Ecole Nationale Supérieure de Techniques Avancées ParisTech
PRB202 - Martingales et Algorithmes Stochastiques
Corrigé de la PC6 - 25 janvier 2018

__Exercice 1__ **:** 1. On a, pour tout $n \geq 1$,

$$\begin{aligned} \frac{\gamma_n}{\gamma_{n-1}} &= \frac{\alpha^2\,\beta}{\sigma^2 + n\,(\alpha\beta)^2}\,\frac{\sigma^2 + (n-1)\,(\alpha\beta)^2}{\alpha^2\,\beta}\,, \\ &= 1 - \beta\,\frac{\alpha^2\,\beta}{\sigma^2 + n\,(\alpha\beta)^2}\,, \\ &= 1 - \beta\,\gamma_n\,. \end{aligned} \tag{1}$$

2. Pour tout $n \in \mathbb{N}\,, \tilde{X}_n = X_n - x_a$, de sorte que :

$$\begin{aligned} \tilde{X}_{n+1} - \tilde{X}_n &= X_{n+1} - X_n\,, \\ &= \gamma_n(a - Y_{n+1})\,, \quad n \in \mathbb{N}\,. \end{aligned} \tag{2}$$

Ainsi, quel que soit $n \in \mathbb{N}$,

$$\begin{aligned} ||\tilde{X}_{n+1}||^2 &= ||\tilde{X}_n + (\tilde{X}_{n+1} - \tilde{X}_n)||^2\,, \\ &= ||\tilde{X}_n||^2 + ||(\tilde{X}_{n+1} - \tilde{X}_n)||^2 + 2 < \tilde{X}_n, \tilde{X}_{n+1} - \tilde{X}_n >\,, \\ &= ||\tilde{X}_n||^2 + \gamma_n^2||a - Y_{n+1}||^2 + 2\,\gamma_n < \tilde{X}_n, a - Y_{n+1} >\,, \text{ d'après (2)} \\ &= ||\tilde{X}_n||^2 + \gamma_n^2||a - Y_{n+1}||^2 + 2\,\gamma_n < \tilde{X}_n, (a - U(X_n)) + (U(X_n) - Y_{n+1}) >\,, \\ &= ||\tilde{X}_n||^2 + \gamma_n^2||a - Y_{n+1}||^2 - 2\,\gamma_n < \tilde{X}_n, U(X_n) - a > + 2\,\gamma_n < \tilde{X}_n, U(X_n) - Y_{n+1} >\,. \end{aligned}$$

En prenant l'espérance dans l'égalité précédente, il vient, pour tout $n \in \mathbb{N}$:

$$\mathbb{E}(||\tilde{X}_{n+1}||^2) = \mathbb{E}(||\tilde{X}_n||^2) + \gamma_n^2\mathbb{E}(||a - Y_{n+1}||^2) - 2\,\gamma_n\mathbb{E}(< \tilde{X}_n, U(X_n) - a >) + 2\,\gamma_n\mathbb{E}(< \tilde{X}_n, U(X_n) - Y_{n+1} >)\,. \tag{3}$$

Soit $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ l'espace des (classes de) variables aléatoires Z définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans $\mathbb{R}^d$ telles que : $\mathbb{E}(||Z||^2) < +\infty$.
$\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ est un espace de Hilbert muni du produit scalaire $<< Z_1, Z_2 >>= \mathbb{E}(< Z_1, Z_2 >)$, pour tout $(Z_1, Z_2) \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)^2$.
Notons $\mathbb{L}^2(\Omega, \mathcal{F}, \sigma(X_0, \cdots, X_n), \mathbb{P}; \mathbb{R}^d)\,, n \in \mathbb{N}$ l'ensemble des (classes de) variables aléatoires U définies sur $(\Omega, \mathcal{F}, \mathbb{P})$ et à valeurs dans $\mathbb{R}^d\,, \sigma(X_0, \cdots, X_n)$ -mesurables telles que : $\mathbb{E}(||U||^2) < +\infty$.
Il apparaît que pour tout $n \in \mathbb{N}, \mathbb{L}^2(\Omega, \mathcal{F}, \sigma(X_0, \cdots, X_n), \mathbb{P}; \mathbb{R}^d)$ est un sous-espace vectoriel fermé de $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$.
Or, d'après l'énoncé, $\mathbb{E}[||X_n||^2] < +\infty$, pour tout $n \in \mathbb{N}$, de sorte que $\mathbb{E}[||\tilde{X}_n||^2] \leq 2\,(\mathbb{E}[||X_n||^2] + x_a^2) < +\infty$.
Par ailleurs, quel que soit $n \in \mathbb{N}\,, X_n$ donc $\tilde{X}_n = X_n - x_a$ est $\sigma(X_0, \cdots, X_n)$ -mesurable.
On en déduit que, pour tout $n \in \mathbb{N}$:

$$\tilde{X}_n = X_n - x_a \in \mathbb{L}^2(\Omega, \mathcal{F}, \sigma(X_0, \cdots, X_n), \mathbb{P}; \mathbb{R}^d)\,.$$

