Francis Comets — Université Paris 7 — Denis Diderot
Notes de cours M2 — 12 novembre 2007

Chapitre 6

Couplage et Chaînes de Markov

Cadre : Q matrice stochastique sur un espace discret E.

1 Couplage

DÉFINITION 1.1 *Un couplage de deux variables aléatoires X, Y est un couple $(\bar{X}, \bar{Y})$ défini sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbf{P})$ avec $\bar{X} \overset{loi}{=} X, \bar{Y} \overset{loi}{=} Y$.*

Ainsi, un couplage de deux variables aléatoires est une réalisation de ces variables sur un *même* espace. La dépendance entre les deux variables aléatoires permettra par exemple de les comparer.

EXEMPLE 1.1 *Soit $X \sim$ Bernoulli(p), $Y \sim$ Bernoulli(p'), $0 \leq p \leq p' \leq 1$. Alors,*

$$\bar{X} = \mathbf{1}\{U \leq p\}, \bar{Y} = \mathbf{1}\{U \leq p'\}$$

est un couplage de X et Y sur l'espace $(\Omega, \mathcal{A}, \mathbf{P})$ où la variable uniforme U est définie. On a $\bar{X} \leq \bar{Y}$ p.s.

Soit Q est une matrice stochastique sur E ; il y a plusieurs manières de coupler les chaînes de Markov $(X_n^a)_n$ et $(X_n^b)_n$ issues de deux conditions initiales distinctes $a, b \in E$. Fixons une représentation de la chaîne comme suite récurrente aléatoire, i.e. fixons $F : E \times [0, 1[\to E$ telle que

$$X_{n+1} = F(X_n, U_n)$$

soit une chaîne de transition Q ($(U_n)_n$ est i.i.d. uniforme sur $[0, 1[$). Voila des **exemples** importants **de couplages** $((X_n)_n, (Y_n)_n)$ de $(X_n^a)_n, (X_n^b)_n$ (avec toujours $X_0 = a, Y_0 = b$) :

1. *Couplage libre* :
$$X_{n+1} = F(X_n, U_n) \ , \quad Y_{n+1} = F(Y_n, U_n)$$
On utilise le même U, comme dans l'exemple 1.1
2. *Couplage antithétique* :
$$X_{n+1} = F(X_n, U_n) \ , \quad Y_{n+1} = F(Y_n, 1 - U_n)$$
Utilisé dans les méthodes de réduction de la variance.

3. *Couplage indépendant*: Avec $(U_n)_n$ et $(U'_n)_n$ deux suites i.i.d. uniformes et indépendantes entre elles,
$$X_{n+1} = F(X_n, U_n) \ , \quad Y_{n+1} = F(Y_n, U'_n)$$
Les deux chaînes sont indépendantes.
4. *Couplage indépendant-coalescent* (Doeblin 1938) : On définit
$$V_n = \begin{cases} U_n & \text{si} \quad X_n = Y_n \\ U'_n & \text{si} \quad X_n \neq Y_n \end{cases}$$
et
$$X_{n+1} = F(X_n, U_n) \ , \quad Y_{n+1} = F(Y_n, V_n)$$
On va voir que la suite (V_n) est i.i.d. uniforme. On a donc bien un couplage, mais ici X_n et Y_n évoluent indépendamment jusqu'à ce qu'ils se rencontrent, pour rester ensuite toujours ensemble.

□ Pour montrer que la suite (V_n) définie plus haut est i.i.d. uniforme, il suffit de montrer que la loi de V_n sachant $V_0, \dots V_{n-1}$ est uniforme sur $[0,1[$. En fait on va considérer plus generalement, avec la notation $u_l^k = (u_l, \dots u_k)$ pour $0 \leq l \leq k$,
$$\mathbf{P}(V_n < u | U_0^{n-1}, U_0^{n-1\prime}) = \left\{ \begin{array}{ll} \mathbf{P}(U_n < u | U_0^{n-1}, U_0^{n-1\prime}) & \text{sur} \quad \{X_n = Y_n\} \\ \mathbf{P}(U'_n < u | U_0^{n-1}, U_0^{n-1\prime}) & \text{sur} \quad \{X_n \neq Y_n\} \end{array} \right\} = u$$
■

2 Temps de couplage

On considère un couplage de X_n^a et X_n^b, que l'on notera (X_n^a, X_n^b) "sans les ˜" pour simplifier les notations.

