Processus aléatoires et applications

Nils Berglund
Institut Denis Poisson – UMR 7013
Universite d'Orleans, Universite de Tours, CNRS

Notes de cours

Version provisoire

— Version : November 6, 2023 —

Contents

Chapitre 1

Chaînes de Markov à espace dénombrable

1.1 Notations – formalisme des générateurs

On considère une chaîne de Markov $(X_n)_{n\geqslant 0}$ sur un ensemble dénombrable $\mathscr{X}$.

Définition 1.1.1: Mesures signées

Une *mesure signée finie* sur $\mathscr{X}$ est une application $\mu : \mathscr{X} \to \mathbb{R}$ telle que

$$\|\mu\|_1 := \sum_{x\in\mathscr{X}} |\mu(x)| < \infty .$$

On notera $\mathscr{E}_1$ l'espace de Banach des mesures signées finies.
Si $\mu : \mathscr{X} \to [0,1]$, et $\|\mu\|_1 = 1$, alors μ est une *mesure de probabilité*.

Notons que la somme de deux mesures de probabilité n'est pas une mesure de probabilité. Le sous-ensemble des mesures de probabilité n'est donc pas un sous-espace de $\mathscr{E}_1$. Cependant, la combinaison convexe de deux mesures de probabilité est une mesure de probabilité.

Définition 1.1.2: Fonctions test

Une *fonction test* (ou *observable*) sur $\mathscr{X}$ est une application $f : \mathscr{X} \to \mathbb{R}$ telle que

$$\|f\|_\infty := \sup_{x\in\mathscr{X}} |f(x)| < \infty .$$

On notera $\mathscr{E}_\infty$ l'espace de Banach des fonctions test.

Voici quelques notations utiles.

- Pour une mesure signée finie μ et une fonction test f, nous écrirons

$$\mu(f) = \sum_{x\in\mathscr{X}} \mu(x) f(x) .$$

Cette quantité est bien définie, car

$$|\mu(f)| \leqslant \sum_{x\in\mathscr{X}} |\mu(x)||f(x)| \leqslant \sup_{x\in\mathscr{X}} |f(x)| \sum_{x\in\mathscr{X}} |\mu(x)| = \|f\|_\infty \|\mu\|_1 < \infty .$$

- Si μ est une mesure de probabilité, nous écrirons aussi $\mu(f) = \mathbb{E}^{\mu}[f]$.
- Si δ_x dénote la mesure de Dirac en x (c'est-à-dire que $\delta_x(x) = 1$ et $\delta_x(y) = 0$ si $y \neq x$), on abrège $\mathbb{E}^{\delta_x}[f]$ par $\mathbb{E}^x[f]$.
- Pour $A \subset \mathscr{X}$, on écrit
$$\mu(A) = \mu(\mathbb{1}_A) = \sum_{x\in A} \mu(x) .$$
- Si μ est une mesure de probabilité, alors $\mu(A)$ est aussi la probabilité de A.
- Pour une mesure de probabilité μ et une fonction test f, on écrira
$$\mathbb{E}^{\mu}[f(X_n)] = \mu P^n f = \sum_{x\in\mathscr{X}} \sum_{y\in\mathscr{X}} \mu(x)(P^n)_{xy} f(y) ,$$
où $(P^n)_{xy}$ est l'élément de matrice (x, y) de P^n.

Définition 1.1.3: Distance en variation totale

La *distance en variation totale* entre deux mesures $\mu, \nu \in \mathscr{E}_1$ est
$$\|\mu - \nu\|_{\mathrm{TV}} = 2 \sup\{|\mu(A) - \nu(A)| \colon A \subset X\} .$$

Intuitivement, deux mesures sont d'autant plus proches en variation totale qu'elles donnent des probabilités proches aux événements.

Pour des mesures de probabilité, le résultat suivant montre que la distance en variation totale est en fait équivalente à la norme ℓ^1.

Lemme 1.1.4: Équivalence des distances

Si μ et ν sont deux mesures de probabilité, alors
$$\|\mu - \nu\|_{\mathrm{TV}} = \sum_{x\in\mathscr{X}} |\mu(x) - \nu(x)| = \|\mu - \nu\|_1 .$$