De plus, quel que soit $n \geq 1\,, \mathbb{E}[||Y_n||^2] < +\infty$ et $\mathbb{E}[Y_{n+1}|\sigma(X_0, \cdots, X_n)] = U(X_n)\,, n \in \mathbb{N}$.
Ainsi, pour tout $n \in \mathbb{N}\,, U(X_n)$ est la projection orthogonale de $Y_{n+1} \in \mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ sur le sous-espace $\mathbb{L}^2(\Omega, \mathcal{F}, \sigma(X_0, \cdots, X_n), \mathbb{P}; \mathbb{R}^d)$, de sorte que $Y_{n+1} - U(X_n)$ est orthogonal dans $\mathbb{L}^2(\Omega, \mathcal{F}, \mathbb{P}; \mathbb{R}^d)$ à tout vecteur de $\mathbb{L}^2(\Omega, \mathcal{F}, \sigma(X_0, \cdots, X_n), \mathbb{P}; \mathbb{R}^d)$ et en particulier à $\tilde{X}_n$.
Ainsi, quel que soit $n \in \mathbb{N}\,, << \tilde{X}_n, Y_{n+1} - U(X_n) >>= \mathbb{E}(< \tilde{X}_n, Y_{n+1} - U(X_n) >) = 0$, et on obtient alors de l'égalité (3) l'identité recherchée :

$$\mathbb{E}(||\tilde{X}_{n+1}||^2) = \mathbb{E}(||\tilde{X}_n||^2) + \gamma_n^2\mathbb{E}(||a - Y_{n+1}||^2) - 2\gamma_n\mathbb{E}(< U(X_n) - a, X_n - x_a >)\,. \tag{4}$$

3. Comme $U(x_a)=a$, il vient, pour tout $n\geq 1$ tel que $\tilde{X}_n\neq 0$:

$$\begin{aligned}
||\tilde{X}_n||^2+\gamma_n^2||a-U(X_n)||^2-2\gamma_n < U(X_n)-a, X_n-x_a >) &= \sum_{i=1}^{d}((X_n^i-x_a^i)^2+\gamma_n^2(U^i(X_n)-U^i(x_a))^2 \\
&\quad -2\gamma_n(U^i(X_n)-U^i(x_a))(X_n^i-x_a^i))\,, \\
&= \sum_{i=1}^{d}((X_n^i-x_a^i)^2\left(1+\gamma_n^2\left(\frac{U^i(X_n)-U^i(x_a)}{X_n^i-x_a^i}\right)^2\right. \\
&\quad \left. -2\gamma_n\frac{U^i(X_n)-U^i(x_a)}{X_n^i-x_a^i}\right)\,, \\
&= \sum_{i=1}^{d}((X_n^i-x_a^i)^2\left(1-\gamma_n\frac{U^i(X_n)-U^i(x_a)}{X_n^i-x_a^i}\right)^2 .
\end{aligned} \tag{5}$$

Or, d'après l'énoncé, on a, pour tout $x=(x_1,\cdots,x_d)\in\mathbb{R}^d$ vérifiant $x\neq x_a$:

$$0<\beta\leq\frac{U^i(x)-U^i(x_a)}{x^i-x_a^i}\leq\frac{1}{\gamma_1}\,,\quad 1\leq i\leq d\,. \tag{6}$$

Comme $\gamma_n>0$, quel que soit $n\geq 1$, on déduit de l'inégalité (6) que, si $\tilde{X}_n\neq 0$:

$$1-\frac{\gamma_n}{\gamma_1}\leq 1-\gamma_n\frac{U^i(X_n)-U^i(x_a)}{X_n^i-x_a^i}\leq 1-\gamma_n\,\beta\,,$$

soit encore, pour tout $n\geq 1$:

$$\left(1-\gamma_n\frac{U^i(X_n)-U^i(x_a)}{X_n^i-x_a^i}\right)^2\leq(1-\gamma_n\,\beta)^2\,, \tag{7}$$

puisque $(\gamma_n)_{n\in\mathbb{N}}$ étant une suite décroissante, $\forall n\geq 1, \gamma_n\geq\gamma_1$ et $1-\frac{\gamma_n}{\gamma_1}\geq 0$.

