^{-2n}}{1-\tau^{-2}} \leq \frac{1}{1-\tau^{-2}}$ et la suite (Y_n) est bornée dans $\mathbb{L}^1$.
3. En déduire que (Y_n) converge presque sûrement vers une variable aléatoire Y, puis en utilisant le lemme de Cesàro que

$$\lim_{n\to+\infty}\frac{1}{\tau^{2n}}\sum_{k=1}^{n}X_k^2 = \left(\frac{\tau^2}{\tau^2-1}\right)Y \quad p.s..$$

   Solution : D'après le cours, on sait que la suite Y_n converge alors presque sûrement, lorsque n tend vers l'infini, vers une variable aléatoire Y. D'après le lemme de Cesáro, en posant $b_k = \tau^{2k}$ qui est une suite strictement croissante, il vient :

$$\lim_{n\to+\infty}\frac{1}{\tau^{2n}}\sum_{k=1}^{n}(\tau^{2k}-\tau^{2k-2})Y_k = Y,$$

   c'est-à-dire :

$$(1-\tau^{-2}) \times \lim_{n\to+\infty}\frac{1}{\tau^{2n}}\sum_{k=1}^{n}X_k^2 = Y, \text{ d'où le résultat}$$

   Dans toute la suite, nous supposerons que la variable limite Y est strictement positive presque sûrement. Comme dans l'exercice précédent, on souhaite donner un estimateur consistant du paramètre inconnu θ. On introduit les suites :

$$C_n := \sum_{k=1}^{n}\frac{X_kX_{k-1}}{1+X_{k-1}^2}, \quad D_n := \sum_{k=1}^{n}\frac{X_{k-1}^2}{1+X_{k-1}^2}, \quad M_n := C_n - \theta D_n.$$

   On pose alors $\widetilde{\theta}_n := C_n/D_n$ de sorte que $\widetilde{\theta}_n - \theta = M_n/D_n$.

4. Montrer que la suite (M_n) est une martingale et que lorsque n tend vers l'infini $\langle M\rangle_n = O(n)$.
Solution : La suite M_n est clairement adaptée à filtration $\mathcal{F}_n$. Par ailleurs, en majorant brutalement, on a $|D_n| \leq n$ et $|C_n| \leq \sum_{k=1}^n |X_k|$, donc M_n est intégrable (et de carré intégrable) pour tout $n \geq 0$. Par ailleurs, on a

$$\begin{aligned}
\mathbb{E}[M_{n+1}|\mathcal{F}_n] &= \mathbb{E}[C_{n+1}|\mathcal{F}_n] - \theta\mathbb{E}[D_{n+1}|\mathcal{F}_n] \\
&= C_n + \mathbb{E}\left[\frac{X_{n+1}X_n}{1+X_n^2} \mid \mathcal{F}_n\right] - \theta D_n - \theta\mathbb{E}\left[\frac{X_n^2}{1+X_n^2}|\mathcal{F}_n\right] \\
&= C_n - \theta D_n + \frac{X_n}{1+X_n^2}\mathbb{E}\left[X_{n+1}|\mathcal{F}_n\right] - \theta\frac{X_n^2}{1+X_n^2} \\
&= M_n \ \text{ car } \ \mathbb{E}[X_{n+1}|\mathcal{F}_n] = \theta X_n.
\end{aligned}$$

La suite (M_n) est donc bien une martingale de carré intégrable. Le calcul du compensateur donne :

$$\begin{aligned}
\langle M\rangle_n &= \textstyle\sum_{k=1}^n \mathbb{E}[(M_k - M_{k-1})^2|\mathcal{F}_{k-1}] \\
&= \sum_{k=1}^n \mathbb{E}\left[\left(\frac{X_{k-1}}{1+X_{k-1}^2}\right)^2 (X_k - \theta X_{k-1})^2|\mathcal{F}_{k-1}\right] \\
&= \sum_{k=1}^n \left(\frac{X_{k-1}}{1+X_{k-1}^2}\right)^2 \mathbb{E}\left[\left(X_k^2 - 2\theta X_k X_{k-1} + \theta^2 X_{k-1}^2\right)|\mathcal{F}_{k-1}\right] \\
&= \sum_{k=1}^n \left(\frac{X_{k-1}}{1+X_{k-1}^2}\right)^2 \left(\mathbb{E}\left[X_k^2|\mathcal{F}_{k-1}\right] - 2\theta X_{k-1}\mathbb{E}\left[X_k|\mathcal{F}_{k-1}\right] + \theta^2 X_{k-1}^2\right) \\
&= \sum_{k=1}^n \left(\frac{X_{k-1}}{1+X_{k-1}^2}\right)^2 \left(\tau^2 X_{k-1}^2 + 1 - 2\theta^2 X_{k-1}^2 + \theta^2 X_{k-1}^2\right) \\
&= \sum_{k=1}^n \left(\frac{X_{k-1}}{1+X_{k-1}^2}\right)^2 \left(1 + (\tau^2 - \theta^2)X_{k-1}^2\right) \leq n(1 + (\tau^2 - \theta^2)).
\end{aligned}$$

5. Montrer que lorsque n tend vers l'infini, $n^{-1}D_n$ tend presque sûrement vers 1 et en déduire que

$$\lim_{n\to+\infty} \widetilde{\theta}_n = \theta \ \ p.s..$$

Solution : La suite (Y_n) converge presque sûrement lorsque n tend vers l'infini vers $Y > 0$ (par hypothèse). La suite (X_n) tend donc presque sûrement vers l'infini avec n et par suite, D_n/n tend vers 1 presque sûrement (lemme de Cesáro). Dès lors en écrivant le rapport M_n/D_n comme $M_n/D_n = M_n/n \times n/D_n$, et en utilisant la loi des grands nombre pour les martingales de carré intégrable, on conclut que presque sûrement M_n/D_n tend vers zéro lorsque n tend vers l'infini, i.e. $\widetilde{\theta}_n$ est un estimateur consistant de θ