Démonstration. Soit $B = \{x \in \mathscr{X} \colon \mu(x) > \nu(x)\}$. Alors on a
$$0 \leqslant \mu(B) - \nu(B) = 1 - \mu(B^c) + (1 - \nu(B^c)) = \nu(B^c) - \mu(B^c) , \tag{1.1.1}$$
ce qui implique
$$\begin{aligned}\sum_{x\in\mathscr{X}} |\mu(x) - \nu(x)| &= \sum_{x\in B} (\mu(x) - \nu(x)) + \sum_{x\in B^c} (\nu(x) - \mu(x)) \\ &= \mu(B) - \nu(B) + \nu(B^c) - \mu(B^c) \\ &= 2\big[\mu(B) - \nu(B)\big] \end{aligned} \tag{1.1.2}$$
par (1.1.1). De plus, pour tout $A \subset \mathscr{X}$,
$$\mu(A) - \nu(A) \leqslant \sum_{x\in A\cap B} (\mu(x) - \nu(x)) \leqslant \sum_{x\in B} (\mu(x) - \nu(x)) = \mu(B) - \nu(B) ,$$
où nous avons utilisé à deux reprises le fait que $\mu(x) \leqslant \nu(x)$ sur $A \cap B^c$. De même,
$$\nu(A) - \mu(A) \leqslant \sum_{x\in A\cap B^c} (\nu(x) - \mu(x)) \leqslant \nu(B^c) - \mu(B^c) = \mu(B) - \nu(B) .$$
Il suit de (1.1.2) que
$$|\mu(A) - \nu(A)| \leqslant \mu(B) - \nu(B) = \frac{1}{2}\|\mu - \nu\|_1 .$$
De plus, si $A = B$, on a égalité. □

Définition 1.1.5: Générateur

Soit P la matrice de transition d'une chaîne de Markov sur un ensemble dénombrable $\mathscr{X}$. Le *générateur* de la chaîne de Markov est l'application $\mathscr{L} : \mathscr{E}_\infty \to \mathscr{E}_\infty$ donnée par

$$(\mathscr{L}f)(x) = \sum_{y\in\mathscr{X}} p_{xy}[f(y) - f(x)] . \tag{1.1.3}$$

Remarquons que comme $\sum_{y\in\mathscr{X}} p_{xy} = 1$, on a l'expression équivalente

$$(\mathscr{L}f)(x) = \biggl[\sum_{y\in\mathscr{X}} p_{xy} f(y)\biggr] - f(x) = \mathbb{E}^x[f(X_1)] - f(x) .$$

On peut donc écrire $\mathscr{L} = P - I$, où I dénote la matrice identité.

1.2 Fonctions de Lyapounov

Dans la suite, nous supposons que P est la matrice de transition d'une chaîne de Markov *irréductible* sur $\mathscr{X}$. De plus, nous supposons que $\mathscr{X}$ est équipé d'une norme $\|\cdot\|$. Par exemple, si $\mathscr{X} \subset \mathbb{Z}$, on peut prendre $\|x\| = |x|$. Si $\mathscr{X} \subset \mathbb{Z}^d$, on peut prendre la norme Euclidienne (ou toute autre norme équivalente).

Définition 1.2.1: Fonction de Lyapounov

Une *fonction de Lyapounov* est une fonction $V : \mathscr{X} \to \mathbb{R}_+ = [0,\infty[$ satisfaisant

$$V(x) \to +\infty \qquad \text{pour } \|x\| \to \infty .$$

Proposition 1.2.2: Formule de Dynkin

Pour toute fonction de Lyapounov V, on a

$$\mathbb{E}^x[V(X_n)] = V(x) + \mathbb{E}^x\biggl[\sum_{m=0}^{n-1} (\mathscr{L}V)(x_m)\biggr] . \tag{1.2.1}$$

De plus, si τ est un temps d'arrêt tel que $\mathbb{E}^x[\tau] < \infty$, alors

$$\mathbb{E}^x[V(X_\tau)] = V(x) + \mathbb{E}^x\biggl[\sum_{m=0}^{\tau-1} (\mathscr{L}V)(x_m)\biggr] .$$

Démonstration. Montrons (1.2.1). On procède par récurrence sur n. L'initialisation se fait pour $n = 1$, où la définition (1.1.3) du générateur implique

$$\mathbb{E}^x[V(X_1)] = V(x) + (\mathscr{L}V)(x) .$$

Pour vérifier l'hérédité, on écrit

$$\mathbb{E}^x[V(X_{n+1})] = \mathbb{E}^x[V(X_n)] + \mathbb{E}^x[V(X_{n+1}) - V(X_n)] .$$

Or, si $\mathscr{F}_n$ dénote la tribu engendrée par $(X_0, X_1, \dots, X_n)$, on a

$$\begin{aligned}\mathbb{E}^x[V(X_{n+1}) - V(X_n)] &= \mathbb{E}^x\bigl[\mathbb{E}^x[V(X_{n+1}) - V(X_n) \bigm| \mathscr{F}_n]\bigr] \\ &= \mathbb{E}^x\bigl[\mathbb{E}^{X_n}[V(X_{n+1}) - V(X_n)]\bigr] = \mathbb{E}^x[(\mathscr{L}V)(X_n)] .\end{aligned}$$

Avec l'hypothèse de récurrence, ceci conclut la démonstration. □

Théorème 1.2.3: Croissance sous-exponentielle

Supposons qu'il existe une fonction de Lyapounov V et $c > 0$ tel que

$$(\mathscr{L}V)(x) \leqslant cV(x) \qquad \forall x \in \mathscr{X} .$$

Alors on a

$$\mathbb{E}^x[V(X_n)] \leqslant (1+c)^n V(x)$$

pour tout $n \in \mathbb{N}$ et tout $x \in \mathscr{X}$.