Combinant (5) et (7), on obtient, pour tout $n\geq 1$ tel que $\tilde{X}_n\neq 0$:

$$||\tilde{X}_n||^2+\gamma_n^2||a-U(X_n)||^2-2\gamma_n < U(X_n)-a, X_n-x_a >)\leq||\tilde{X}_n||^2(1-\gamma_n\beta)^2\,. \tag{8}$$

L'inégalité précédente est encore trivialement vérifiée lorsque $\tilde{X}_n=X_n-x_a=0$, puisque $a-U(X_n)=a-U(x_a)=0$.

4. $U\ :\mathbb{R}^d\to\mathbb{R}^d$ étant continue, la fonction $x\mapsto U(x)-a$ est continue sur $\mathbb{R}^d$ donc $(\mathcal{B}(\mathbb{R}^d),\mathcal{B}(\mathbb{R}^d))$-mesurable. Pour tout $n\in\mathbb{N}, X_n$ est $(\sigma(X_0,\cdots,X_n),\mathcal{B}(\mathbb{R}^d))$- mesurable de sorte que $U(X_n)-a$ est $(\sigma(X_0,\cdots,X_n),\mathcal{B}(\mathbb{R}^d))$- mesurable comme étant la composée de deux fonctions $(\sigma(X_0,\cdots,X_n),\mathcal{B}(\mathbb{R}^d))$- et $(\mathcal{B}(\mathbb{R}^d),\mathcal{B}(\mathbb{R}^d))$- mesurable.
Utilisant (6), il vient pour tout $n\in\mathbb{N}$ tel que $\tilde{X}_n\neq 0$:

$$\left(\frac{U^i(X_n)-U^i(x_a)}{X_n^i-x_a^i}\right)^2\leq\frac{1}{\gamma_1^2}\,,\quad 1\leq i\leq d\,,$$

soit :

$$||U(X_n)-U(x_a)||^2\leq\frac{1}{\gamma_1^2}||X_n-x_a||^2\,,$$

l'inégalité précédente étant encore vérifiée si $X_n=x_a$.
On en déduit alors que pour tout $n\in\mathbb{N}, U(X_n)-a=U(X_n)-U(x_a)\in\mathbb{L}^2(\Omega,\mathcal{F},\sigma(X_0,\cdots,X_n),\mathbb{P};\mathbb{R}^d)$ puisque $X_n-x_a\in\mathbb{L}^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^d)$.
Comme $U(X_n)-Y_{n+1}$ est orthogonal dans $\mathbb{L}^2(\Omega,\mathcal{F},\mathbb{P};\mathbb{R}^d)$ à $\mathbb{L}^2(\Omega,\mathcal{F},\sigma(X_0,\cdots,X_n),\mathbb{P};\mathbb{R}^d)$, il est en particulier orthogonal à $U(X_n)-a$, quel que soit $n\in\mathbb{N}$.
Ainsi, d'après le théorème de Pythagore, on a, pour tout $n\in\mathbb{N}$:

$$\mathbb{E}(||a-Y_{n+1}||^2)=\mathbb{E}(||(a-U(X_n))+(U(X_n)-Y_{n+1})||^2)=\mathbb{E}(||a-U(X_n)||^2)+\mathbb{E}(||U(X_n)-Y_{n+1}||^2)\,. \tag{9}$$

Injectant (9) dans (4), il vient, quel que soit $n\in\mathbb{N}$:

$$\mathbb{E}(||\tilde{X}_{n+1}||^2)=\mathbb{E}(||\tilde{X}_n||^2)+\gamma_n^2\mathbb{E}(||a-U(X_n)||^2)-2\gamma_n\mathbb{E}(< U(X_n)-a,\tilde{X}_n >)+\gamma_n^2\mathbb{E}(||U(X_n)-Y_{n+1}||^2)\,.$$

Les trois premiers termes se majorent directement en prenant l'espérance dans l'inégalité (8), valide pour tout $n \geq 1$; le dernier terme est inférieur à $\gamma_n^2 \sigma^2$, d'après l'énoncé.

On a donc, quel que soit $n \geq 1$,

$$\begin{aligned}\mathbb{E}(||\tilde{X}_{n+1}||^2) &\leq \mathbb{E}(||\tilde{X}_n||^2)(1-\gamma_n\beta)^2 + \gamma_n^2\sigma^2\,,\\ &= \left(\frac{\gamma_n}{\gamma_{n-1}}\right)^2 \mathbb{E}(||\tilde{X}_n||^2) + \gamma_n^2\sigma^2\,,\end{aligned} \tag{10}$$

d'après (1).

