A characterisation of transient random walks on stochastic matrices with Dirichlet distributed limits

Shaun McKinlay

Australia New Zealand Applied Probability Workshop
University of Queensland

9 July 2013

Products of random i.i.d. stochastic matrices

Let $\{X(n)\}_{n\geq 1}$ be a sequence of random i.i.d. $d \times d$ stochastic matrices. We consider the limit of the left products

$$X(n,1) := X(n)X(n-1)\cdots X(1)$$

as $n \to \infty$ for a certain class of random stochastic matrices $X(1)$.

The right product is given by

$$X(1,n) := X(1)X(2)\cdots X(n) \stackrel{d}{=} X(n,1).$$

These products generate the left and right random walks

$$n \mapsto X(n,1) \quad \text{and} \quad n \mapsto X(1,n),$$

respectively.

A theorem by Chamayou and Letac (1994)

In Chamayou and Letac (1994) ("CL94"), the authors study the left products for random stochastic matrices X satisfying:

[I] The rows of X are independent.

[II] The rows of X are Dirichlet distributed.

[III] Letting $(\alpha_{i,1}, \ldots, \alpha_{i,d})$ be the Dirichlet parameters of the ith row of X, we have $\sum_{j=1}^{d} \alpha_{i,j} = \sum_{j=1}^{d} \alpha_{j,i}$ for $i = 1, \ldots, d$.

They show that the above conditions are sufficient to ensure that:

[A1] The products $X(n, 1)$ converge a.s. to some random matrix $\widehat{X}$ as $n \to \infty$.

[A2] The limit $\widehat{X}$ has identical rows a.s.

[A3] The rows of $\widehat{X}$ are Dirichlet distributed.

This extends a result by Van Assche (1986) who proved it for $d = 2$ and all $\alpha_{i,j} = p > 0$. Volodin, Kotz and Johnson (1993) also independently proved this for all $\alpha_{i,j} = p > 0$, and any $d \geq 2$.

A characterisation theorem

It turns out assertions [**A1**]–[**A3**] remain true under much broader conditions.

Denote by $\mathscr{K}_d$ the class of all distributions of a random $d \times d$ stochastic matrix X such that [**A1**]–[**A3**] hold.

We extend the result in CL94 by providing a charaterisation theorem for the class $\mathscr{K}_d$.

Some notation

We denote matrix row and column sums of $A = (\alpha_{i,j})_{i=1}^{r}{}_{j=1}^{c}$ by

$$p_{i\bullet} := \sum_{j=1}^{c} p_{i,j} \text{ for } i = 1, \ldots, r$$

$$p_{\bullet j} := \sum_{i=1}^{r} p_{i,j} \text{ for } j = 1, \ldots, c$$

For a vector $(y_1, \ldots, y_c)$, we denote the sum of its components by $y_\bullet := \sum_{i=1}^{c} y_i$, and set $\mathbb{R}_+ := (0, \infty)$.

For a vector $\mathbf{a} = (a_1, \dots, a_d) \in \mathbb{R}_+^d$, we denote by

- $D_{\mathbf{a}}$: the *Dirichlet distribution* with parameter vector $\mathbf{a}$
- $G_{\mathbf{a}}$: the distribution $\Gamma_{a_1} \otimes \cdots \otimes \Gamma_{a_d}$, that is $(Z_1, \dots, Z_d) \sim G_{\mathbf{a}}$ iff all $Z_i \sim \Gamma_{a_i}$, and $Z_1, \dots, Z_d$ are independent

In what follows, the matrix $A = (\alpha_{i,j})_{i=1}^{r}{}_{j=1}^{c}$ will be the set of parameters for the following distributions:

- D_A: the law of the matrix $X = (X_{i,j})$, with $X^{(i)} := (X_{i,1}, \dots, X_{i,c}) \sim D_{(\alpha_{i,1}, \dots, \alpha_{i,c})}$ and $X^{(1)}, \dots, X^{(r)}$ are independent.
- G_A: the law of the matrix $Z = (Z_{i,j})$, such that $Z^{(i)} \sim G_{(\alpha_{i,1}, \dots, \alpha_{i,c})}$ and $Z^{(1)}, \dots, Z^{(r)}$ are independent.

