Math 639: Lecture 13

Stationary measures, the hypercube, riffle shuffles

Bob Hough

March 23, 2017

Excessive measures

Recall that S is the countable state space of a Markov chain. Let $p(\cdot, \cdot)$ be the chain's homogeneous transition probability.

Definition

We say that a non-zero $\mu : S \to [0, \infty]$ is an *excessive measure* if

$$\mu(y) \geqslant \sum_{x \in S} \mu(x) p(x, y), \qquad \forall y \in S.$$

Invariant measure

Theorem

Let T_z denote the possibly infinite return time to a state z by homogeneous Markov chain $\{X_n\}$. Then

$$\mu_z(y) = \mathrm{E}_z\left[\sum_{n=0}^{T_z-1} \mathbf{1}(X_n = y)\right]$$

is an excessive measure for $\{X_n\}$, the support of which is the closed set of all states accessible from z. If z is a recurrent state then $\mu_z(\cdot)$ is an invariant measure, whose support is closed and recurrent.

Invariant measure

Proof.

- Set $h_k(\omega, y) = \sum_{n=0}^{T_z(\omega)-1} \mathbf{1}(\omega_{n+k} = y)$. Thus $\mu_z(y) = \mathrm{E}_z[h_0(\omega, y)]$.
- Calculate

$$\begin{aligned}
\mathrm{E}_z[h_1(\omega, y)] &= \mathrm{E}_z\left[\sum_{n=0}^{\infty} \mathbf{1}(T_z > n)\mathbf{1}(X_{n+1} = y) \sum_{x \in S} \mathbf{1}(X_n = x)\right] \\
&= \sum_{x \in S} \sum_{n=0}^{\infty} \mathrm{E}_z\left[\mathbf{1}(T_z > n)\mathbf{1}(X_n = x)\,\mathrm{Prob}_z(X_{n+1} = y | \mathscr{F}_n)\right] \\
&= \sum_{x \in S} \sum_{n=0}^{\infty} \mathrm{E}_z\left[\mathbf{1}(T_z > n)\mathbf{1}(X_n = x)\right] p(x, y) \\
&= \sum_{x \in S} \mu_z(x) p(x, y).
\end{aligned}$$

$\square$

Invariant measure

Proof.

- Observe that if $\omega_0 = z$, $h_0(\omega, y) \geqslant h_1(\omega, y)$ with equality when $y \neq z$ or $T_z(\omega) < \infty$.
- It follows that

$$\mu_z(y) = \mathrm{E}_z[h_0(\omega, y)] \geqslant \mathrm{E}_z[h_1(\omega, y)] = \sum_{x \in S} \mu_z(x) p(x, y),$$

  with equality when $y \neq z$ or z is recurrent. This proves that μ_z is excessive.
- Iterating, for any $k \geqslant 1$,

$$\mu_z(y) \geqslant \sum_{x \in S} \mu_z(x) \, \mathrm{Prob}_x(X_k = y)$$

  with equality if z is recurrent.

□

Invariant measure

Proof.

- If $\rho_{zy} = 0$ then $\mu_z(y) = 0$, while if $\rho_{zy} > 0$ then $\mathrm{Prob}_z(X_k = y) > 0$ for some finite k, so that $\mu_z(y) \geqslant \mu_z(z)\,\mathrm{Prob}_z(X_k = y)$. Thus when z is recurrent, the support of μ_z is its irreducible component.
- If $x \leftrightarrow z$ then $1 = \mu_z(z) \geqslant \mu_z(x)\,\mathrm{Prob}_x(X_k = z)$ for some k, whence μ_z is invariant and σ-finite.

$\square$

Invariant measure

Theorem

If R is a recurrent equivalence class of states then the invariant measure whose support is contained in R is unique and has R as its support. In particular, the invariant measure of an irreducible, recurrent chain is unique.

Invariant measure

Proof.

- Since R is closed, the restriction of $p(\cdot,\cdot)$ to R is a transition probability, so we may assume $S = R$.
- Hence there exists a strictly positive invariant measure $\mu = \mu_z$ on R
- Define transition probability $q(x,y) = \frac{\mu(y)p(y,x)}{\mu(x)}$.
- Let ν be any excessive probability for $p(\cdot,\cdot)$. Then for any y,

$$\nu(y) \geqslant \sum_{x \in S} \nu(x)p(x,y) = \sum_{x \in S} \nu(x)q(y,x)\frac{\mu(y)}{\mu(x)}$$

so that $\frac{\nu}{\mu}$ is a superharmonic function for q.