5. L'inégalité demandée est satisfaite pour $n = 1$ puisque $\frac{\sigma^2\gamma_0}{\beta} = \alpha^2$ qui majore bien $\mathbb{E}(||\tilde{X}_1||^2)$ d'après l'énoncé.

   On procède ensuite par récurrence sur $n \in \mathbb{N}^*$.

   Utilisant (10), on a :

$$\begin{aligned}\mathbb{E}(||\tilde{X}_{n+1}||^2) &\leq \left(\frac{\gamma_n}{\gamma_{n-1}}\right)^2 \mathbb{E}(||\tilde{X}_n||^2) + \gamma_n^2\sigma^2\,,\\ &\leq \left(\frac{\gamma_n}{\gamma_{n-1}}\right)^2 \frac{\sigma^2}{\beta}\gamma_{n-1} + \gamma_n^2\sigma^2\,,\ \text{d'après l'hypothèse de récurrence,}\\ &= \frac{\sigma^2}{\beta}\gamma_n\left(\frac{\gamma_n}{\gamma_{n-1}} + \gamma_n\beta\right)\,,\\ &= \frac{\sigma^2}{\beta}\gamma_n\,,\ \text{grâce à la relation (1).}\end{aligned}$$

Exercice 2 : L'algorithme proposé est exactement un algorithme de Robbins-Monro avec :

$$U(x) = \mathbb{P}(X_1 \leq x) = \mathbb{P}(Z_1 \leq x-\theta) = \int_{-\infty}^{x-\theta} f(t)\,dt\,,\ \text{pour tout } x \in \mathbb{R}\,. \tag{11}$$

Ainsi, pour tout $x \in \mathbb{R}, U(x) \in [0,1]$ et U est bien une fonction bornée.

Par ailleurs, on a :

$$\begin{aligned}\int_{-\infty}^{+\infty} f(t)\,dt = 1\,,\ &\text{car } f \text{ est une densité de probabilité sur } \mathbb{R}\,,\\ &= \int_{-\infty}^{0} f(t)\,dt + \int_{0}^{+\infty} f(t)\,dt\,,\\ &= \int_{-\infty}^{0} f(t)\,dt + \int_{-\infty}^{0} f(-u)\,du\,,\ \text{en effectuant le changement de variable } u=-t \text{ dans la seconde intégrale,}\\ &= 2\int_{-\infty}^{0} f(t)\,dt\,,\ \text{puisque } f \text{ est paire sur } \mathbb{R}\,,\end{aligned}$$

de sorte que : $\int_{-\infty}^{0} f(t)\,dt = \frac{1}{2}$ et d'après (11), il vient alors : $U(\theta) = \int_{-\infty}^{0} f(t)\,dt = \frac{1}{2}$.

Comme f est continue sur $\mathbb{R}$ par hypothèse, U est continue sur $\mathbb{R}$, au vu de la définition de la fonction U donnée en (11).

De plus, $\forall x \in \mathbb{R}, U'(x) = f(x-\theta) > 0$, puisque f est à valeurs strictement positives.

U est alors strictement croissante sur $\mathbb{R}$.

Ainsi, $U(x) > U(\theta)$ si $x > \theta$ et $U(x) < U(\theta)$ lorsque $x < \theta$, de sorte que, pour tout $x \in \mathbb{R}$ tel que $x \neq \theta$,

$$< U(x) - U(\theta), x - \theta >\, > 0\,.$$

Par ailleurs, U définit une bijection de $\mathbb{R}$ sur $]0,1[$ et θ est l'unique solution de l'équation $U(\theta) = \frac{1}{2}$.

Les conditions sur $U, a = \frac{1}{2}$ et $x_a = \theta$ sont donc bien vérifiées.

Pour tout $n \in \mathbb{N}$, posons $Y_{n+1} = \mathbf{1}_{\{X_{n+1} \leq \hat{\theta}_n\}}$.

On a bien alors, quel que soit $n \in \mathbb{N}$,

$$\hat{\theta}_{n+1} - \hat{\theta}_n = \gamma_n\left(\frac{1}{2} - Y_{n+1}\right)\,.$$

Avec $\mathcal{F}_n = \sigma(\hat{\theta}_1, \cdots, \hat{\theta}_n)\,, n \geq 1\,, \hat{\theta}_n$ est $\mathcal{F}_n$ -mesurable, pour tout $n \geq 1$.