Démonstration. Notons $f_n(x) = \mathbb{E}^x[V(X_n)]$. Alors la formule de Dynkin implique

$$\begin{aligned} f_n(x) &= V(x) + \mathbb{E}^x\biggl[\sum_{m=0}^{n-1} (\mathscr{L}V)(x_m)\biggr] \\ &\leqslant V(x) + c\sum_{m=0}^{n-1} f_m(x) . \end{aligned}$$

En utilisant $f_0(x) = V(x)$ comme initialisation, le résultat suit facilement par récurrence sur n. □

Théorème 1.2.4: Non-explosion

Supposons qu'il existe $d \geqslant 0$ et un ensemble borné $K \subset \mathscr{X}$ tel que pour tout $x \in \mathscr{X}$, on ait

$$(\mathscr{L}V)(x) \leqslant d\mathbb{1}_K(x) = \begin{cases} d & \text{si } x \in K , \\ 0 & \text{sinon} . \end{cases}$$

Alors

$$\mathbb{P}^x\Bigl\{\lim_{n\to\infty} \|X_n\| = \infty\Bigr\} = 0 \qquad \forall x \in \mathscr{X} .$$

Démonstration (idée). Soit Ω_1 l'événement

$$\Omega_1 = \Bigl\{\omega\colon \lim_{n\to\infty} \|X_n(\omega)\| = \infty\Bigr\} .$$

Considérons le cas $d = 0$. Alors la formule de Dynkin implique

$$\mathbb{E}^x[V(X_n)] \leqslant V(x) .$$

Or, on a aussi

$$\mathbb{E}^x[V(X_n)] = \underbrace{\mathbb{E}^x[V(X_n)\mathbb{1}_{\Omega_1^c}]}_{\geqslant 0} + \mathbb{E}^x[V(X_n)\mathbb{1}_{\Omega_1}]$$

Par conséquent, $\mathbb{E}^x[V(X_n)\mathbb{1}_{\Omega_1}] \leqslant \mathbb{E}^x[V(X_n)] \leqslant V(x)$. Comme $V(X_n)$ tend vers l'infini sur Ω_1, ceci n'est possible que si $\mathbb{E}^x[\mathbb{1}_{\Omega_1}] = \mathbb{P}^x\{\Omega_1\} = 0$. □

Théorème 1.2.5: Récurrence positive

Soit $f : \mathscr{X} \to [1,\infty[$ et V une fonction de Lyapounov telle que

$$(\mathscr{L}V)(x) \leqslant -cf(x) + d\mathbb{1}_K(x) \qquad \forall x \in K\,,$$

pour un ensemble borné $K \subset \mathscr{X}$ et des constantes $c > 0$ et $d \geqslant 0$. Supposons de plus qu'il existe $\delta > 0$ tel que K satisfait

$$p_{xy} \geqslant \delta \qquad \forall x, y \in K\,. \tag{1.2.2}$$

Alors la chaîne de Markov est récurrente positive, et admet donc une mesure de probabilité invariante π. De plus,

$$\pi(f) < \infty\,.$$

Démonstration. Commençons par considérer le cas où $K = \{x_0\}$ contient un seul point. Nous voulons montrer que x_0 est récurrent positif, car dans ce cas, la chaîne de Markov étant irréductible, elle sera récurrente positive. Soit donc

$$\tau_K = \tau_{x_0} = \inf\{n \geqslant 1 \colon X_n \in K\}\,.$$

Par la formule de Dynkin, nous avons

$$\begin{aligned}
0 \leqslant \mathbb{E}^{x_0}[V(X_{\tau_K})] &= V(x_0) + \mathbb{E}^{x_0}\biggl[\sum_{m=0}^{\tau_K-1} (\mathscr{L}V)(X_m)\biggr] \\
&\leqslant V(x_0) - c\mathbb{E}^{x_0}\biggl[\sum_{m=0}^{\tau_K-1} f(X_m)\biggr] + d \qquad\qquad (1.2.3)\\
&\leqslant V(x_0) - c\mathbb{E}^{x_0}[\tau_K] + d\,.
\end{aligned}$$

Il suit que

$$\mathbb{E}^{x_0}[\tau_{x_0}] = \mathbb{E}^{x_0}[\tau_K] \leqslant \frac{V(x_0)+d}{c}\,,$$

ce qui est bien fini, impliquant que la chaîne de Markov est récurrente positive.

Dans le cas où K contient au moins deux points, nous savons déjà par le raisonnement ci-dessus que $\mathbb{E}^{x_0}[\tau_K]$ est fini pour tout $x_0 \in K$. Il nous faut montrer que $\mathbb{E}^{x_0}[\tau_{x_0}]$ est également fini pour tout $x_0 \in K$.

Soit $\tau_{K,n}$ le temps du $n$ième passage de la chaîne en K, et soit

$$Y_n = X_{\tau_{K,n}}\,.$$

(La chaîne de Markov $(Y_n)_{n\geqslant 0}$ est appelée la *trace* de $(X_n)_{n\geqslant 0}$ sur K.) On peut déduire de l'hypothèse (1.2.2) que Y_n atteint x_0 en un temps d'espérance finie. Ceci suit du fait qu'une chaîne récurrente sur un ensemble fini est récurrente positive. Comme le temps de retour vers K est borné, il suit qu'on a bien $\mathbb{E}^{x_0}[\tau_{x_0}] < \infty$.