Chamayou and Letac's first theorem and our extension

The following theorem is the first main result in CL94:

Theorem

If $(\mathbf{Y}, X) \sim D_{(\alpha_{1\bullet},\ldots,\alpha_{r\bullet})} \otimes D_A$, *then* $\mathbf{Y}X \sim D_{(\alpha_{\bullet 1},\ldots,\alpha_{\bullet c})}$.

We extend this theorem as follows:

Theorem

Let $\mathbf{t} = (t_1, \ldots, t_r) \in \mathbb{R}_+^r$ *and* $\mathbf{s} = (s_1, \ldots, s_c) \in \mathbb{R}_+^c$ *with* $t_\bullet = s_\bullet$. *Suppose* X *is an* $r \times c$ *non-negative random matrix independent of both* $\mathbf{Y} \sim D_{\mathbf{t}}$ *and* $\mathbf{V} \sim G_{\mathbf{t}}$. *Then*

$$\mathbf{Y}X \sim D_{\mathbf{s}} \quad \textit{iff} \quad \mathbf{V}X \sim G_{\mathbf{s}}.$$

Two properties of the gamma and Dirichlet distributions

Let $\mathbf{Z} = (Z_1, \dots, Z_d) \sim G_{\mathbf{t}}$ for some $\mathbf{t} \in \mathbb{R}_+^d$. Then

$$\left(\frac{Z_1}{Z_\bullet}, \dots, \frac{Z_d}{Z_\bullet}\right) \sim D_{\mathbf{t}}. \tag{1}$$

The second property is

$$(Z_1, \dots, Z_d) \stackrel{d}{=} \left(\frac{Z_1}{Z_\bullet}, \dots, \frac{Z_d}{Z_\bullet}\right) \widetilde{Z}_\bullet, \tag{2}$$

where $(\widetilde{Z}_1, \dots, \widetilde{Z}_d)$ is an independent copy of $\mathbf{Z}$.

Our theorem is indeed an extension

For $(\mathbf{V}, Z) \sim G_{(\alpha_{1\bullet},\dots,\alpha_{r\bullet})} \otimes G_A$, then property (1) implies that

$$X := \begin{pmatrix} \frac{Z_{1,1}}{Z_{1\bullet}} & \dots & \frac{Z_{1,c}}{Z_{1\bullet}} \\ \vdots & \ddots & \vdots \\ \frac{Z_{r,1}}{Z_{r\bullet}} & \dots & \frac{Z_{r,c}}{Z_{r\bullet}} \end{pmatrix} \sim D_A \text{ is independent of } \mathbf{V}.$$

Now

$$\begin{aligned} \mathbf{V}X &= \sum_{k=1}^{r} \left(\frac{Z_{k,1}}{Z_{k\bullet}}, \dots, \frac{Z_{k,c}}{Z_{k\bullet}} \right) V_k \\ &\stackrel{d}{=} \sum_{k=1}^{r} (Z_{k,1}, \dots, Z_{k,c}) = (Z_{\bullet 1}, \dots, Z_{\bullet c}) \sim G_{(\alpha_{\bullet 1},\dots,\alpha_{\bullet c})}. \end{aligned}$$

It follows from our theorem that for a random vector $\mathbf{Y}$ satisfying $(\mathbf{Y}, X) \sim D_{(\alpha_{1\bullet},\dots,\alpha_{r\bullet})} \otimes D_A$, one has $\mathbf{Y}X \sim D_{(\alpha_{\bullet 1},\dots,\alpha_{\bullet c})}$.

A theorem by Pitman

The proof our extension to the first main theorem in CL94 is based on an extension of the following remarkable observation from Pitman (1937).

Let $\mathbf{Z} = (Z_1, \ldots, Z_d) \sim G_{\mathbf{t}}$, and $f : \mathbb{R}^d \to \mathbb{R}$ be a *scale independent* function, i.e., for any $a \neq 0$,

$$f(ax_1, \ldots, ax_d) \equiv f(x_1, \ldots, x_d).$$

Then $f(\mathbf{Z})$ is independent of $Z_\bullet$.