Prenant $\mathcal{F}_0 = \{\emptyset, \Omega\}$, comme $\hat{\theta}_0 \in \mathbb{R}, \hat{\theta}_n$ est alors $\mathcal{F}_n$ -mesurable, pour tout $n \in \mathbb{N}$.

Par ailleurs, puisque $\hat{\theta}_{n+1} = \hat{\theta}_n - \gamma_n \left(\mathbf{1}_{\{Z_{n+1} \leq \hat{\theta}_n - \theta\}} - \frac{1}{2}\right)$, quel que soit $n \in \mathbb{N}$, il vient : $\hat{\theta}_{n+1} = g_n(\hat{\theta}_n, Z_{n+1})$, où $g_n(y,z) = y - \gamma_n \left(\mathbf{1}_{\{z \leq y - \theta\}} - \frac{1}{2}\right)$, $n \in \mathbb{N}, (y,z) \in \mathbb{R}^2$.

On montre alors que, pour tout $n \geq 1$: $\mathcal{F}_n = \sigma(\hat{\theta}_1, \cdots, \hat{\theta}_n) = \sigma(Z_1, \cdots, Z_n)$.

Comme $(Z_n)_{n\geq 1}$ est une suite de variables aléatoires indépendantes, $X_{n+1} = Z_{n+1} + \theta$ est indépendante de $\mathcal{F}_n$, quel que soit $n \in \mathbb{N}$.

Complément : Soit $(\Omega, \mathcal{F}, \mathbb{P})$ un espace de probabilité et $\mathcal{G}$ une sous-tribu de $\mathcal{F}$.

Etant donné une variable aléatoire Y, $\mathcal{G}$ - mesurable, à valeurs dans un espace mesurable $(E, \mathcal{E})$ et Z une variable aléatoire indépendante de $\mathcal{G}$ prenant ses valeurs dans $(H, \mathcal{H})$.

Pour toute fonction borélienne Φ, bornée ou à valeurs positives, définie sur $(E \times H, \mathcal{E} \otimes \mathcal{H})$, la fonction ϕ donnée par :

$$\forall y \in E, \phi(y) = \mathbb{E}[\Phi(y, Z)]$$

est borélienne sur $(E, \mathcal{E})$ et on a : et on a :

$$\mathbb{E}[\Phi(Y,Z)|\mathcal{G}] = \phi(Y)\,, \mathbb{P}\text{- p.s.} \quad .$$

Ce résultat a été démontré sur une situation particulière dans le cadre de la question **1.** de **l'Exercice 3** de la **PC1**.

Il exprime que, sous les hypothèses énoncées, on peut calculer $\mathbb{E}[\Phi(Y,Z)|\mathcal{G}]$ comme si Y était une constante.

□

Quel que soit $n \in \mathbb{N}\,, \mathbb{E}[Y_{n+1}|\mathcal{F}_n] = \mathbb{E}[\mathbf{1}_{\{X_{n+1} \leq \hat{\theta}_n\}}|\mathcal{F}_n]$ est de la forme $\mathbb{E}[\Phi(\hat{\theta}_n, X_{n+1})|\mathcal{F}_n]$, avec

$\Phi(y,z) = \mathbf{1}_{\{z \leq y\}}$, qui est borélienne à valeurs positives, X_{n+1} est indépendante de $\mathcal{F}_n$ et $\hat{\theta}_n$ est $\mathcal{F}_n$ -mesurable, pour tout $n \in \mathbb{N}$.

Ainsi, quel que soit $n \in \mathbb{N}$:

$$\mathbb{E}[Y_{n+1}|\mathcal{F}_n] = \phi(\hat{\theta}_n)\,,$$

où, pour tout $y \in \mathbb{R}\,, \phi(y) = \mathbb{E}[\Phi(y, X_{n+1})]$.

Comme, quel que soit $x \in \mathbb{R}$,

$$\begin{aligned} U(x) &= \mathbb{P}(X_1 \leq x)\,, \\ &= \mathbb{P}(Z_1 \leq x - \theta)\,, \\ &= \int_{-\infty}^{x-\theta} f(t)\,dt\,, \\ &= \int_{-\infty}^{x} f(u-\theta)\,du\,, \end{aligned}$$

X_{n+1}, comme X_1 (les Z_n , $n \geq 1$, donc les $X_n = \theta + Z_n$, $n \geq 1$, sont identiquement distribués), admet alors pour densité $h(u) = f(u - \theta)$, pour tout $u \in \mathbb{R}$.

