Convergence of some time inhomogeneous Markov chains via spectral techniques

Jessica Zúñiga
with Laurent Saloff-Coste

June 2008

Motivation

- Mironov (2002)
  (Not so) random shuffles of RC4.
  - *total variation upper bound of* $O(n \log n)$ *for cyclic-to-random shuffle*
  - *lower bound of* $O(n)$ *for cyclic-to-random shuffle*

- *Mossel, Peres, Sinclair (2004)*
  Shuffling by semi-random transpositions.
  - total variation upper bound of $O(n \log n)$ for any semi-random transposition shuffle
  - lower bound of $O(n \log n)$ for cyclic-to-random shuffle

- M. Ganapathy developed independently and by similar arguments an improved upper bound for semi-random transpositions. (2006)

Basic notation

V a finite set equipped with $(K_n)_1^\infty$ s.t.

$$K_n(x,y) \geq 0 \text{ and } \sum_y K_n(x,y) = 1.$$

The associated Markov chain: $X = (X_n)_0^\infty$

$$\begin{aligned} &P(X_n = y | X_{n-1} = x, X_{n-2} = x_{n-2}, \ldots, X_0 = x_0) \\ &= P(X_n = y | X_{n-1} = x) \\ &= K_n(x,y) \end{aligned}$$

Let μ_0 be the distribution of X_0. The distribution μ_n of X_n is

$$\mu_n(x) = \sum_{x \in V} \mu_0(x) K_{0,n}(x,y)$$

where $K_{n,n} = I$ and for $m \geq n$

$$K_{n,m}(x,y) = \sum_{z \in V} K_{n,m-1}(x,z) K_m(z,y).$$

K_n are matrices: $K_{n,m} = K_{n+1} \ldots K_m$.

Major assumption: There exists a measure π that is invariant for the sequence $(K_n)_1^\infty$, i.e.

$$\pi K_n = \pi.$$

Large class of examples: random walks on groups.

Let $V = G$ a finite group equipped with a probability measure p. The Markov kernel

$$K(x, y) = p(x^{-1}y)$$

has $\pi = 1/|G|$ as an invariant measure.

Any sequence of probability measures $(p_i)_1^\infty$ on G gives a sequence $(K_i)_1^\infty$ of Markov kernels with invariant measure π.

$$K_{n,m}(x, y) = p_{n+1} * \cdots * p_m(x^{-1}y)$$

where

$$u * v(x) = \sum_{y \in G} u(y)v(y^{-1}x)$$

Distances:

$$\begin{aligned} d_{TV}(\mu,\pi) &= \sup_{A\in V} |\mu(A)-\pi(A)| \\ d_2(\mu,\pi)^2 &= \sum_y \left|\frac{\mu(y)}{\pi(y)}-1\right|^2 \pi(y) \end{aligned}$$

Jensen's inequality: $2d_{TV}(\mu,\pi) \le d_2(\mu,\pi)$

(K,π) is a linear operator on $\ell^2(V,\pi)$.

Let $f,g \in \ell^2(V,\pi)$.

- $Kf = \sum_{y\in V} K(\cdot,y)f(y)$
- The adjoint K^* is determined by $\langle Kf,g\rangle = \langle f,K^*g\rangle$
- K^* has associated Markov kernel

$$K^*(x,y) = \pi(y)K(y,x)/\pi(x).$$

Spectral analysis

Eigenvalues of K:

$$1 = \beta_0(K) \geq \beta_1(K) \geq \cdots \geq \beta_{|V|-1}(K) \geq -1.$$

Singular values of K on $\ell^2(V, \pi)$:

$$1 = \sigma_0(K) \geq \sigma_1(K) \geq \cdots \geq \sigma_{|V|-1}(K) \geq 0$$

where

$$\sigma_i(K) = \sqrt{\beta_i(KK^*)} = \sqrt{\beta_i(K^*K)}.$$

Theorem: Consider (K, π). Let $(\psi_i)_0^{|V|-1}$ be a basis of $\ell^2(V, \pi)$ such that $KK^*\psi_i = \sigma_i(K)^2\psi_i$ (assume $\psi_0 = 1$.) Then

$$d_2(K(x, \cdot), \pi)^2 = \sum_{i=1}^{|V|-1} \sigma_i(K)^2 |\psi_i(x)|^2$$

and

$$\sum_{x \in V} d_2(K(x, \cdot), \pi)^2 \pi(x) = \sum_{i=1}^{|V|-1} \sigma_i(K)^2.$$

Main technical result

Theorem: Let $(K_i)_1^\infty$ be a sequence in V with positive invariant measure π. For $j \geq 1$ and $0 \leq i \leq |V| - 1$ let $\sigma_i(K_j)$ be the singular values of K_j on $\ell^2(V, \pi)$ then

$$d_2(K_{0,n}(x, \cdot), \pi) \leq (\pi(x)^{-1} - 1)^{1/2} \prod_1^n \sigma_1(K_j)$$

and

$$\sum_{x \in V} d_2(K_{0,n}(x, \cdot), \pi)^2 \pi(x) \leq \sum_{i=1}^{|V|-1} \prod_{j=1}^n \sigma_i(K_j)^2.$$

Main idea of proof:
For all $k = 1, \ldots, |V| - 1$

$$\sum_{j=1}^k \sigma_j(K_1 \cdots K_n)^2 \leq \sum_{j=1}^k \prod_{i=1}^n \sigma_j(K_i)^2$$

The inequality follows from chapter 3 in *Topics in matrix analysis* by Horn and Johnson.