$\square$

Invariant measure

Proof.

- By considering paths, we can check that $\rho_{x,y} > 0$ for p implies $\rho_{y,x} > 0$ for q, and hence the Markov chain with transition probability q is irreducible.
- Considering loops, the probability of a return from x to x at step k under p is equal to the same probability under q (by running each loop in reverse). Hence the q-chain is recurrent.
- Check as an exercise that the only positive super-harmonic function for an irreducible recurrent chain is a constant, and hence ν is a scalar multiple of μ.

$\square$

Reversible chains

Definition

A non-zero $\mu : S \to [0, \infty)$ is called a *reversible measure* for the transition probability $p(\cdot, \cdot)$ if for all $x, y \in S$, $\mu(x)p(x, y) = \mu(y)p(y, x)$. The transition probability $p(\cdot, \cdot)$ is *reversible* if it has a reversible measure.

Time-reversed chain

Definition

If $\mu(\cdot)$ is an invariant measure for transition probability $p(x, y)$, then $q(x, y) = \mu(y)p(y, x)/\mu(x)$ is a transition probability on the support of $\mu(\cdot)$, called the *adjoint* or *dual* of p with respect to μ. The corresponding Markov chain is called the *time-reversed* chain.

Random walk on a graph

Definition

- A *network* consists of a countable (finite or infinite) set of vertices V with a symmetric weight function $w : V \times V \mapsto [0, \infty)$ (i.e. $w_{xy} = w_{yx}$ for all $x, y \in V$). Set $\mu(x) = \sum_{y \in V} w_{xy}$.
- A *random walk* on the network is a homogeneous Markov chain of state space V and transition probability

$$p(x, y) = \frac{w_{xy}}{\mu(x)}.$$

Recurrent states

Definition

Let T_z denote the first return time to state z. A recurrent state z is called *positive recurrent* if $\mathrm{E}_z[T_z] < \infty$ and *null recurrent* otherwise.

Recurrence

Theorem

If $\pi(\cdot)$ is an invariant probability measure, then all states z with $\pi(z) > 0$ are positive recurrent. Further, if the support of $\pi(\cdot)$ is an irreducible set R of positive recurrent states then $\pi(z) = 1/\,\mathrm{E}_z[T_z]$ for all $z \in R$.

Recurrence

Proof.

- Starting the chain from the invariant distribution π one easily verifies that π is supported on recurrent states.
- Calculate, starting from a recurrent state z,

$$\mu_z(S) = \sum_{y \in S} \mu_z(y) = \mathrm{E}_x \left[\sum_{y \in S} \sum_{n=0}^{T_z - 1} \mathbf{1}(X_n = y) \right] = \mathrm{E}_z[T_z].$$

  Thus, if μ_z is a finite measure then z is positive recurrent.
- If π is supported on a single irreducible then $\pi(z) = \frac{\mu_z(z)}{\mu_z(S)} = \frac{1}{\mathrm{E}[T_z]}$.
- To complete the proof, note that an invariant probability measure is a mixture of invariant probability measures supported on single irreducibles.

$\square$

Markovian coupling

Theorem

Let $\{X_n\}$ and $\{Y_n\}$ be two independent copies of an aperiodic, irreducible Markov chain. Suppose further that the irreducible chain $Z_n = (X_n, Y_n)$ is recurrent. Then, regardless of the initial distribution (X_0, Y_0), the first meeting time $\tau = \min\{\ell \geqslant 0 : X_\ell = Y_\ell\}$ of the two processes is a.s. finite, and for any n,

$$\|\mathscr{L}_{X_n} - \mathscr{L}_{Y_n}\|_{\mathrm{TV}} \leqslant 2\,\mathrm{Prob}(\tau > n).$$

Markovian coupling

Proof.