La chaîne de Markov étant récurrente positive, elle admet une unique probabilité invariante π. Il reste à montrer que $\pi(f)$ est fini. En fait, la mesure μ donnée par

$$\mu(y) = \mathbb{E}^{x_0}\biggl[\sum_{n=1}^{\tau_{x_0}} \mathbb{1}_{X_n=y}\biggr]$$

est invariante pour tout $x_0 \in \mathscr{X}$. La probabilité invariante π est obtenue en normalisant μ. Comme

$$\sum_{y\in\mathscr{X}} \mu(y) = \mathbb{E}^{x_0}\biggl[\sum_{n=1}^{\tau_{x_0}} \underbrace{\sum_{y\in\mathscr{X}} 1_{X_n=y}}_{=1}\biggr] = \mathbb{E}^{x_0}[\tau_{x_0}]\,,$$

on conclut que

$$\pi(y) = \frac{1}{\mathbb{E}^{x_0}[\tau_{x_0}]}\mathbb{E}^{x_0}\biggl[\sum_{n=1}^{\tau_{x_0}} 1_{X_n=y}\biggr]\,.$$

Il suit que

$$\pi(f) = \frac{1}{\mathbb{E}^{x_0}[\tau_{x_0}]}\mathbb{E}^{x_0}\biggl[\sum_{n=1}^{\tau_{x_0}} \underbrace{\sum_{y\in\mathscr{X}} f(y) 1_{X_n=y}}_{=f(X_n)}\biggr] = \frac{1}{\mathbb{E}^{x_0}[\tau_{x_0}]}\mathbb{E}^{x_0}\biggl[\sum_{n=1}^{\tau_{x_0}} f(X_n)\biggr]\,.$$

Or (1.2.3) implique

$$\mathbb{E}^{x_0}\biggl[\sum_{n=1}^{\tau_{x_0}} f(X_n)\biggr] = \mathbb{E}^{x_0}\biggl[\sum_{n=0}^{\tau_{x_0}-1} f(X_n)\biggr] \leqslant \frac{V(x_0)+d}{c}\,,$$

ce qui permet de majorer $\pi(f)$. □

Remarque 1.2.6

On peut affaiblir l'hypothèse (1.2.2) sur K de plusieurs manières.

- Il suffit de supposer qu'il existe un $k \geqslant 1$ tel que $(P^k)_{xy} \geqslant \delta$ pour tout $x, y \in K$ et un $\delta > 0$.
- Une condition suffisante encore plus faible est qu'il existe des réels $a_1, a_2, \dots \geqslant 0$, de somme égale à 1, tels que
$$\sum_{i=1}^{\infty} a_i (P^k)_{xy} \geqslant \delta$$
pour tout $x, y \in K$ et un $\delta > 0$. Des ensembles K satisfaisant ce critère sont appelés *petits* («petite set» en anglais).

1.3 Normes à poids

Pour obtenir des résultats de convergence de la loi de X_n vers π, il est plus utile de travailler avec des normes à poids.

Définition 1.3.1: Norme à poids sur les fonctions test

Soit $W : \mathscr{X} \to [1,\infty[$. La *norme à poids* W d'une fonction test est définie comme

$$\|f\|_W = \sup_{x\in\mathscr{X}} \frac{|f(x)|}{W(x)}\,.$$

On notera $\mathscr{E}_\infty^W$ l'espace de Banach des fonctions test f telles que $\|f\|_W < \infty$.

Notons les propriétés suivantes :

- Pour tout $x \in \mathscr{X}$, on a $|f(x)| \leqslant \|f\|_W W(x)$.
- On a $\|f\|_\infty \leqslant \|f\|_W$. Par conséquent $\mathscr{E}_\infty^W \subset \mathscr{E}_\infty$.
- Plus généralement, si W_1 et W_2 sont deux poids tels que $W_1(x) \leqslant W_2(x)$ pour tout $x \in \mathscr{X}$, alors $\|f\|_{W_1} \leqslant \|f\|_{W_2}$. Par conséquent $\mathscr{E}_\infty^{W_2} \subset \mathscr{E}_\infty^{W_1}$.

On peut également définir une métrique duale à $\|\cdot\|_W$ entre mesures signées finies de la manière suivante.