DÉFINITION 2.1 *Le temps de rencontre du couplage (X_n^a, X_n^b) est*
$$\tau^{ab} = \inf\{n \geq 0; X_n^a = X_n^b\}$$

Il dépend du couplage choisi :

EXEMPLE 2.1 *Considérons la promenade aléatoire simple sur $\mathbb{Z}$ partant de $a = 0$ et $b = 2$. Pour le couplage libre, $\tau^{ab} = \infty$ p.s., tandis que pour le couplage indépendant, $\tau^{ab} < \infty$ est la moitié du temps de premier passage en 2 pour une promenade aléatoire simple.*

Attention, il se peut qu'il n'existe pas de couplage avec $\tau^{ab} < \infty$ p.s.

Exemple 2.2 *(Chaine de Wright-Fisher) : $E = \{0, 1, \ldots, m\}$, et $Q(x, \cdot) = \mathcal{B}(m, x/m)$. Dans ce cas, les états 0 et m sont absorbants, et la propriété*

$$\mathbf{P}(X^a \text{absorbé en } m) = \frac{a}{m} \neq \mathbf{P}(X^b \text{absorbé en } m) \;, \quad a \neq b, a, b \in E$$

entraine que

$$\begin{aligned}\mathbf{P}(\tau^{ab} = \infty) &\geq \mathbf{P}(X^a, X^b \text{absorbés en des points différents}) \\ &\geq |\mathbf{P}(X^a \text{absorbé en } m) - \mathbf{P}(X^b \text{absorbé en } m)| \\ &> 0 \quad \text{si } a \neq b \;.\end{aligned}$$

■

Vu l'exemple 4 ci-dessus, on pourra se limiter à des couplages particuliers :

Un couplage est dit *coalescent* si

$$n \geq \tau^{ab} \implies X_n^a = X_n^b$$

et alors on appelle plutôt τ^{ab} le *temps de couplage*.

On voit que les couplages libre et indépendant-coalescent sont des couplages coalescents.

Il y a un lien avec la distance en variation entre les lois des deux chaînes de Markov. Rappelons quelques définitions et formules. Soient $\mu, \nu \in \mathcal{P}(E)$. La *distance en variation* entre ces lois est

$$\begin{aligned}\|\mu - \nu\|_{\text{var}} &:= \frac{1}{2} \sum_x |\mu(x) - \nu(x)| \;\equiv\; \frac{1}{2} \|\mu - \nu\|_{\ell_1(E)} \\ &= \frac{1}{2} \sup\{|\sum_x f(x)\mu(x) - \sum_x f(x)\nu(x)|; f : E \to \mathbb{R}, \|f\|_\infty \leq 1\} \\ &= \sup\{|\mu(A) - \nu(A)|; A \subset E\}\end{aligned}$$

(le premier sup est atteint pour $f(x) = \text{signe}(\mu(x) - \nu(x))$, le deuxième pour $A = \{x : \mu(x) > \nu(x)\}$). Un point important est qu'il s'agit d'une norme sur l'espace des mesures finies sur E : cette distance est égale à la moitié de la norme de la variation totale de la mesure $\mu - nu$ (i.e., de la norme dans $\ell_1(E)$ du vecteur $\mu - nu$). En particulier, $0 \leq \|\mu - \nu\|_{\text{var}} \leq 1$.