- The Markov chain $Z_n = (X_n, Y_n)$ on S^2 is irreducible by independence. Since $\{Z_n\}$ is recurrent, $\tau_z = \min\{\ell \geqslant 0 : Z_\ell = z\}$ is a.s. finite for each $z \in S^2$. Thus, so is

$$\tau = \inf\{\tau_z : z = (x, x), \text{ some } x \in S\}.$$

- For the remaining claim, let $g \in b\mathscr{S}$ bounded by 1, and verify that, for $k \leqslant n$,

$$\mathbf{1}(\tau = k) \, \mathrm{E}_{X_k}[g(X_{n-k})] = \mathbf{1}(\tau = k) \, \mathrm{E}_{Y_k}[g(Y_{n-k})]$$

or $\mathrm{E}[\mathbf{1}(\tau = k) g(X_n)] = \mathrm{E}[\mathbf{1}(\tau = k) g(Y_n)]$. Thus

$$\mathrm{E}[g(X_n)] - \mathrm{E}[g(Y_n)] = \mathrm{E}[\mathbf{1}(\tau > n)(g(X_n) - g(Y_n))] \leqslant 2 \, \mathrm{Prob}(\tau > n).$$

□

Random walk on the hypercube

This model is closely related to the Ehrenfest urn model of diffusion. Let $G = (\mathbb{Z}/2\mathbb{Z})^n$ and let Q be the measure

$$Q(x) = \begin{cases} \frac{1}{n+1} & \|x\|_1 \leqslant 1 \\ 0 & \|x\|_1 \geqslant 2 \end{cases}.$$

Let Q^{*0} be a point mass at $x = 0$, and $Q^{*N} = Q * Q^{N-1}$ for $N \geqslant 1$. This defines the distribution of a random walk on G.

Random walk on the hypercube

- Random walk on the hypercube is a Markov chain with stationary distribution U which is uniform on G.
- Recall that we define the total variation distance by

$$\left\|Q^{*N} - U\right\|_{\mathrm{TV}} = \sup_{A \subset G} \left|Q^{*N}(A) - U(A)\right| = \frac{1}{2}\left\|Q^{*N} - U\right\|_1.$$

Random walk on the hypercube

Theorem

Let $c \in \mathbb{R}$ and let N be the integer nearest $\frac{n+1}{4}(\log n + c)$. There is a function $f(x)$ which satisfies $f(x) \to 0$ as $|x| \to \infty$ such that for fixed c, as $n \to \infty$,

$$\left| \left\| Q^{*N} - U \right\|_{\mathrm{TV}} - \delta_{(c<0)} \right| \leqslant f(c) + o(1).$$

Informally, random walk on the hypercube has mixing time $\frac{1}{4} n \log n$ with a cut-off window of order n.

Random walk on the hypercube

- The characters of G are given as follows. For $x, y \in (\mathbb{Z}/2\mathbb{Z})^n$, $\chi_y(x) = (-1)^{x \cdot y}$.
- Given $y \in (\mathbb{Z}/2\mathbb{Z})^n$, the Fourier transform of Q at y is

$$\hat{Q}(y) = \sum_x (-1)^{x \cdot y} Q(x) = 1 - \frac{2\|y\|_1}{n+1}.$$

- The uniform measure has $\hat{U}(y) = \delta_{(y=0)}$.

Random walk on the hypercube

- The Fourier transform has the usual property of converting convolution to pointwise multiplication, so

$$\widehat{P * Q}(y) = \hat{P}(y)\hat{Q}(y).$$

  In other words, the characters are eigenvectors of the transition kernel with the Fourier coefficients as eigen-values.

- The Parseval relation is

$$\sum_{x \in (\mathbb{Z}/2\mathbb{Z})^n} |f(x)|^2 = \frac{1}{2^n} \sum_{y \in (\mathbb{Z}/2\mathbb{Z})^n} \left|\hat{f}(y)\right|^2$$

Random walk on the hypercube

- Let $c > 0$. Cauchy-Schwarz gives

$$
\begin{aligned}
\left\| Q^{*N} - U \right\|_{\mathrm{TV}}^2 &\leqslant \frac{2^n}{4} \left\| Q^{*N} - U \right\|_2^2 \\
&= \frac{1}{4} \sum_{0 \neq y \in (\mathbb{Z}/2\mathbb{Z})^2} \left| \hat{Q}(y) \right|^{2N} \\
&= \frac{1}{4} \sum_{j=1}^{n} \binom{n}{j} \left(1 - \frac{2j}{n+1}\right)^{2N} \\
&\leqslant \frac{1}{2} \sum_{1 \leqslant j \leqslant n/2} \frac{n^j}{j!} e^{-\frac{4jN}{n+1}} + O(n^{-1/2}) \\
&\lesssim \frac{1}{2} \sum_{1 \leqslant j} \frac{e^{-jc}}{j!} = \frac{e^{e^{-c}} - 1}{2}.
\end{aligned}
$$

This gives the upper bound.