An extension of Pitman's theorem

Lemma

Let $(\Omega, \mathcal{F}, \mathbb{P})$ *be a probability space,* $(E, \mathcal{E})$ *a measurable space, and* $X : \Omega \to E$ *a random element. Suppose* $H : \mathbb{R}^r \times E \to [0, \infty)$ *is jointly measurable and, for any* $a \neq 0$ *and* $\omega \in \Omega$,

$$H(ay_1, \ldots, ay_r, X(\omega)) = H(y_1, \ldots, y_r, X(\omega))$$

for all $(y_1, \ldots, y_r) \in \mathbb{R}^r$. *If* $\mathbf{V} = (V_1, \ldots, V_r) \sim G_{\mathbf{t}}$, $\mathbf{t} \in \mathbb{R}^r_+$, *is independent of* X, *then* $V_\bullet$ *is independent of* $H(\mathbf{V}, X)$.

To prove this lemma we show that the joint Laplace transform $\phi(s, u) = \mathbb{E} e^{-sV_\bullet - uH(\mathbf{V}, X)}$ can be expressed as the product of two functions, one depending on s, and the other on u.

A proof of our first theorem

For the forward implication, let $\mathbf{V}X \sim G_{\mathbf{s}}$ for $\mathbf{V} \sim G_{\mathbf{t}}$ independent of X.

The previous lemma implies that for the function $H(v_1, \ldots, v_r, X) = (H_1(v_1, \ldots, v_r, X), \ldots, H_c(v_1, \ldots, v_r, X))$ defined by

$$H_j(v_1, \ldots, v_r, X) := \sum_{i=1}^{r} \frac{v_i X_{i,j}}{v_\bullet}, \quad 1 \leq j \leq c,$$

the random vector $H(V_1, \ldots, V_r, X) \equiv \left(\frac{V_1}{V_\bullet}, \ldots, \frac{V_r}{V_\bullet}\right) X$ is independent of $V_\bullet$.

Therefore

$$\mathbf{V}X = \left(\frac{V_1}{V_\bullet}, \dots, \frac{V_r}{V_\bullet}\right) X V_\bullet \stackrel{d}{=} \left(\frac{V_1}{V_\bullet}, \dots, \frac{V_r}{V_\bullet}\right) X \widetilde{V}_\bullet,$$

where $(\widetilde{V}_1, \dots, \widetilde{V}_r) \sim G_{\mathbf{t}}$ is independent of $(\mathbf{V}, X)$.

Since $\mathbf{V}X \sim G_{\mathbf{s}}$, for $\mathbf{Z} := (Z_1, \dots, Z_c) \sim G_{\mathbf{s}}$, one has

$$\mathbf{V}X \stackrel{d}{=} \mathbf{Z} \stackrel{d}{=} \left(\frac{Z_1}{Z_\bullet}, \dots, \frac{Z_c}{Z_\bullet}\right) \widetilde{Z}_\bullet,$$

$(\widetilde{Z}_1, \dots, \widetilde{Z}_c)$ being an independent copy of $\mathbf{Z}$.

Taking logarithms on the components of the vectors above

$$\left(\ln\left(\frac{\sum_{i=1}^{r} V_i X_{i,1}}{V_\bullet}\right), \ldots, \ln\left(\frac{\sum_{i=1}^{r} V_i X_{i,c}}{V_\bullet}\right)\right) + \ln(\widetilde{V}_\bullet)(1, \ldots, 1)$$
$$\stackrel{d}{=} \left(\ln\left(\frac{Z_1}{Z_\bullet}\right), \ldots, \ln\left(\frac{Z_c}{Z_\bullet}\right)\right) + \ln(\widetilde{Z}_\bullet)(1, \ldots, 1),$$

Since $t_\bullet = s_\bullet$, one has $\widetilde{V}_\bullet \stackrel{d}{=} \widetilde{Z}_\bullet$, and so letting ψ, φ and χ denote the characteristic functions of the first, second (and fourth), and third terms above, respectively, we have

$$\psi(u_1, \ldots, u_c)\varphi(u_1, \ldots, u_c) = \chi(u_1, \ldots, u_c)\varphi(u_1, \ldots, u_c).$$

We conclude that $\psi \equiv \chi$ (since $\varphi(u_1, \ldots, u_c) \neq 0$), and therefore

$$\left(\frac{V_1}{V_\bullet}, \ldots, \frac{V_r}{V_\bullet}\right) X = \left(\frac{\sum_{i=1}^r V_i X_{i,1}}{V_\bullet}, \ldots, \frac{\sum_{i=1}^r V_i X_{i,c}}{V_\bullet}\right)$$
$$\stackrel{d}{=} \left(\frac{Z_1}{Z_\bullet}, \ldots, \frac{Z_c}{Z_\bullet}\right) \sim D_{\mathbf{s}}$$

Since the left hand side above has the form $\mathbf{Y}X$ for $\mathbf{Y} \sim D_{\mathbf{t}}$ independent of X, we have $\mathbf{Y}X \sim D_{\mathbf{s}}$ as required.