Proposition 2.1 *(**Inégalité du couplage**). Avec μ_n^a la loi de X_n^a, on a*

$$\|\mu_n^a - \mu_n^b\|_{\text{var}} \leq \mathbf{P}(\tau^{ab} > n)$$

pour tout couplage coalescent τ^{ab}.

□ On utilise la deuxième formule : l'intéret de disposer d'un couplage est de pouvoir mettre la différence d'espérances sous le même signe espérance,

$$\begin{aligned}
\sum_x f(x)\mu_n^a(x) - \sum_x f(x)\mu_n^b(x) &= \mathbf{E}f(X_n^a) - \mathbf{E}f(X_n^b)] \\
&= \mathbf{E}[f(X_n^a) - f(X_n^b)] \\
&= \mathbf{E}[(f(X_n^a) - f(X_n^b))(\mathbf{1}\{\tau^{ab} \leq n\} + \mathbf{1}\{\tau^{ab} > n\})] \\
&= \mathbf{E}[(f(X_n^a) - f(X_n^b))\mathbf{1}\{\tau^{ab} > n\}] \\
&\leq 2\|f\|_\infty \mathbf{P}(\tau^{ab} > n)
\end{aligned}$$

■

3 Couplage réussi

Lorsque $\tau^{ab} < \infty$ p.s., on dira que le couplage a réussi. Dans ce cas, on a necessairement

$$\mu_n^a - \mu_n^b \to 0$$

quand $n \to \infty$, propriété que l'on qualifie de *perte de mémoire* : avec le temps, la chaîne oublie quel est son point de départ. La queue de distribution $\mathbf{P}(\tau^{ab} > n)$ du temps de rencontre permet de quantifier la vitesse à laquelle se perd la mémoire. On a donc intéret à considerer un couplage pour lequel le temps de rencontre est court.

Lorsque $\tau^{ab} < \infty$ p.s. pour tous $a, b \in E$ et s'il existe une probabilité invariante π, on fait démarrer Y non pas d'un état fixé b mais de la loi π, et on montre alors que $\mu_n^a \to \pi$ à une vitesse qui est controlée par les queues $\mathbf{P}(\tau^{ab} > n)$.

Proposition 3.1 *Si Q admet une probabilité invariante π, alors pour tout $a \in E$,*

$$\mu_n^a \to \pi\ , \qquad n \to \infty\ .$$

On a de plus

$$\|\mu_n^a - \pi\|_{\mathrm{var}} \leq \sup\{\mathbf{P}(\tau^{ab} > n); b \in E\}$$

□ Puisque π est invariante, on a $\sum_b \pi(b)\mu_n^b = \pi$, et par convexité de la norme

$$\begin{aligned}
\|\mu_n^a - \pi\|_{\mathrm{var}} &= \|\sum_b \pi(b)[\mu_n^a - \mu_n^b]\|_{\mathrm{var}} \\
&\leq \sum_b \pi(b)\|\mu_n^a - \mu_n^b\|_{\mathrm{var}} \\
&\leq \sum_b \pi(b)\mathbf{P}(\tau^{ab} > n) \qquad (3.1) \\
&\leq \sup\{\mathbf{P}(\tau^{ab} > n); b \in E\}
\end{aligned}$$

où l'inégalité (3.1) résulte de l'inégalité de couplage (2.1). D'autre part, on peut majorer (3.1) différemment, en utilisant des parties finies $K \subset E$:

$$\|\mu_n^a - \pi\|_{\text{var}} \leq \sup\{\mathbf{P}(\tau^{ab} > n); b \in K\} + \pi(K^c) ,$$

si bien que

$$\lim_{n\to\infty} \|\mu_n^a - \pi\|_{\text{var}} \leq \lim_{K \nearrow E} \limsup_{n\to\infty} \big(\sup\{\mathbf{P}(\tau^{ab} > n); b \in K\} + \pi(K^c)\big) = 0$$