Random walk on the hypercube

- To prove the lower bound, let $f(x) = \sum_{\|y\|_1=1} \chi_y(x)$.
- Calculate

$$\mathrm{E}_U[f] = 0, \qquad \mathrm{E}_U\left[f^2\right] = \mathrm{E}_U\left[n + 2\sum_{\|y\|_1=2} \chi_y\right] = n$$

$$\mathrm{E}_{Q*N}[f] = n\left(1 - \frac{2}{n+1}\right)^N,$$

$$\mathrm{E}_{Q*N}[f^2] = n + (n^2 - n)\left(1 - \frac{4}{n+1}\right)^N$$

$$\begin{aligned}\mathrm{Var}_{Q*N}[f] &= n + (n^2 - n)\left(1 - \frac{4}{n+1}\right)^N - n^2\left(1 - \frac{2}{n+1}\right)^{2N} \\ &\leqslant n\left(1 - \left(1 - \frac{2}{n+1}\right)^{2N}\right) < n.\end{aligned}$$

Random walk on the hypercube

- Assume $c < 0$ and observe

$$\mathrm{E}_{Q^{*N}}[f] = n\left(1 - \frac{2}{n+1}\right)^N = (1 + o(1))n^{\frac{1}{2}}e^{-\frac{c}{2}}.$$

- Define $A = \left\{x \in (\mathbb{Z}/2\mathbb{Z})^n : f(x) > \frac{1}{2}\mathrm{E}_{Q^{*N}}[f]\right\}$. By Chebyshev's inequality,

$$U(A) \leqslant (4 + o(1))e^c, \qquad Q^{*N}(A^c) \leqslant (4 + o(1))e^c,$$

so $\left\|Q^{*N} - U\right\|_{\mathrm{TV}} \geqslant 1 - (4 + o(1))e^c$.

Riffle shuffling

The following discussion is based on the paper "Trailing the dovetail shuffle to its lair" by Bayer and Diaconis, *Annals of Applied Probability*, vol 2., No. 2, (1992) pp. 294–313.

Riffle shuffling

The Gilbert-Shannon-Reeds model of riffle shuffling proceeds as follows.

- A deck of n cards is cut into two portions according to the binomial distribution, so that the probability that the top k cards are cut off is $\binom{n}{k}/2^n$.
- The two packets of cards are then riffled together so that one by one the cards drop from left or right packet with probability proportional to the size of the packet.

The G-S-R model defines a Markov chain on the symmetric group S_n with stationary measure the uniform measure on S_n.

Riffle shuffling

Definition

Let $\pi \in S_n$. A *rising sequence* in π is a maximal set $\{i, i+1, i+2, ..., k\}$ such that $\pi(i) < \pi(i+1) < \cdots < \pi(k)$. Define $R(\pi)$ the number of rising sequences in π.

Theorem

If n cards are shuffled m times according to the G-S-R model, then the chance that the deck is in arrangement π is $\binom{2^m+n-R(\pi)}{n}/2^{mn}$.

Riffle shuffling

Let Q^m denote the distribution of m shuffles of n cards from the G-S-R model, and let $U(\pi) = \frac{1}{n!}$ denote the uniform measure on S_n. Recall that the total variation distance is defined by

$$\|Q^m - U\|_{\mathrm{TV}} = \max_{A \subset S_n} |Q^m(A) - U(A)|.$$

Theorem

If n cards are shuffled m times with $m = \frac{3}{2}\log_2 n + \theta$, then for large n,

$$\|Q^m - U\|_{\mathrm{TV}} = 1 - 2\Phi\left(\frac{-2^{-\theta}}{4\sqrt{3}}\right) + O\left(\frac{1}{n^{\frac{1}{2}}}\right),$$

with

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}}\, dt.$$

a-shuffles

Let $a \geqslant 2$.