One can obtain the backward implication by reversing these steps.

Chamayou and Letac's transient random walk

The following theorem is the second main result in CL94.

Theorem

If $r = c = d$, $X \sim D_A$, and

$$(\alpha_{1\bullet}, \ldots, \alpha_{d\bullet}) = (\alpha_{\bullet 1}, \ldots, \alpha_{\bullet d}),$$

then $\mathcal{L}(X) \in \mathscr{K}_d$, and $\widehat{X}^{(1)} \sim D_{(\alpha_{1\bullet}, \ldots, \alpha_{d\bullet})}$. Furthermore, if $\mathbf{Y}$ is a random vector in the d-dimensional simplex that is independent of X, then $\mathbf{Y}X \stackrel{d}{=} \mathbf{Y}$ iff $\mathbf{Y} \stackrel{d}{=} \widehat{X}^{(1)}$.

An characterisation theorem for $\mathcal{L}(X) \in \mathscr{K}_d$

The following theorem is an extension of the second main theorem in CL94.

Theorem

(i) *$\mathcal{L}(X) \in \mathscr{K}_d$ iff*

[C1] *there exists a* $\mathbf{t} \in \mathbb{R}_+^d$ *such that, for a random vector* $\mathbf{V} \sim G_{\mathbf{t}}$ *independent of* X*, one has* $\mathbf{V}X \stackrel{d}{=} \mathbf{V}$*; and*

[C2] *for an i.i.d. sequence* $\{X(n)\}_{n\geq 1}$ *with* $X(1) \stackrel{d}{=} X$*,* $\exists m < \infty$ *such that* $\mathbb{P}(X(m,1)$ *is positive*$) > 0$.

(ii) *If* $\mathcal{L}(X) \in \mathscr{K}_d$*, then* $\widehat{X}^{(1)} \sim D_{\mathbf{t}}$*, where the vector* $\mathbf{t}$ *is the same as in* [**C1**]*, and if* $\mathbf{Y}$ *is a random vector in the d-dimensional simplex that is independent of* X*, then* $\mathbf{Y}X \stackrel{d}{=} \mathbf{Y}$ *iff* $\mathbf{Y} \stackrel{d}{=} \widehat{X}^{(1)}$.

Random exchange models

Suppose we have $d < \infty$ bins holding amounts $q_k(n)$, $k = 1, \ldots, d$, of a homogeneous commodity at times $n = 0, 1, 2, \ldots$, respectively.

The dynamics of the model are as follows: at time $n \geq 1$, the vector

$$\mathbf{q}(n-1) := (q_1(n-1), \ldots, q_d(n-1))$$

changes to

$$\mathbf{q}(n) := \mathbf{q}(n-1)X(n).$$

Then

$$\mathbf{q}(n) = \mathbf{q}(0)X(1,n), \quad n \geq 1,$$

is a Markov chain, with stationary distribution $\mathcal{L}(\widehat{X}^{(1)})$ (where we assume w.l.o.g. that $\sum_{k=1}^{d} \mathbf{q}_k(0) = 1$).

Example 1

We now extend the definition of $D_{\mathbf{a}}$ to include vectors containing zeros as follows. The components of $\mathbf{Y} \sim D_{\mathbf{a}}$ that correspond to zero components of $\mathbf{a}$ are identically zero, whereas the subvector of $\mathbf{Y}$ consisting of the components Y_j of that random vector that correspond to $a_j > 0$ form a usual Dirichlet distributed vector. We define $G_{\mathbf{a}}$, D_A and G_A in a similar way.