■

COROLLAIRE 3.1 *S'il existe des couplages réussis $\tau^{ab} < \infty$ p.s. pour tous $a, b \in E$, Q admet au plus une une probabilité invariante.*

□ En effet, la dernière proposition entraine que toute probabilité invariante est telle que

$$\pi = \lim_{n\to\infty} \mu_n^a$$

et l'affirmation découle de l'unicité de la limite. ■

4 Coefficient d'ergodicité de Doeblin

Il est défini par

$$\beta(Q) = \sum_{y \in E} \inf_{x \in E} Q(x, y) \in [0, 1] \tag{4.2}$$

EXEMPLE 4.1 *Chateau de cartes : $E = \mathbb{N}$, $Q(x, x+1) = p_x, Q(x, 0) = 1 - p_x$.*
Dans ce cas, on a $\beta = 1 - \sup_x(p_x)$. Si $\beta > 0$, on a, à partir de tout état x, une probabilité uniformément positive (et bornée par β) de retourner à 0. On peut alors attendre le premier temps où les deux chaines issues d'états différents retournent ensemble à 0 (ce qui se produit avec probabilité au moins β^2 à chaque instant si les chaines sont indépendantes), et ensuite les faire rester ensemble.

Le résultat suivant est bien meilleur, il utilise de manière avantageuse la possibilité d'instaurer une dépendance entre les chaines.

THÉORÈME 4.1 *Supposons $\beta(Q) = \beta > 0$. Alors, il existe une fonction $F : E \times [0, 1[\to E$ telle que pour le couplage libre on ait*

$$\mathbf{P}(\sup_{a,b \in E} \tau^{ab} > n) \leq (1 - \beta)^n$$

□ Quitte à numéroter les points, on peut supposer que $E = \{1, 2, \ldots\}$. On définit

$$q(y) = \inf_z Q(z, y)$$

On fixe $x \in E$, et on construit $F(x, \cdot)$ en partitionnant [0,1[en intervalles (fermés à gauche, ouverts à droite, indexés par $y \in E$) construits récursivement comme indiqués figure 1.

– Pour $y = 1$: on considère les 2 intervalles

$$\left\{ \begin{array}{l} I_1^{x,-} \\ I_1^{x,+} \end{array} \right. \text{ situé à} \left\{ \begin{array}{l} \text{droite de } 0 \\ \text{gauche de } 1 \end{array} \right. \text{ de longueur } \left\{ \begin{array}{c} q(1) \\ Q(x,1) - q(1) \end{array} \right.$$

Notons que $Q(x, 1) - q(1) \geq 0$, et que

$$I_1^x = I_1^{x,+} \bigcup I_1^{x,-} \text{est de longueur Q(x, 1)} \tag{4.3}$$

– Pour $y = 2$: on considère les 2 intervalles

$$\left\{ \begin{array}{l} I_2^{x,-} \\ I_2^{x,+} \end{array} \right. \text{ situé à } \left\{ \begin{array}{l} \text{droite de } I_1^{x,-} \\ \text{gauche de } I_1^{x,+} \end{array} \right. \text{ de longueur } \left\{ \begin{array}{c} q(2) \\ Q(x,2) - q(2) \end{array} \right.$$

Comme plus haut, $I_2^x = I_2^{x,+} \bigcup I_2^{x,-}$ est de longueur $Q(x, 2)$.