Quel que soit $y \in \mathbb{R}$, il vient :

$$\begin{aligned} \phi(y) &= \mathbb{E}[\Phi(y, X_{n+1})]\,, \\ &= \int_{-\infty}^{+\infty} \Phi(y,u)\,h(u)\,du\,, \\ &= \int_{-\infty}^{+\infty} \mathbf{1}_{\{u \leq y\}}\,f(u-\theta)\,du\,, \\ &= \int_{-\infty}^{+\infty} \mathbf{1}_{\{t \leq y-\theta\}}\,f(t)\,dt\,, \\ &= \int_{-\infty}^{y-\theta} f(t)\,dt\,. \end{aligned}$$

On a alors, pour tout $n \in \mathbb{N}$,

$$\mathbb{E}[Y_{n+1}|\mathcal{F}_n] = \phi(\hat{\theta}_n) = \int_{-\infty}^{\hat{\theta}_n - \theta} f(t)\,dt = U(\hat{\theta}_n)\,,\ \text{d'après (11).}$$

Il reste à remarquer que $0 \leq Y_{n+1} = \mathbf{1}_{\{X_{n+1} \leq \hat{\theta}_n\}} \leq 1$, quel que soit $n \in \mathbb{N}$, de sorte que : $\forall n \geq 1, |Y_n| \leq b$, avec $b = 1$.

Ainsi, toutes les hypothèses du résultat de convergence de l'algorithme de Robbins-Monro sont satisfaites et $(\hat{\theta}_n)_{n\in\mathbb{N}}$ converge $\mathbb{P}$- presque-sûrement vers θ.

<u>Exercice 3</u> : 1. Pour tout $n \in \mathbb{N}$, on a :

$$X_{n+1} - X_n = \gamma_n(Y_{n+1} - a)\,, \text{ avec } a = x^*\theta, \tag{12}$$

et

$$Y_{n+1} = \theta\, X_n + \epsilon_{n+1}\,. \tag{13}$$

Injectant (13) dans (12), il vient, quel que soit $n \in \mathbb{N}$:

$$X_{n+1} - x^* = (1 + \theta\gamma_n)(X_n - x^*) + \gamma_n\epsilon_{n+1}\,. \tag{14}$$

Montrons par récurrence sur $n \in \mathbb{N}$ que :

$$X_{n+1} - x^* = \beta_n(X_0 - x^*) + \beta_n\, M_{n+1}\,, \tag{15}$$

avec

$$M_{n+1} = \sum_{k=0}^{n} \frac{\gamma_k}{\beta_k}\epsilon_{k+1}\,, n \in \mathbb{N}\,.$$

L'égalité (14) pour $n = 0$ donne : $X_1 - x^* = (1 + \theta\gamma_0)(X_0 - x^*) + \gamma_0\epsilon_1$.

Or, $\beta_0 = 1 + \theta\gamma_0\,, M_1 = \displaystyle\sum_{k=0}^{0} \frac{\gamma_k}{\beta_k}\epsilon_{k+1} = \frac{\gamma_0}{\beta_0}\epsilon_1$ de sorte que $\beta_0 M_1 = \gamma_0\epsilon_1$.

Ainsi, (15) est valide lorsque $n = 0$.

D'apès l'égalité (14), on obtient, pour $n \geq 1$:

$$\begin{aligned}
X_{n+1} - x^* &= (1 + \theta\gamma_n)(X_n - x^*) + \gamma_n\epsilon_{n+1}\,,\\
&= (1 + \theta\gamma_n)(\beta_{n-1}(X_0 - x^*) + \beta_{n-1}\, M_n) + \gamma_n\epsilon_{n+1}\,, \text{ d'après l'hypothèse de récurrence,}\\
&= (1 + \theta\gamma_n)\beta_{n-1}(X_0 - x^*) + (1 + \theta\gamma_n)\beta_{n-1}\, M_n + \beta_n\frac{\gamma_n}{\beta_n}\epsilon_{n+1}\,,\\
&= \beta_n(X_0 - x^*) + \beta_n(M_n + \frac{\gamma_n}{\beta_n}\epsilon_{n+1})\,,\\
&= \beta_n(X_0 - x^*) + \beta_n\, M_{n+1}\,.
\end{aligned}$$