- In an a-shuffle the deck is cut into a piles according to the multinomial distribution.
- Then cards are dropped one at a time from the piles, with probability that a given pile is chosen proportional to the size of the pile.

Thus the G-S-R model of riffle shuffling is a 2-shuffle.

Shuffling models

An a-shuffle admits several descriptions as follows.

- (Geometric description) Place n points uniformly and independently in the interval $[0, 1]$, at locations $x_1 < x_2 < \cdots < x_n$. The map $x \mapsto ax \bmod 1$ rearranges the order of the points x_i, giving a measure on S_n
- (Maximum entropy description) All possible ways of cutting a deck into a packets, then interleaving the packets are equally likely. Empty packets are allowed.
- (Inverse description) Deal the deck into a piles by turning cards up one at a time and choosing a pile uniformly at random. Then place the piles one on top of the other.

Shuffling models

Lemma

Each of the above three descriptions is equivalent to an a-shuffle. Moreover, in each model an a-shuffle followed by a b shuffle is equivalent to an ab-shuffle.

Shuffling models

Proof.

Each of the descriptions results in a multinomial distribution on packet sizes.

- In the geometric description, the packets are those points in interval $[(i-1)/a, i/a]$, which is multinomial distributed.
- In the maximum entropy description, the number of ways of interleaving a packets is a multinomial coefficient.
- This is immediate in the inverse-shuffle.

$\square$

Shuffling models

Proof.

Conditional on the packet sizes, the maximum entropy description makes all possible interleavings equally likely. We now check this in the other descriptions.

- In the initial a-shuffle description, given packets of size $j_1, ..., j_a$, the probability of any specific ordering is

$$\frac{j_1! \cdots j_a!}{(j_1 + \cdots + j_a)!} = \binom{j_1 + \cdots + j_a}{j_1, ..., j_a}^{-1}.$$

- For the geometric description, use that the conditional distribution of the j_i points in the ith interval is the same as dropping j_i points uniformly in the interval, then overlap the intervals one on top of the other.
- In the inverse model this is immediate.

□

Shuffling models

Proof.

- To prove the product rule of shuffles (an a-shuffle followed by a b-shuffle is an ab-shuffle) it suffices to prove this in the geometric model, since the various models induce the same transition probability on S_n.
- In the geometric model, note that the points $ax \bmod 1$ are still independent and uniform in the interval $[0, 1]$, so that successive shuffles can reuse the points.
- The product rule follows from the identity $b(ax \bmod 1) \bmod 1 = abx \bmod 1$.

$\square$

a-Shuffles

Theorem

The probability that an a-shuffle results in a permutation π is $\frac{\binom{a+n-R(\pi)}{n}}{a^n}$.

Combined with the product rule, this gives our formula for the probabilities of repeated G-S-R shuffles.

a-Shuffles

Proof.

- By the maximum entropy description, the probability is $\frac{N}{a^n}$ where N is the number of ways of cutting the deck into a packets so that π is an interleaving.
- Since the r rising sequences of π are the result of merging the a packets of the cut, the number of ways of doing this is $\binom{a+n-r}{n}$.

$\square$

Worpitzky's identity

The number of permutations of S_n with r rising sequences is the Eulerian number $A_{n,r}$. *Worpitzky's identity* states

$$a^n = \sum_{r=1}^{n} A_{n,r} \binom{a+n-r}{n}.$$

The group algebra

Definition

The group algebra $L(S_n)$ is the set of all functions from S_n to the rational numbers. Multiplication is defined by convolution, alternatively, by formally multiplying the linear expressions.

In the group algebra, define

$$A_i = \sum_{\pi : R(\pi) = i} \pi.$$

The group algebra

Theorem

Let $\mathscr{A}$ be the subalgebra of $L(S_n)$ generated by $A_1, \dots, A_n$. $\mathscr{A}$ is a commutative, semisimple algebra of dimension n. A basis of primitive idempotents is given by

$$e_n(l) = \sum_{r=1}^{n} \sigma_l(n-r, \cdots, 1-r) A_r,$$

with σ_l the lth elementary symmetric function.

The group algebra

Proof.