– Etc ... pour $y \geq 1$

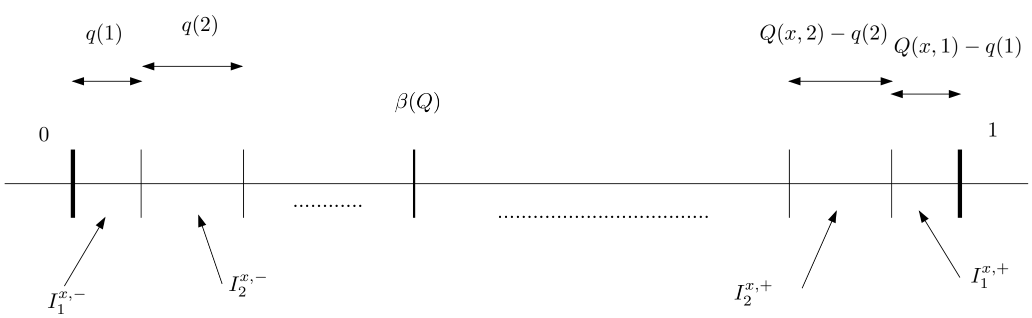

FIG. 1 – **Partition de [0,1**[pour construire la fonction F

On répète la construction pour tout $x \in E$. Observons que

$$[0, 1[\setminus\{\beta(Q)\} \subset \bigcup_{y \in E} I_y^x \subset [0, 1[\quad , \qquad I_y^{x,-} \text{ne dépend pas de } x. \tag{4.4}$$

Plus précisement, $\bigcup_{y \in E} I_y^x = [0, 1[$ si E est fini, et $\bigcup_{y \in E} I_y^x = [0, 1[\setminus\{\beta(Q)\}$ sinon. On définit alors $F : E \times [0, 1[\to E$ par

$$F(x, u) = y \quad \text{avec} \quad u \in I_y^x \tag{4.5}$$

i.e., $F(x, u)$ est l'indice y de l'intervalle où tombe u. Dans le cas où $\bigcup_{y \in E} I_y^x = [0, 1[\setminus\{\beta(Q)\}$, il faut aussi attribuer une valeur à $F(x, \beta(Q))$ afin que F soit bien définie partout ; mais cette valeur n'a aucune importance puisque les U_n sont des variables aléatoires à densités. D'après (4.3), on a $\mathbf{P}(F(x,U)=y)=|I_y^x|=Q(x,y)$, et la suite récurrente aléatoire associée à F est de transition Q. D'après la deuxième partie de (4.4), on a la propriété remarquable

$$u<\beta \implies F(x,u)=F(x',u)\ \forall x,x'\in E \tag{4.6}$$

Il s'en suit que pour tout $a,b\in E$,

$$\tau^{ab}\le\overline{\tau}:=\inf\{n\ge 0; U_n<\beta\}$$

Mais $\overline{\tau}$ suit la loi géométrique $\mathbf{P}(\overline{\tau}>n)=(1-\beta)^n, n\ge 0$, et le théorème est montré. ■

Ce couplage est appelé le *couplage maximal.*

En corollaire du théorème, on déduit directement des résultats généraux des proposition 3.1 et corollaire 3.1 l'énoncé suivant :

Corollaire 4.1 *Si $\beta(Q)>0$ et si Q admet une probabilité invariante π, elle est nécessairement unique et alors pour tout $a\in E$,*

$$\|\mu_n^a-\pi\|_{\text{var}}\le[1-\beta(Q)]^n$$

Ainsi, on a convergence à vitesse exponentielle vers l'état stationnaire. Notons que sous l'hypothèse $\beta(Q)>0$, la chaine n'est pas forcément irreductible : par exemple $Q(x,\cdot)=\nu$ avec $\nu\in\mathcal{P}(E)$ est telle que $\beta=1$ même si ν s'annulle en certains points de E.

Remarque 4.1 *En fait, la condition $\beta(Q)>0$ entraine qu'il existe une probabilité invariante. (Laissé en exo)*

5 Cas d'un espace d'état fini

On suppose ici E fini. Par compacité, il existe une probabilité invariante, qui sera unique si $\beta(Q)>0$ d'après le corollaire. Mais il est plus réaliste de faire l'hypothèse suivante :