Définition 1.3.2: Distance à poids entre mesures

Pour une fonction poids $W : \mathscr{X} \to [1,\infty[$ et deux mesures signées finies μ, ν, on pose

$$\begin{aligned}\rho_W(\mu,\nu) &= \sup_{f\,:\,\|f\|_W \leqslant 1} \sum_{x\in\mathscr{X}} f(x)|\mu(x)-\nu(x)| \\ &= \sup_{f\,:\,\|f\|_W \neq 0} \frac{1}{\|f\|_W} \sum_{x\in\mathscr{X}} f(x)|\mu(x)-\nu(x)|\,.\end{aligned} \tag{1.3.1}$$

On a alors les propriétés suivantes :

- On peut remplacer $f(x)$ par $|f(x)|$ dans la définition (1.3.1). En effet, cette transformation ne change pas $\|f\|_W$.
- On peut également remplacer $|\mu(x)-\nu(x)|$ par $\mu(x)-\nu(x)$. En effet, il suffit de changer le signe de $f(x)$ selon le signe de $\mu(x)-\nu(x)$ pour trouver le même résultat.
- Si $W(x) = 1$ pour tout $x \in \mathscr{X}$, alors on a

$$\rho_W(\mu,\nu) = \|\mu-\nu\|_{\mathrm{TV}}\,.$$

En effet, le supremum dans (1.3.1) est alors atteint pour la fonction test f valant 1 partout.
- Pour un poids W général, on a

$$\rho_W(\mu,\nu) = \sum_{x\in\mathscr{X}} W(x)|\mu(x)-\nu(x)|\,. \tag{1.3.2}$$

En effet, le supremum dans (1.3.1) est atteint pour $f(x) = W(x)$ pour tout $x \in \mathscr{X}$.
- On a la majoration

$$\big|(\mu-\nu)(f)\big| \leqslant \|f\|_W \rho_W(\mu,\nu)\,. \tag{1.3.3}$$

En effet,

$$\big|(\mu-\nu)(f)\big| \leqslant \sum_{x\in\mathscr{X}} |\mu(x)-\nu(x)||f(x)| \leqslant \|f\|_W \sum_{x\in\mathscr{X}} |\mu(x)-\nu(x)|W(x)\,,$$

d'où le résultat, par (1.3.2).

1.4 Un critère de convergence

Les théorèmes de la section 1.2 sont dûs à Meyn et Tweedie [2]. Leurs travaux fournissent également un résultat de convergence de $\mathbb{E}^x[f(X_n)]$ vers $\pi(f)$, mais les hypothèses sont assez difficiles à vérifier (notamment, tous les ensembles bornés doivent être petits), et les bornes obtenues ne sont pas explicites.

Le résultat suivant est dû à Hairer et Mattingly [1]. Les hypothèses sont plus faciles à vérifier en pratique, et la majoration obtenue pour $\mathbb{E}^x[f(X_n)] - \pi(f)$ a l'avantage de faire intervenir des constantes explicites.

Théorème 1.4.1: Critère de convergence pour espérances

Supposons que les deux conditions suivantes soient satisfaites.

1. **Condition de dérive géométrique :** Il existe $d \geqslant 0$, $\gamma \in]0,1[$ et une fonction de Lyapounov V tels que
$$(\mathscr{L}V)(x) \leqslant -(1-\gamma)V(x) + d \qquad \forall x \in \mathscr{X} . \tag{1.4.1}$$
2. **Condition de minoration :** Pour un $R > 2d/(1-\gamma)$, soit $K = \{x \in \mathscr{X} : V(x) < R\}$. Alors il existe $\alpha \in]0,1[$ et une mesure de probabilité ν telle que
$$\inf_{x\in K} p_{xy} = \inf_{x\in K} \mathbb{P}^x\{X_1 = y\} \geqslant \alpha\nu(y) \qquad \forall y \in \mathscr{X} . \tag{1.4.2}$$

Alors il existe des constantes $M > 0$ et $\bar\gamma < 1$ telles que
$$\|\mathbb{E}^\cdot[f(X_n)] - \pi(f)\|_{1+V} \leqslant M\bar\gamma^n \|f - \pi(f)\|_{1+V} \tag{1.4.3}$$
pour toute fonction test $f \in \mathscr{E}_\infty^{1+V}$.

Précisons qu'on a
$$\|\mathbb{E}^\cdot[f(X_n)] - \pi(f)\|_{1+V} = \sup_{x\in\mathscr{X}} \frac{|\mathbb{E}^x[f(X_n)] - \pi(f)|}{1+V(x)} .$$
La majoration (1.4.3) peut donc s'écrire
$$\Bigl|\mathbb{E}^x[f(X_n)] - \pi(f)\Bigr| \leqslant (1+V(x))M\bar\gamma^n\|f-\pi(f)\|_{1+V} \qquad \forall x \in \mathscr{X} .$$
Comme $\bar\gamma < 1$, on a donc convergence exponentielle de $\mathbb{E}^x[f(X_n)]$ vers $\pi(f)$. Comme en pratique, on peut souvent choisir x tel que $V(x)$ ne soit pas trop grand, la dépendance en $V(x)$ ne pose pas de problème.

La présence de $\|f - \pi(f)\|_{1+V}$ n'est pas vraiment restrictive non plus. Si par exemple $f \in \mathscr{E}_\infty$, ou si f est à support compact (c'est-à-dire nulle en-dehors d'un ensemble compact), cette quantité est finie.