- An a-shuffle may be represented in $\mathscr{A}$ as
$$B_a = \frac{1}{a^n} \sum_{r=1}^{n} \binom{a+n-r}{n} A_r.$$
- B_2^2 is a positive linear combination of A_1, which is the identity, and A_2, A_2^2. Since $B_2^2 = B_4 \in \mathscr{A}$, $A_2^2 \in \mathscr{A}$.
- Since $B_2 B_3 \in \mathscr{A}$ is a positive linear combination of A_2A_3, A_2^2, A_2, A_3 and A_1 we have $A_2A_3 \in \mathscr{A}$. Since $B_2B_3 = B_3B_2 = B_6$, $A_2A_3 = A_3A_2$.
- Continuing inductively proves $\mathscr{A}$ is commutative.

□

The group algebra

Proof.

- Write
$$B_2^m = B_{2^m} = \frac{1}{2^{mn}} \sum_{r=1}^{n} \binom{2^m + n - r}{n} A_r$$
$$= \frac{1}{n!} \sum_{l=0}^{n-1} \frac{1}{2^{lm}} \sum_{r=1}^{n} \sigma_l(n-r, \ldots, 1-r) A_r.$$
It follows that B_2 is a linear map $\mathscr{A} \to \mathscr{A}$ with eigenvalues $\{\frac{1}{2^l}\}_{l=0}^{n-1}$, with the $e_n(\ell)$ as eigenvectors.
- Since, as a linear map, B_2 has distinct eigenvalues and commutes with $\mathscr{A}$, it follows that B_2 generates $\mathscr{A}$. Thus the $e_n(l)$ simultaneously diagonalizes $\mathscr{A}$, so $\mathscr{A}$ is semisimple.

□

Local limit theorem

Now we restrict to n which is even and prove our mixing time theorem in this case.

Lemma

Let $Q^m(r) = \frac{\binom{2^m+n-r}{n}}{2^{mn}}$ be the probability of r rising sequences after m riffle shuffles. Let $r = \frac{n}{2} + h$, $-\frac{n}{2} + 1 \leqslant h \leqslant \frac{n}{2}$, and $m = \lfloor \log_2(n^{\frac{3}{2}} c) \rfloor$ with $0 < c < \infty$ fixed. Then

$$Q^m(r) = \frac{1}{n!} \exp\left(\frac{1}{c\sqrt{n}} \left(-h + \frac{1}{2} + O_c\left(\frac{h}{n}\right)\right) - \frac{1}{24c^2} - \frac{1}{2}\left(\frac{h}{cn}\right)^2 + O_c\left(\frac{1}{n}\right)\right).$$

Proof.

Calculate

$$Q^m(r) = \frac{1}{n!}\left(\frac{2^m+n-r}{2^m}\cdots\frac{2^m+1-r}{2^m}\right)$$
$$= \frac{1}{n!}\exp\left(\sum_{i=0}^{n-1}\log\left(1+\frac{\frac{n}{2}-h-i}{cn^{\frac{3}{2}}}\right)\right).$$

Use the bounds in $-\frac{1}{2} < x < 1$,

$$x - \frac{x^2}{2} + \frac{x^3}{3} - x^4 \leqslant \log(1+x) \leqslant x - \frac{x^2}{2} + \frac{x^3}{3}.$$

□

Local limit theorem

Proof.

Now use the estimates

$$\frac{1}{cn^{\frac{3}{2}}}\sum_{i=0}^{n-1}\left(\frac{n}{2}-h-i\right) = \frac{-h+\frac{1}{2}}{c\sqrt{n}},$$

$$\frac{1}{2c^2n^3}\sum_{i=0}^{n-1}\left(\frac{n}{2}-h-i\right)^2 = \frac{1}{24c^2}+\frac{1}{2}\left(\frac{h}{cn}\right)^2+O_c\left(\frac{1}{n}\right),$$

$$\frac{1}{3c^3n^{\frac{9}{2}}}\sum_{i=0}^{n-1}\left(\frac{n}{2}-h-i\right)^3 = O_c\left(\frac{h}{n^{\frac{3}{2}}}\right),$$

$$\frac{1}{c^4n^6}\sum_{i=0}^{n-1}\left(\frac{n}{2}-h-i\right)^4 = O_c\left(\frac{1}{n}\right).$$

$\square$

Monotonicity

Note that, for fixed m, $Q^m(r) = \frac{\binom{2^m+n-r}{n}}{2^{mn}}$ is monotone decreasing in r.