Proposition 5.1 *Soit Q une matrice de transition sur E fini, telle que*

$$\exists k\ge 1:\qquad Q^k(x,y)>0\quad\forall x,y\in E \tag{5.7}$$

Alors, Q est irréductible, elle admet une unique probabilité invariante π, et $\exists\delta>0$ tel que pour tout $n\ge 0$,

$$\|\mu_n^a-\pi\|_{\text{var}}\le[1-\delta]^{n-(k-1)}$$

L'hypothèse (5.7) est vérifiée si Q est irréductible apériodique. La preuve de cette dernière affirmation est élementaire, mais nous ne la donnerons pas ici (voir [1] p.42-43). $\Box$ Puisque E est fini, l'hypothèse (5.7) implique que

$$\beta(Q^k) \geq \min\{Q^k(x,y); x,y \in E\} > 0 \ ,$$

de sorte que pour tout entier m,

$$\inf_{x\in E} \|\mu^x_{mk} - \pi\|_{\mathrm{var}} \leq [1-\beta(Q^k)]^m$$

d'après le théorème 4.1. Par ailleurs, $\sup_x \|\mu^x_n - \pi\|_{\mathrm{var}}$ est croissant en n, puisque

$$\begin{aligned}
\|\mu^x_{n+1} - \pi\|_{\mathrm{var}} &= \|\sum_{y\in E} Q(x,y)\mu^y_n - \pi\|_{\mathrm{var}} \\
&= \|\sum_{y\in E} Q(x,y)[\mu^y_n - \pi]\|_{\mathrm{var}} \\
&\leq \sum_{y\in E} Q(x,y)\|\mu^y_n - \pi\|_{\mathrm{var}} \\
&\leq \sup_{y\in E} \|\mu^y_n - \pi\|_{\mathrm{var}}
\end{aligned}$$

par convexité de la norme. Par conséquent, $1-\delta := [1-\beta(Q^k)]^{1/k}$ convient. ■

Voila un résultat dans le cas périodique, indiquant comment se comporte la loi de la chaine à l'instant t.

La période de $x \in E$ est défini par

$$d(x) = \mathrm{PGCD}\{n \geq 0; Q^n(x,x) > 0\}$$

Si Q est irreductible, on montre que tous les points x ont même période, appelé période de Q. Alors, il existe une partition de E, $E = E_1 \bigcup E_2 \ldots \bigcup E_d$ telle que $Q : E_1 \to E_2 \to \ldots \to E_d \to E_1$ (i.e., $Q(x,y) = 0$ si $x \in E_i, y \in E_j$ avec $j \neq i+1$ modulo d).

Proposition 5.2 *Soit Q irréductible périodique sur E, de période $d \geq 2$. Alors,*

1. *Q^d est apériodique (et irréductible) sur chacun des espaces $E_i, i \leq d$;*
2. *Notant π_i l'unique probabilité invariante de Q^d sur $E_i (i = 1, \ldots d)$, la chaine Q a pour unique probabilité invariante*

$$\pi = \frac{1}{d}\sum_{i=1}^d \pi_i$$

3. *Il existe $\delta > 0$ tel que*

$$\|\mu^a_{nd} - \pi_i\|_{\mathrm{var}} \leq [1-\delta]^n \qquad \forall a \in E_i,\ i \leq d$$

Pour une preuve, voir [1], section 3.3 Ch.3.

6 Coefficient d'ergodicité de Dobrushin

Il s'agit d'une amélioration du précédent, défini par

$$\alpha(Q) = \inf_{a,b\in E} \sum_{y\in E} \min\{Q(a,y), Q(b,y)\} \in [0,1] \tag{6.8}$$

Clairement, $\alpha(Q) \geq \beta(Q)$. Il est adapté à la construction suivante, dans laquelle on se contente de coupler deux chaines à la fois, et non toutes. En contrepartie, il faudra représenter la chaine double (X_n^a, X_n^b) comme une suite récurrente aléatoire (cf (6.12)).