La condition de minoration (1.4.2) est un peu plus faible que la condition (1.2.2) du Théorème 1.2.5. Le point crucial est que l'on ait une borne inférieure sur les probabilités de transition qui soit indépendante du point de départ dans K.

Dans la suite, nous utiliserons les notations
$$(\mu\mathscr{P})(y) = \mathbb{P}^\mu\{X_1 = y\} = \sum_{x\in\mathscr{X}} \mu(x)p_{xy} ,$$
$$(\mathscr{P}f)(x) = \mathbb{E}^x[f(X_1)] = \sum_{y\in\mathscr{X}} p_{xy}f(y) .$$
La condition de dérive géométrique (1.4.1) est équivalente à
$$(\mathscr{P}V)(x) \leqslant \gamma V(x) + d \qquad \forall x \in \mathscr{X} .$$
L'ingrédient essentiel de la démonstation du Théorème 1.4.1 est la borne suivante.

Proposition 1.4.2: L'application $\mathscr{P}$ est contractante pour la distance $\rho_{1+\beta V}$

Il existe $\bar\gamma \in]0,1[$ et $\beta > 0$ tels que
$$\rho_{1+\beta V}(\mu\mathscr{P}, \nu\mathscr{P}) \leqslant \bar\gamma\rho_{1+\beta V}(\mu,\nu) .$$

En fait, la proposition donne des expressions explicites pour les constantes β et $\bar{\gamma}$: pour tout choix de α_0 et γ_0 tels que

$$0 < \alpha < \alpha_0 \qquad \text{et} \qquad \gamma + \frac{2d}{R} < \gamma_0 < 1\,, \tag{1.4.4}$$

on peut prendre

$$\beta = \frac{\alpha_0}{d} \qquad \text{et} \qquad \bar{\gamma} < \max\left\{1 - (\alpha - \alpha_0), \frac{2 + R\beta\gamma_0}{2 + R\beta}\right\}. \tag{1.4.5}$$

Ces expressions ne sont par particulièrement élégantes, mais elles ont le mérite d'être explicites, ce qui peut servir dans les applications. Notons que si d tend vers 0, alors β tend vers l'infini, et on peut prendre $\bar{\gamma}$ arbitrairement proche de γ.

Nous allons d'abord montrer que la Proposition 1.4.2 implique bien le Théorème 1.4.1.

Démonstration du Théorème 1.4.1. Nous donnons tout d'abord une démonstation de l'existence de π, même si celle-ci suit en fait du Théorème 1.2.5. C'est une application du théorème du point fixe de Banach.

Fixons $x_0 \in \mathscr{X}$, et soit $\mu_0^{x_0} = \delta_{x_0}$ la mesure de Dirac en x_0. Soit $\mu_n^{x_0} = \mu_0^{x_0}\mathscr{P}^n$. La proposition implique

$$\rho_{1+\beta V}(\mu_{n+1}^{x_0}, \mu_n^{x_0}) \leqslant \bar{\gamma}\rho_{1+\beta V}(\mu_n^{x_0}, \mu_{n-1}^{x_0}) \leqslant \dots \leqslant \bar{\gamma}^n \rho_{1+\beta V}(\mu_1^{x_0}, \mu_0^{x_0})\,.$$

On a donc une suite de Cauchy, et comme on sait que la distance en variation totale est complète, donc a fortiori la distance $\rho_{1+\beta V}$, on en conclut que la suite des $\mu_n^{x_0}$ converge vers une mesure π en variation totale. De plus, $\pi(1+V)$ est finie, car $\rho_{1+\beta V}(\pi, \mu_0^{x_0})$ l'est.

Pour montrer que π est invariante, il suffit d'observer que

$$\pi\mathscr{P} = \lim_{n\to\infty} \pi\mathscr{P}^{n+1} \lim_{n\to\infty} \pi\mathscr{P}^n = \pi\,.$$

Afin de démontrer (1.4.3), il est utile de centrer f. Posons donc

$$\hat{f}(x) = f(x) - \pi(f) \qquad \forall x \in \mathscr{X}\,.$$

Alors $\pi(\hat{f}) = \pi(f) - \pi(\pi(f)) = 0$, puisque $\pi(\pi(f)) = \pi(f)$. Ainsi

$$\mathscr{P}^n f - \pi(f) = \mathscr{P}^n \hat{f} + \mathscr{P}^n \pi(f) - \pi(\hat{f}) - \pi(\pi(f)) = \mathscr{P}^n \hat{f}\,.$$

Ceci permet d'écrire

$$\|\mathbb{E}^{\cdot}[f(X_n)] - \pi(f)\|_{1+V} = \|\mathscr{P}^n f - \pi(f)\|_{1+V} = \|\mathscr{P}^n \hat{f}\|_{1+V}\,.$$

Il s'agit donc de montrer qu'il existe une constante M telle que pour toute fonction test $\hat{f}$ satisfaisant $\pi(\hat{f}) = 0$, on ait

$$\|\mathscr{P}^n \hat{f}\|_{1+V} \leqslant M\bar{\gamma}^n \|\hat{f}\|_{1+V}\,.$$