2. Posons $\mathcal{F}_n = \sigma(\epsilon_1, \cdots, \epsilon_n)$, pour tout $n \geq 1$.
Pour tout $k \in \{0, \cdots, n\}$, $\frac{\gamma_k}{\beta_k}\epsilon_{k+1}$ est $(\mathcal{F}_{k+1}, \mathcal{B}(\mathbb{R}))$-mesurable donc $(\mathcal{F}_{n+1}, \mathcal{B}(\mathbb{R}))$-mesurable, puisque $\mathcal{F}_{k+1} \subset \mathcal{F}_{n+1}\,, 0 \leq k \leq n$. Ainsi, M_{n+1} est $\mathcal{F}_{n+1}$-mesurable, comme étant la somme de $n+1$ variables aléatoires $(\mathcal{F}_{n+1}, \mathcal{B}(\mathbb{R}))$-mesurables.
Puisque $\epsilon_{k+1} \hookrightarrow \mathcal{N}(0,1)\,, 0 \leq k \leq n\,, M_{n+1}\,, n \in \mathbb{N}$ est par ailleurs une variable aléatoire intégrable comme étant la combinaison linéaire de variables aléatoires intégrables.
De plus, pour tout $n \in \mathbb{N}$, il vient :

$$\begin{aligned}
\mathbb{E}[M_{n+1}|\mathcal{F}_n] &= M_n + \mathbb{E}\left[\frac{\gamma_n}{\beta_n}\epsilon_{n+1}\middle|\mathcal{F}_n\right]\,, \text{ car } M_n \text{ est } \mathcal{F}_n\text{-mesurable,}\\
&= M_n + \frac{\gamma_n}{\beta_n}\mathbb{E}[\epsilon_{n+1}]\,, \text{ vu que } \epsilon_{n+1} \text{ est indépendante de } \mathcal{F}_n\,,\\
&= M_n\,, \text{ puisque } \epsilon_{n+1} \hookrightarrow \mathcal{N}(0,1)\,.
\end{aligned}$$

Ainsi, $(M_n)_{n\geq 1}$ est une $(\mathcal{F}_n)_{n\geq 1}$-martingale.
<u>Rappel</u> : Si $Y_1, \cdots, Y_n\,, n \geq 1$ sont n variables aléatoires indépendantes telles que $Y_k \hookrightarrow \mathcal{N}(m_k, \sigma_k^2)$, pour $1 \leq k \leq n$, alors $Y_1 + \cdots + Y_n \hookrightarrow \mathcal{N}(m_1 + \cdots + m_n, \sigma_1^2 + \cdots + \sigma_n^2)$. □
Pour tout $n \in \mathbb{N}\,, M_{n+1}$ est alors une variable aléatoire gaussienne telle que :

$$\mathbb{E}[M_{n+1}] = \sum_{k=0}^{n} \frac{\gamma_k}{\beta_k}\mathbb{E}[\epsilon_{k+1}] = 0 \tag{16}$$

et :
$$\mathbb{E}[M_{n+1}^2] = \mathrm{Var}(M_{n+1}) = \sum_{k=0}^{n} \mathrm{Var}\left(\frac{\gamma_k}{\beta_k}\epsilon_{k+1}\right) = \sum_{k=0}^{n} \frac{\gamma_k^2}{\beta_k^2}\mathrm{Var}(\epsilon_{k+1}) = \sum_{k=0}^{n} \frac{\gamma_k^2}{\beta_k^2}\,. \tag{17}$$

Or, pour tout $n \in \mathbb{N}\,, \log(\beta_n) = \log((1+\theta\gamma_0)\cdots(1+\theta\gamma_n)) = \sum_{k=0}^{n} \log(1+\theta\gamma_k)\,.$

Comme $\theta > 0$ et $\gamma_m > 0$, quel que soit $m \in \mathbb{N}\,, \log(1+\theta\gamma_m) \underset{m\to+\infty}{\sim} \theta\gamma_m > 0\,.$ Or, d'après l'énoncé, la série $\sum_{m\geq 0} \gamma_m$ est convergente de sorte que la série $\sum_{m\geq 0} \log(1+\theta\gamma_m)$ est également convergente; ainsi, $\log(\beta_n) = \sum_{k=0}^{n} \log(1+\theta\gamma_k)$ admet une limite finie lorsque $n \to +\infty$ et on en déduit que β_n converge vers un certain $\beta_\infty \in \mathbb{R}_*^+$.

Il vient alors : $\frac{\gamma_m^2}{\beta_m^2} \underset{m\to+\infty}{\sim} \frac{\gamma_m^2}{\beta_\infty^2}$ et la série $\sum_{m\geq 0} \gamma_m^2$ étant convergente (on suppose que toutes les hypothèses de l'algorithme de Robbins-Monro sont vérifiées sauf la divergence de la série $\sum_{m\geq 0} \gamma_m$), la série $\sum_{m\geq 0} \frac{\gamma_m^2}{\beta_m^2}$ est également convergente.