On note

$$q^{a,b}(y) := \min\{Q(a,y), Q(b,y)\}\ , \quad q^{a,b} := \sum_y q^{a,b}(y) \geq \alpha(Q)$$

Couplage de Dobrushin. Pour tous a, b fixés dans E, on construit une fonction $F(a, b; u)$ à l'aide de deux partitions [0,1[en intervalles (fermés à gauche, ouverts à droite, indexés par $y \in E$).

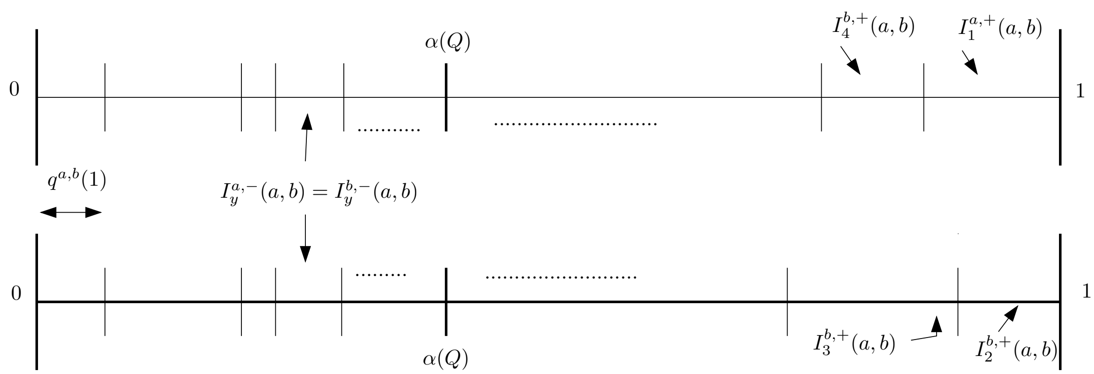

FIG. 2 – **Partitions de [0,1**[, couplage de Dobrushin. La ligne du haut représente les $I_y^{a,\pm}(a,b)$, celle du bas les $I_y^{b,\pm}(a,b)$

On pose $I_0^{a,-}(a,b) = I_0^{b,-}(a,b) = \{0\}$, $I_0^{a,+}(a,b) = I_0^{b,+}(a,b) = \{1\}$.

Pour $y = 1, 2, \ldots$, on considère récursivement, d'abord les 2 intervalles $I_y^{a,-}(a,b), I_y^{a,+}(a,b)$,

$$\left\{\begin{array}{c} I_y^{a,-}(a,b) \\ I_y^{a,+}(a,b) \end{array}\right. \text{situé à} \left\{\begin{array}{c} \text{droite de } I_{y-1}^{a,-}(a,b) \\ \text{gauche de } I_{y-1}^{a,+}(a,b) \end{array}\right. \text{de longueur } \left\{\begin{array}{c} q^{a,b}(y) \\ Q(a,y) - q^{a,b}(y) \end{array}\right. ,$$

et similairement $I_y^{b,-}(a,b), I_y^{b,+}(a,b)$,

$$\left\{\begin{array}{c} I_y^{b,-}(a,b) \\ I_y^{b,+}(a,b) \end{array}\right. \text{situé à} \left\{\begin{array}{c} \text{droite de } I_{y-1}^{b,-}(a,b) \\ \text{gauche de } I_{y-1}^{b,+}(a,b) \end{array}\right. \text{de longueur } \left\{\begin{array}{c} q^{a,b}(y) \\ Q(b,y) - q^{a,b}(y) \end{array}\right. .$$

Pout tout y, l'un des deux intervalles $I_y^{a,+}(a,b), I_y^{b,+}(a,b)$ est vide. Cette fois, les ensembles

$$I_y^a(a,b) = I_y^{a,-}(a,b) \bigcup I_y^{a,+}(a,b) \quad , \qquad I_y^b(a,b) = I_y^{b,-}(a,b) \bigcup I_y^{b,+}(a,b) \tag{6.9}$$

sont de longueurs respectives $Q(a,y), Q(b,y)$, et

$$[0,1[\backslash\{q^{a,b}\} \subset \bigcup_{y\in E} I_y^a(a,b) = \bigcup_{y\in E} I_y^b(a,b) \subset [0,1[\quad , \qquad I_y^{a,-}(a,b) = I_y^{b,-}(a,b) \; . \tag{6.10}$$