Or, comme $(\mathscr{P}^n \hat{f})(x) = \delta_x(\mathscr{P}^n \hat{f}) = \mu_n^x(\hat{f})$ et $\pi(\hat{f}) = 0$, on a

$$\|\mathscr{P}^n \hat{f}\|_{1+\beta V} = \sup_{x\in\mathscr{X}} \frac{|(\mu_n^x - \pi)(\hat{f})|}{1 + \beta V(x)} \leqslant \|\hat{f}\|_{1+\beta V} \sup_{x\in\mathscr{X}} \frac{\rho_{1+\beta V}(\mu_n^x, \pi)}{1 + \beta V(x)}$$

en vertu de (1.3.3). Observons que

$$\begin{aligned}
\rho_{1+\beta V}(\mu_n^x, \pi) &= \rho_{1+\beta V}(\mu_n^x, \mu_{n+1}^x) + \rho_{1+\beta V}(\mu_{n+1}^x, \mu_{n+2}^x) + \dots \\
&\leqslant [\bar{\gamma}^n + \bar{\gamma}^{n+1} + \dots]\rho_{1+\beta V}(\mu_1^x, \mu_0^x) \\
&= \frac{\bar{\gamma}^n}{1 - \bar{\gamma}} \rho_{1+\beta V}(\mu_1^x, \mu_0^x)\,.
\end{aligned}$$

De plus,

$$
\begin{aligned}
\rho_{1+\beta V}(\mu_1^x, \mu_0^x) &= \sum_{y\in\mathscr{X}} (1+\beta V(y))|\mu_1^x(y) - \mu_0^x(y)| \\
&\leqslant \sum_{y\in\mathscr{X}} (1+\beta V(y))\mu_1^x(y) + 1 + \beta V(x) \\
&= (1 + \beta(\mathscr{P}V))(x) + 1 + \beta V(x)\,.
\end{aligned}
$$

Comme $(\mathscr{P}V)(x) = (\mathscr{L}V)(x) + V(x) \leqslant \gamma V(x) + d$ par l'hypothèse de dérive géométrique (1.4.1), on obtient finalement

$$
\|\mathscr{P}^n \hat{f}\|_{1+\beta V} \leqslant \|f\|_{1+\beta V} \frac{\bar{\gamma}^n}{1-\bar{\gamma}} \sup_{x\in\mathscr{X}} \frac{2+\beta[V(x)+\gamma V(x)+d]}{1+\beta V(x)} =: M\bar{\gamma}^n \|\hat{f}\|_{1+\beta V}\,.
$$

En particulier, la borne est vraie pour $\beta = 1$, ce qui conclut la démonstation. □

Il nous reste à démontrer la Proposition 1.4.2. L'idée est de travailler avec une définition alternative de la distance ρ_β. On introduit sur $\mathscr{X}$ la distance

$$
d_\beta(x,y) = \begin{cases} 0 & \text{si } x = y\,, \\ 2 + \beta V(x) + \beta V(y) & \text{si } x \neq y\,, \end{cases}
$$

(on vérifie facilement que d_β satisfait bien la définition d'une distance), et la semi-norme de Lipschitz

$$
|||f|||_\beta = \sup_{x\neq y} \frac{|f(x) - f(y)|}{d_\beta(x,y)}\,,
$$

C'est une semi-norme, et non une norme, car $|||f|||_\beta = 0$ n'implique pas $f = 0$ (mais seulement que f est constante). On a alors le résultat suivant.

Lemme 1.4.3: Équivalence des distances

$$
\rho_{1+\beta V}(\mu,\nu) = \sup_{f:|||f|||_\beta\leqslant 1} \sum_{x\in\mathscr{X}} f(x)(\mu(x) - \nu(x))\,.
$$

Notons que la seule différence entre la définition (1.3.1) de $\rho_{1+\beta V}(\mu,\nu)$ et cette égalité est l'ensemble des f sur lequel on prend le supremum.

Montrons tout d'abord que ce lemme implique la proposition.

Démonstration de la Proposition 1.4.2. L'idée est de montrer que $\mathscr{P}$ est contractante dans la seminorme $|||\cdot|||_\beta$, à savoir

$$
|||\mathscr{P}f|||_\beta \leqslant \bar{\gamma} |||f|||_\beta\,. \tag{1.4.6}
$$

En effet, ceci implique

$$
\begin{aligned}
\rho_{1+\beta V}(\mu\mathscr{P}, \nu\mathscr{P}) &= \sup_{f\,:\,|||f|||_\beta\neq 0} \frac{1}{|||f|||_\beta} (\mu\mathscr{P} - \nu\mathscr{P})(f) \\
&= \sup_{f\,:\,|||f|||_\beta\neq 0} \frac{1}{|||f|||_\beta} (\mu - \nu)(\mathscr{P}f) \\
&\leqslant \sup_{f_1\,:\,|||f_1|||_\beta\neq 0} \frac{\bar{\gamma}}{|||f_1|||_\beta} (\mu - \nu)(f_1) \\
&= \bar{\gamma}\rho_{1+\beta V}(\mu,\nu)\,.
\end{aligned}
$$

Afin de démontrer (1.4.6), nous fixons un f tel que $\|f\|_{1+\beta V} \leqslant 1$ et $|||f|||_\beta \leqslant 1$. Il s'agit de montrer que

$$\left|(\mathscr{P}f)(x) - (\mathscr{P}f)(x)\right| \leqslant \bar{\gamma} d_\beta(x,y)$$

pour tout $x,y \in \mathscr{X}$. La relation est vraie pour $x = y$, donc nous supposons $x \neq y$. Nous considérons deux cas séparément.