Ainsi, puisque pour tout $n \in \mathbb{N}\,, \mathbb{E}[M_{n+1}^2] = \sum_{k=0}^{n} \frac{\gamma_k^2}{\beta_k^2}$ d'après (17), on a : $\sup_{n\in\mathbb{N}} \mathbb{E}[M_{n+1}^2] \leq \sum_{n=0}^{+\infty} \frac{\gamma_n^2}{\beta_n^2} < +\infty\,.$

On déduit du théorème de convergence des martingales uniformément bornées dans $\mathbb{L}^2$, que M_n converge alors $\mathbb{P}$ - presque-sûrement et dans $\mathbb{L}^2$ vers une variable aléatoire M_∞.

Rappel : Soit $(Y_n)_{n\in\mathbb{N}}$ une suite de variables aléatoires gaussiennes telles que $Y_n \hookrightarrow \mathcal{N}(m_n, \sigma_n^2)\,, n \in \mathbb{N}$, convergeant en loi vers Y.

Alors :

La variable aléatoire Y est gaussienne et $Y \hookrightarrow \mathcal{N}(m, \sigma^2)$, avec $m = \lim_{n\to+\infty} m_n$ et $\sigma = \lim_{n\to+\infty} \sigma_n$.

□

La convergence au sens $\mathbb{P}$ - presque-sûr et dans $\mathbb{L}^2$ entraînant la convergence en loi, on utilise le rappel précédent pour conclure que la variable aléatoire M_∞ est gaussienne et telle que $M_\infty \hookrightarrow \mathcal{N}\left(0, \sum_{n=0}^{+\infty} \frac{\gamma_n^2}{\beta_n^2}\right)$, utilisant (16) et (17).

3. En passant à la limite lorsque $n \to +\infty$ dans l'égalité (15), on obtient que $(X_n - x^*)_{n\in\mathbb{N}}$ converge $\mathbb{P}$ - presque-sûrement vers une variable aléatoire $L = \beta_\infty(X_0 - x^* + M_\infty)$, non nulle $\mathbb{P}$ - presque-sûrement puisque β_∞ et M_∞ sont non identiquement nulles.

   Ce résultat vient contredire celui de l'algorithme de Robbins-Monro appliqué à la fonction $U(x) = \theta\,x$, pour $x \in \mathbb{R}$.

   L'unique solution de l'équation $U(x) = a$ est trivialement $x_a = x^* = \frac{a}{\theta}$.

   Les suites $(X_n)_{n\in\mathbb{N}}$ et $(Y_n)_{n\geq 1}$ sont pourtant exactement construites suivant les hypothèses de l'algorithme de Robbins-Monro.

   En particulier, on a bien : $\mathbb{E}[Y_{n+1}|\mathcal{F}_n] = U(X_n) = \theta\,X_n$, quel que soit $n \in \mathbb{N}$.

   En effet, comme pour tout $n \in \mathbb{N}\,, X_{n+1} = (1+\theta\gamma_n)X_n + \gamma_n(\epsilon_{n+1} - a) = f_n(X_n, \epsilon_{n+1})$, avec $f_n(x,\epsilon) = (1+\theta\gamma_n)x + \gamma_n(\epsilon - a)\,, n \in \mathbb{N}\,, (x,\epsilon) \in \mathbb{R}^2$, on peut montrer : $\mathcal{F}_n = \sigma(\epsilon_1, \cdots, \epsilon_n) = \sigma(X_1, \cdots, X_n)$, pour tout $n \geq 1$ et X_n est alors $\mathcal{F}_n$ - mesurable.

   Il vient donc :
$$\begin{aligned}\mathbb{E}[Y_{n+1}|\mathcal{F}_n] &= \mathbb{E}[\theta\,X_n + \epsilon_{n+1}|\mathcal{F}_n]\,,\\ &= \theta\,X_n + \mathbb{E}[\epsilon_{n+1}|\mathcal{F}_n]\,, \text{ car } X_n \text{ est } \mathcal{F}_n\text{ - mesurable},\\ &= \theta\,X_n + \mathbb{E}[\epsilon_{n+1}]\,, \text{ puisque } \epsilon_{n+1} \text{ est indépendante de } \mathcal{F}_n\,,\\ &= \theta\,X_n\,, \text{ vu que } \epsilon_{n+1} \hookrightarrow \mathcal{N}(0,1)\,.\end{aligned}$$

   La seule hypothèse non satisfaite est celle de la divergence de la série de terme général γ_n, qui apparaît, au vu de cet exercice, comme étant **absolument cruciale**.