Ergodicity

Definition: Let $\mathcal{Q} = \{Q_1, \ldots . Q_n\}$ have invariant measure π. $(\mathcal{Q}, \pi)$ is ergodic if for any $(K_i)_1^\infty$ with invariant measure π and $K_i \in \mathcal{Q}$ for infinitely many i then

$$\lim_{n \to \infty} K_{0,n}(x, z) - K_{0,n}(y, z) = 0$$

for all $x, y, z \in V$.

Remark:

- Let $\mathcal{Q} = \{Q\}$. $(\mathcal{Q}, \pi)$ is ergodic is stronger than

$$\forall x, y, z \in V \ , \lim_{n \to \infty} Q^{(n)}(x, z) - Q^{(n)}(y, z) = 0.$$

Ergodicity

Remark:

- Let $\mathcal{Q} = \{Q\}$ have invariant measure π.
  (1) $\sigma_1(Q) < 1$ implies $Q^n(x, \cdot) \to \pi$ for all x

  (2) (Q, π) irreducible, aperiodic and reversible w.r.t. π implies $\sigma_1(Q) < 1$

  Counter example: Rudvalis shuffle

Weakness:

- K_i transpose cards $i, i+1$ or nothing with probability $1/2$.

- $\sigma_1(K_1 \cdots K_n) < 1$ but $\sigma_1(K_i) = 1$

Ergodicity

Theorem: Let $\mathcal{Q} = \{Q_1, \ldots, Q_k\}$ be a family of kernels on V with invariant measure π.

- $(\mathcal{Q}, \pi)$ is ergodic iff $\sigma_1(Q_j) < 1$ for all $1 \leq j \leq k$.

- If $(\mathcal{Q}, \pi)$ is ergodic then for any $(K_i)_1^\infty$ with invarient measure π and infinitely many $K_i \in \mathcal{Q}$ we have

$$\forall x, \quad \lim_{n \to \infty} K_{0,n}(x, \cdot) - \pi = 0$$

Time inhomogeneous random walks on finite groups

Main technical result simplifies when $V = G$, in this case $\pi = 1/|G|$.

- $(p_i)_1^\infty$ a sequence of probability measures on G.
- $K_i(x, y) = p_i(x^{-1}y)$
- $K_{n,m}(x, y) = p_{n+1} * \cdots * p_m(x^{-1}y)$
- Set $K_{n,m}(x, y) = p_{n,m}(x^{-1}y)$
- For $0 \leq j \leq |G| - 1$ set $\sigma_j(p_i) = \sigma_j(K_i)$ where singular values are taken w.r.t π.

Time inhomogeneous random walks on finite groups

Theorem: Let $(p_i)_1^\infty$ be a sequence on G. Then

$$d_2(p_{0,n}, \pi) \leq (|G| - 1)^{1/2} \prod_1^n \sigma_1(p_j)$$

$$d_2(p_{0,n}, \pi) \leq \left(\sum_{i=1}^{|G|-1} \prod_{j=1}^n \sigma_i(p_j)^2 \right)^{1/2}.$$

Example: $(S_j)_1^\infty$ a sequence of generating sets of G s.t. $\mathrm{id} \in S_j$. Set

$$K_j(x, y) = \begin{cases} 1/|S_j| & \text{if } x^{-1}y \in S_j \\ 0 & \text{otherwise.} \end{cases}$$

Well know eigenvalue estimates give us

$$\sigma_1(K_j)^2 \leq 1 - \frac{1}{|S_j|^2 d_j^2}$$

where d_j is the diameter of the Caley graph associated to $S_j^\# = S_j \cup S_j^{-1}$.

$$d_2(K_{0,n}(x, \cdot), \pi) \leq |G|^{1/2} \prod_1^n \left(1 - \frac{1}{|S_j|^2 d_j^2} \right)^{1/2}$$

Semi-random transpositions

Let $G = S_n$ and $\pi = 1/n!$. For $1 \leq i \leq n$ set

$$q_i(x) = \begin{cases} 1/n & \text{if } x = (i,j) \text{ for } 1 \leq j \leq n \\ 0 & \text{otherwise.} \end{cases}$$

Theorem: For $n > 1, c > 0$ and $k \geq n(\log n + c)$

$$d_2(q_1^{(k)}, \pi) \leq \sqrt{2}e^{-c}.$$

Note: q_i is reversible w.r.t. π. For any i q_1 and q_i are images of each other under some inner automorphism of S_n.