La fonction $F : E\times E\times [0,1[\to E\times E$,

$$F(a,b;u) = (y,z) \quad \text{avec} \quad u \in I_y^a(a,b), u \in I_z^b(a,b) \; ,$$

est bien définie, le couple $F(a,b;U_0)$ est bien un couplage de X_1^a, X_1^b, et on a la propriété remarquable

$$u < q^{a,b} \implies \text{les coordonnées de } F(a,b;u) \text{ sont égales} \tag{6.11}$$

d'après la deuxième partie de (6.10). La suite récurrente aléatoire

$$(X_{n+1}^a, X_{n+1}^b) := F(X_n^a, X_n^b; U_n) \tag{6.12}$$

est un couplage des chaines issues de a et b, on l'appelle le couplage de Dobrushin. A la différence du couplage maximal, les intervalles de coincidence changent à chaque pas.

Théorème 6.1 *Pour le couplage de Dobrushin,*

$$\mathbf{P}(\sup_{a,b\in E} \tau^{ab} > n) \leq [1-\alpha(Q)]^n$$

□ Avec

$$\overline{\tau^{ab}} = \inf\{n \geq 1 : U_{n-1} < q^{X_{n-1}^a, X_{n-1}^b}\} \; ,$$

il résulte de (6.11) que $\tau^{ab} \leq \overline{\tau^{ab}}$. (On a même égalité, si $a \neq b$.) On calcule alors

$$\begin{aligned} \mathbf{P}(\overline{\tau^{ab}} > n) &= \mathbf{P}(U_0 \geq q^{X_0^a, X_0^b}, \ldots U_{n-1} \geq q^{X_{n-1}^a, X_{n-1}^b}) \\ &\leq \mathbf{P}(U_0 \geq \alpha(Q), \ldots U_{n-1} \geq \alpha(Q)) \\ &= [1-\alpha(Q)]^n \end{aligned}$$

car $q^{(a,b)} \geq \alpha(Q)$. ■

En répétant les calculs du théorème 4.1, on obtient une majoration plus précise :

Théorème 6.2 *Si Q admet une probabilité invariante π, alors pour tout $a \in E$,*

$$\|\mu_n^a - \pi\|_{\mathrm{var}} \leq [1-\alpha(Q)]^n$$

Exercice : Soit $E = \{1,2\}$, et $Q = \begin{pmatrix} 1/3 & 2/3 \\ 2/3 & 1/3 \end{pmatrix}$

1. Montrer qu'il existe n_0 tel que $\forall n \geq n_0$,

$$\begin{aligned} 0.49 &\leq Q^n(1,1) \leq 0.51 \\ 0.49 &\leq Q^n(2,1) \leq 0.51 \end{aligned}$$

2. Trouver un majorant de n_0.

(Indications :

• la probabilite π uniforme sur E est invariante, et Q a tous ses coefficients strictement positifs.

• $\beta(Q) = \alpha(Q) = 2/3$: on résoud $.01 \geq (1-\alpha)^n$, soit $n_0 \geq [2\log 10/\log 3]$, soit $n_0 = 6$.)

Exercice : Chateau de cartes : vérifier que $\beta = \alpha = \inf_x q_x$

Références

[1] R. Fernandez, P. A. Ferrari, A. Galves : *Coupling, renewal and perfect simulation of chains of infinite order.* Notes for the V Brazilian School of Probability, Ubatuba, August 2001. `http ://www.ime.usp.br/ pablo/abstracts/vebp.html`

[2] H. Thorisson : *Coupling, stationarity, and regeneration.* Probability and its Applications. Springer-Verlag, New York, 2000