Let $A = (\alpha_{i,j})_{i=1}^{d}{}_{j=1}^{d}$ be non-negative (i.e. one can have $\alpha_{i,j} = 0$) with $\alpha_{i\bullet} = \alpha_{\bullet i} > 0$, for $i = 1, \ldots, d$. If $X \sim D_A$ and X satisfies [**C2**], then $\mathcal{L}(X) \in \mathscr{K}_d$ with $\widehat{X}^{(1)} \sim D_{(\alpha_{1\bullet},\ldots,\alpha_{d\bullet})}$. Therefore $D_{(\alpha_{1\bullet},\ldots,\alpha_{d\bullet})}$ is the stationary distribution of Markov chain $\{\mathbf{q}(n)\}_{n\geq 1}$.

In particular, we have obtained the stationary distribution for the following simple model: at time $n \geq 1$, a uniform proportion of the commodity previously held in bin k, $k = 1, 2, \ldots, d$, is shifted to the (neighbouring) bin $k + 1 \pmod{d}$.

Example 1

In this case vector $\mathbf{q}(n)$ is defined as above with

$$X := \begin{pmatrix} U_1 & 1-U_1 & 0 & \cdots & 0 \\ 0 & U_2 & 1-U_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & U_{d-1} & 1-U_{d-1} \\ 1-U_d & 0 & \cdots & 0 & U_d \end{pmatrix},$$

where the U_k, $k = 1, \ldots, d$, are i.i.d. uniformly distributed random variables on $(0, 1)$.

Observing that X defined above satisfies [**C2**] for $m = d - 1$, we conclude that $\widehat{X}^{(1)} \sim D_{(2,\ldots,2)}$, and so $D_{(2,\ldots,2)}$ is the stationary distribution of Markov chain $\{\mathbf{q}(n)\}_{n\geq 1}$.

Example 2

In this example, we consider a random stochastic matrix X with all rows dependent. The behaviour of this model is controlled by the decisions of a "leader" as follows:

At time $n \geq 1$, the "leader" shifts a uniform proportion of the commodity held in bin 1 to bin 2. If the proportion shifted is greater than $1/2$, then no other shifts occur in the system at time n. However, if the proportion shifted is less than or equal to $1/2$, then the commodity previously held in bin k, $k = 2, 3, \ldots, d$, $d \geq 2$, is shifted to the (neighbouring) bin $k + 1 \pmod{d}$.

Example 2

In this case, the vector $\mathbf{q}(n)$ is defined as above with

$$X := \begin{pmatrix} U & 1-U & 0 & \cdots & 0 \\ 0 & I & 1-I & \ddots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 \\ 0 & \cdots & 0 & I & 1-I \\ 1-I & 0 & \cdots & 0 & I \end{pmatrix},$$

where $U \sim U(0,1)$, and $I := \mathbf{1}_{\{U>1/2\}}$, $\mathbf{1}_A$ being the indicator function for event A.

We can show that $\mathcal{L}(X) \in \mathscr{K}_d$ with $\widehat{X}^{(1)} \sim D_{(2,\ldots,2)}$ using our extension of the second main theorem in CL94.

It is not hard to directly verify that X satisfies [**C2**] for $m = 2d-2$. To show [**C1**] holds, we let $\mathbf{V} \sim G_{(2,\ldots,2)}$ be independent of X, and show that the characteristic function of $\mathbf{V}X$ is the same as that of $\mathbf{V}$.

Other applications

The second main theorem in CL94 has been used to compute the limiting distribution of random nested tetrahedra in Letac and Scarsini (1998), and to compute the stationary distribution of a donkey walk in the plane in Letac (2002).

Our extension of the second main theorem in CL94 provides natural extensions of these results.

References

- Chamayou, J.-F. and Letac, G. (1994). A transient random walk on stochastic matrices with Dirichlet distributions. *Ann. Prob.* **22**, 424–430.
- Pitman, E. J. G. (1937). The "closest" estimates of statistical parameters. *Proc. Camb. Phil. Soc.* **33**, 212–222.
- Letac, G. (2002). Donkey walk and Dirichlet distributions. *Statist. Prob. Lett.* **57**, 17–22.
- Letac, G. and Scarsini, M. (1998). Random nested tetrahedra. *Adv. Appl. Prob.* **30**, 619–627.
- Van Assche, W. (1986). Products of 2×2 stochastic matrices with random entries. *J. Appl. Prob.* **23**, 1019–1024.
- Volodin, N. A., Kotz, S. and Johnson. N. L. (1993). Use of moments in distribution theory: a multivariate case. *J. Multivariate Anal.* **46**, 112–119.