- q_i's have same singular values
- $d_2(q_i^{(k)}, \pi)^2 = d_2(q_1^{(k)}, \pi)^2 = \sum_{m=1}^{n!-1} \sigma_m(q_1)^{2k}$

Semi-random transpositions

Definition: Let $r = (r_i)_1^\infty$ with $r_i \in \{1, \ldots, n\}$. The r-semi-random transposition shuffle is associated to $(p_i)_1^\infty$ where $p_i = q_{r_i}$.

$p^r_{0,k} = p_1 * \cdots * p_k$: distribution after k steps

Theorem: For any $r = (r_i)_1^\infty$, $r_i \in \{1, \ldots, n\}$

$$d_2(p^r_{0,k}, \pi) \leq d_2(q_1^{(k)}, \pi).$$

For $c > 0$, $k \geq n(\log n + c)$, $c > 0$ then $d_2(p^r_{0,k}, \pi) \leq \sqrt{2}e^{-c}$.

Proof: From techniques just described

$$d_2(p^r_{0,k}, \pi)^2 \leq \sum_{m=1}^{n!-1} \prod_{i=1}^{k} \sigma_m(q_{r_i})^2 = \sum_{m=1}^{n!-1} \sigma_m(q_1)^{2k} = d_2(q_1^{(k)}, \pi)^2.$$

Semi-random transpositions

Remark: Information can be lost with the inequality

$$\sum_{j=1}^{k} \sigma_j(K_1 \cdots K_n)^2 \leq \sum_{j=1}^{k} \prod_{i=1}^{n} \sigma_j(K_i)^2.$$

Let $r = (r_i)_1^\infty$ be a sequence of $i.i.d.$ uniform random variables taking values in $\{1, \ldots, n\}$.

$p_{0,k}^r$:random transposition shuffle

Diaconis-Shahshahani:
For $c > 0,\ k \geq (n/2)(\log n + c)$

$$d_2(p_{0,k}^r, \pi) \leq \beta e^{-c}.$$

Semi-random insertions

$c_{i,j}$: pick card in position i and insert it into position j.

$$c_{i,j} = \begin{cases} \text{id} & \text{if } i = j \\ (j, j-1, \ldots, i+1, i) & \text{if } 1 \le i < j \le n \\ (j, j+1, \ldots, i-1, i) & \text{if } 1 \le j < i \le n. \end{cases}$$

Random insertion measure on S_n:

$$\widetilde{q}(x) = \begin{cases} 1/n & \text{if } x = \text{id}, \\ 2/n^2 & \text{if } x = c_{i,j} \text{ if } |i-j| = 1 \\ 1/n^2 & \text{if } x = c_{i,j} \text{ if } |i-j| > 1 \\ 0 & \text{otherwise.} \end{cases}$$

Theorem (Diaconis and Saloff-Coste): For all $n \ge 28$, $c > 2$ and all $k \ge 2n(logn + c)$

$$d_2(\tilde{q}^{(k)}, \pi) \le 2e^{-(c-2)}.$$

Semi-random insertions

For $i \in \{1, \ldots, n\}$ set

$$\widetilde{q}_i(x) = \begin{cases} 1/n & \text{if } x = c_{i,j} \text{ for some } j,\ 1 \le j \le n, \\ 0 & \text{otherwise.} \end{cases}$$

The adjoint $\tilde{q}_i^*(x) = \tilde{q}_i(x^{-1})$ inserts uniformly chosen card into position i.

Note that:

- For any $1 \le i \le n$, $\tilde{q}_i^* * \tilde{q}_i = \tilde{q}$.
- $\sigma_j(\tilde{q}_i) = \sigma_j(\tilde{q})^{1/2}$, $0 \le j \le n! - 1$.

Definition: Let $r = (r_i)_1^\infty$ with $r_i \in \{1, \ldots, n\}$. The r-semi-random insertion shuffle is associated to $(\tilde{p}_i)_1^\infty$ where $\tilde{p}_i = \tilde{q}_{r_i}$.

$\tilde{p}_{0,k}^r = \tilde{p}_1 * \cdots * \tilde{p}_k$: distribution after k steps.

Semi-random insertions

Theorem: For any $r = (r_i)_1^\infty$, $r_i \in \{1, \ldots, n\}$

$$d_2(\tilde{p}^r_{0,2k}, \pi) \leq d_2(\tilde{q}^{(k)}, \pi).$$

For $n > 28$, $c > 2$ and $k \geq 4n(\log n + c)$

$$d_2(\tilde{p}^r_{0,k}, \pi) \leq 2e^{-(c-2)}.$$

Proof:

$$\begin{aligned} d_2(\tilde{p}^r_{0,2k}, \pi)^2 &\leq \sum_{m=1}^{n!-1} \prod_{i=1}^{2k} \sigma_m(\tilde{q}_{r_i})^2 \\ &= \sum_{m=1}^{n!-1} \sigma_m(\tilde{q})^{2k} \\ &= d_2(\tilde{q}^{(k)}, \pi) \end{aligned}$$