Lemma

Let h^* be an integer such that $Q^m\left(\frac{n}{2}+h\right) \geqslant \frac{1}{n!}$ if and only if $h \leqslant h^*$. Then for any fixed c, as $n \to \infty$,

$$h^* = \frac{-\sqrt{n}}{24c} + \frac{1}{12c^3} + B + O_c\left(\frac{1}{\sqrt{n}}\right),$$

where $-1 \leqslant B \leqslant 1$.

Proof.

This follows on solving for a non-negative exponent in the local limit theorem. □

Recall the statement of the mixing time theorem.

Theorem

Let Q^m be the G-S-R distribution on the symmetric group S_n and let U be the uniform distribution. For $m = \lfloor \log_2(n^{\frac{3}{2}} c) \rfloor$, with $0 < c < \infty$ fixed, as $n \to \infty$,

$$\|Q^m - U\|_{\mathrm{TV}} = 1 - 2\Phi\left(\frac{-1}{4c\sqrt{3}}\right) + O_c\left(\frac{1}{n^{\frac{1}{2}}}\right)$$

and $\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-\frac{t^2}{2}} dt$.

Mixing

Proof.

- Let $R_{nh} = A_{n,\frac{n}{2}+h}$ denote the number of permutations with $\frac{n}{2} + h$ rising sequences.
- Tanny and Stanley have shown that $\frac{A_{n,j}}{n!}$ is the probability that the sum of n random variables uniform on $[0, 1]$ is between $j - 1$ and j.
- In particular, $\frac{R_{nh}}{n!}$ obeys a local limit theorem: if $x_h = \frac{h}{\sqrt{\frac{n}{12}}}$ then, uniformly in h, and $A > 0$,

$$\frac{R_{nh}}{n!} = \frac{e^{-\frac{x_h^2}{2}}}{\sqrt{\frac{2\pi n}{12}}}\left(1 + O\left(\frac{1}{\sqrt{n}}\right)\right) + O_A(n^{-A}).$$

$\square$

Mixing

Proof.

- Putting together the two local limit theorems, if C is sufficiently large, then

$$\sum_{|h|<C\sqrt{n\log n}} \frac{R_{nh}}{n!} = 1 + O(1/\sqrt{n})$$
$$\sum_{|h|<C\sqrt{n\log n}} R_{nh} Q^m\left(\frac{n}{2}+h\right) = 1 + O(1/\sqrt{n}).$$

Since both distributions have total mass 1, we can truncate to $|h| < C\sqrt{n\log n}$ while making error $O(1/\sqrt{n})$.

□

Mixing

Proof.

- Calculate (recall $h^* = -\frac{\sqrt{n}}{24c} + O(1)$)

$$\sum_{-C\sqrt{n\log n}<h\leqslant h^*} \frac{R_{nh}}{n!} = O\left(\frac{1}{\sqrt{n}}\right) + \sum_{-C\sqrt{n\log n}<h\leqslant h^*} \frac{e^{-\frac{6h^2}{n}}}{\sqrt{\frac{2\pi n}{12}}}$$
$$= \Phi\left(\frac{-1}{4c\sqrt{3}}\right) + O\left(\frac{1}{\sqrt{n}}\right).$$

$\square$

Mixing

Proof.

- Similarly

$$\sum_{-C\sqrt{n\log n}<h\leqslant h^*} R_{nh}Q^{*m}\left(\frac{n}{2}+h\right) = O\left(\frac{1}{\sqrt{n}}\right)$$

$$+\frac{e^{-\frac{1}{24c^2}}}{\sqrt{\frac{2\pi n}{12}}}\sum_{-C\sqrt{n\log n}<h\leqslant h^*}\exp\left(-\frac{6h^2}{n}-\frac{h}{c\sqrt{n}}+O\left(\frac{1+|h|}{n}\right)\right)$$

$$=\Phi\left(\frac{1}{4c\sqrt{3}}\right)+O\left(\frac{1}{\sqrt{n}}\right).$$

- Combined, this gives

$$\|Q^m - U\|_{\mathrm{TV}} = 1-2\Phi\left(\frac{-1}{4c\sqrt{3}}\right)+O_c\left(\frac{1}{n^{\frac{1}{2}}}\right).$$

□