- **Cas 1:** $V(x) + V(y) \geqslant R$. Dans ce cas, on a

$$\left|(\mathscr{P}f)(x)\right| = \left|\sum_{y\in\mathscr{X}} p_{xy} f(y)\right| \leqslant \underbrace{\|f\|_{1+\beta V}}_{\leqslant 1} \sum_{y\in\mathscr{X}} (1+\beta V(y)) p_{xy} \leqslant 1 + \beta(\mathscr{P}V)(x) \,, \tag{1.4.7}$$

  d'où, par la condition (1.4.1) de dérive géométrique,

$$\begin{aligned}\left|(\mathscr{P}f)(x) - (\mathscr{P}f)(y)\right| &\leqslant 2 + \beta(\mathscr{P}V)(x) + \beta(\mathscr{P}V)(y) \\ &\leqslant 2 + \beta\gamma V(x) + \beta\gamma V(y) + 2\beta d \\ &\leqslant 2 + \beta\gamma_0[V(x) + V(y)]\end{aligned}$$

  pour tout γ_0 satisfaisant (1.4.4). Un calcul élémentaire montre alors qu'en posant

$$\gamma_1 = \frac{2 + \beta R \gamma_0}{2 + \beta R} \,,$$

  on obtient bien la majoration requise

$$\left|(\mathscr{P}f)(x) - (\mathscr{P}f)(y)\right| \leqslant \gamma_1(2 + V(x) + V(y)) = \gamma_1 d_\beta(x,y) \,.$$

- **Cas 2:** $V(x) + V(y) < R$. Dans ce cas, on a $x,y \in K$. La matrice $\widetilde{\mathscr{P}}$ d'éléments

$$\tilde{p}_{xy} = \frac{1}{1-\alpha} p_{xy} - \frac{\alpha}{1-\alpha} \nu(y)$$

  est une matrice stochastique. En effet, la condition de minoration (1.4.2) montre que ces éléments sont tous positifs ou nuls, et on vérifie immédiatement que la somme sur y des $\tilde{p}_{xy}$ vaut 1. De plus, on a

$$(\mathscr{P}f)(x) = (1-\alpha)(\widetilde{\mathscr{P}}f)(x) + \alpha\nu(f) \,,$$

  d'où, par un calcul analogue à (1.4.7),

$$\begin{aligned}\left|(\mathscr{P}f)(x) - (\mathscr{P}f)(y)\right| &= (1-\alpha)\left|(\widetilde{\mathscr{P}}f)(x) - (\widetilde{\mathscr{P}}f)(y)\right| \\ &\leqslant (1-\alpha)[2 + \beta(\widetilde{\mathscr{P}}V)(x) + \beta(\widetilde{\mathscr{P}}V)(y)] \\ &\leqslant (1-\alpha) + \beta(\mathscr{P}V)(x) + \beta(\mathscr{P}V)(y) \\ &\leqslant 2(1-\alpha) + \beta\gamma[V(x) + V(y)] + 2\beta d \\ &\leqslant \gamma_2 d_\beta(x,y) \,,\end{aligned}$$

  où β est donné par (1.4.5) et $\gamma_2 = \max\{\gamma, 1-(\alpha-\alpha_0)\}$. La troisième ligne suit du fait que

$$(\widetilde{\mathscr{P}}V)(x) \leqslant \frac{1}{1-\alpha}(\mathscr{P}V)(x) \,.$$

Le résultat suit, avec $\bar{\gamma} = \max\{\gamma_1, \gamma_2\}$. □

Bibliographie

[1] Martin Hairer and Jonathan C. Mattingly. Yet another look at Harris' ergodic theorem for Markov chains. In *Seminar on Stochastic Analysis, Random Fields and Applications VI*, volume 63 of *Progr. Probab.*, pages 109–117. Birkhäuser/Springer Basel AG, Basel, 2011.

[2] Sean P. Meyn and R. L. Tweedie. Stability of Markovian processes. I. Criteria for discrete-time chains. *Adv. in Appl. Probab.*, 24(3):542–574, 1992.

Nils Berglund
Institut Denis Poisson (IDP)
Universite d'Orleans, Universite de Tours, CNRS – UMR 7013
Bâtiment de Mathematiques, B.P. 6759
45067 Orleans Cedex 2, France
E-mail address: `nils.berglund@univ-orleans.fr`
`https://www.idpoisson.fr/berglund`