Modern Discrete Probability

VI - Spectral Techniques

Background

Sébastien Roch
UW–Madison
Mathematics

December 1, 2014

1. Review
2. Bounding the mixing time via the spectral gap
3. Applications: random walk on cycle and hypercube
4. Infinite networks

Mixing time I

Theorem (Convergence to stationarity)

Consider a finite state space V. Suppose the transition matrix P is irreducible, aperiodic and has stationary distribution π. Then, for all x, y, $P^t(x, y) \to \pi(y)$ as $t \to +\infty$.

For probability measures μ, ν on V, let their *total variation distance* be $\|\mu - \nu\|_{\mathrm{TV}} := \sup_{A \subseteq V} |\mu(A) - \nu(A)|$.

Definition (Mixing time)

The *mixing time* is

$$\mathrm{t_{mix}}(\varepsilon) := \min\{t \geq 0 \,:\, d(t) \leq \varepsilon\},$$

where $d(t) := \max_{x \in V} \|P^t(x, \cdot) - \pi(\cdot)\|_{\mathrm{TV}}$.

Mixing time II

Definition (Separation distance)

The *separation distance* is defined as

$$s_x(t) := \max_{y \in V} \left[1 - \frac{P^t(x,y)}{\pi(y)} \right],$$

and we let $s(t) := \max_{x \in V} s_x(t)$.

Because both $\{\pi(y)\}$ and $\{P^t(x,y)\}$ are non-negative and sum to 1, we have that $s_x(t) \geq 0$.

Lemma (Separation distance v. total variation distance)

$$d(t) \leq s(t).$$

Mixing time III

Proof: Because $1 = \sum_y \pi(y) = \sum_y P^t(x,y)$,

$$\sum_{y:P^t(x,y)<\pi(y)} \left[\pi(y) - P^t(x,y)\right] = \sum_{y:P^t(x,y)\geq\pi(y)} \left[P^t(x,y) - \pi(y)\right].$$

So

$$\begin{aligned} \|P^t(x,\cdot) - \pi(\cdot)\|_{\mathrm{TV}} &= \frac{1}{2}\sum_y \left|\pi(y) - P^t(x,y)\right| \\ &= \sum_{y:P^t(x,y)<\pi(y)} \left[\pi(y) - P^t(x,y)\right] \\ &= \sum_{y:P^t(x,y)<\pi(y)} \pi(y)\left[1 - \frac{P^t(x,y)}{\pi(y)}\right] \\ &\leq s_x(t). \end{aligned}$$

Reversible chains

Definition (Reversible chain)

A transition matrix P is *reversible* w.r.t. a measure η if $\eta(x)P(x,y) = \eta(y)P(y,x)$ for all $x, y \in V$. By summing over y, such a measure is necessarily stationary.

Example I

Recall:

Definition (Random walk on a graph)

Let $G = (V, E)$ be a finite or countable, locally finite graph. *Simple random walk* on G is the Markov chain on V, started at an arbitrary vertex, which at each time picks a uniformly chosen neighbor of the current state.

Let (X_t) be simple random walk on a connected graph G. Then (X_t) is reversible w.r.t. $\eta(v) := \delta(v)$, where $\delta(v)$ is the degree of vertex v.

Example II

Definition (Random walk on a network)

Let $G = (V, E)$ be a finite or countable, locally finite graph. Let $c : E \to \mathbb{R}_+$ be a positive edge weight function on G. We call $\mathcal{N} = (G, c)$ a *network*. Random walk on $\mathcal{N}$ is the Markov chain on V, started at an arbitrary vertex, which at each time picks a neighbor of the current state proportionally to the weight of the corresponding edge.

Any countable, reversible Markov chain can be seen as a random walk on a network (not necessarily locally finite) by setting $c(e) := \pi(x)P(x, y) = \pi(y)P(y, x)$ for all $e = \{x, y\} \in E$. Let (X_t) be random walk on a network $\mathcal{N} = (G, c)$. Then (X_t) is reversible w.r.t. $\eta(v) := c(v)$, where $c(v) := \sum_{x \sim v} c(v, x)$.

Eigenbasis I

We let $n := |V| < +\infty$. Assume that P is irreducible and reversible w.r.t. its stationary distribution $\pi > 0$. Define

$$\langle f, g\rangle_\pi := \sum_{x \in V} \pi(x) f(x) g(x), \quad \|f\|_\pi^2 := \langle f, f\rangle_\pi,$$

$$(Pf)(x) := \sum_y P(x,y) f(y).$$

We let $\ell^2(V,\pi)$ be the Hilbert space of real-valued functions on V equipped with the inner product $\langle \cdot, \cdot\rangle_\pi$ (equivalent to the vector space $(\mathbb{R}^n, \langle \cdot, \cdot\rangle_\pi)$).

Theorem

There is an orthonormal basis of $\ell^2(V,\pi)$ formed of eigenfunctions $\{f_j\}_{j=1}^n$ of P with real eigenvalues $\{\lambda_j\}_{j=1}^n$.

Eigenbasis II

Proof: We work over $(\mathbb{R}^n, \langle \cdot, \cdot \rangle_\pi)$. Let D_π be the diagonal matrix with π on the diagonal. By reversibility,

$$M(x,y) := \sqrt{\frac{\pi(x)}{\pi(y)}} P(x,y) = \sqrt{\frac{\pi(y)}{\pi(x)}} P(y,x) =: M(y,x).$$

So $M = (M(x,y))_{x,y} = D_\pi^{1/2} P D_\pi^{-1/2}$, as a symmetric matrix, has real eigenvectors $\{\phi_j\}_{j=1}^n$ forming an orthonormal basis of $\mathbb{R}^n$ with corresponding real eigenvalues $\{\lambda_j\}_{j=1}^n$. Define $f_j := D_\pi^{-1/2}\phi_j$. Then

$$Pf_j = PD_\pi^{-1/2}\phi_j = D_\pi^{-1/2} D_\pi^{1/2} P D_\pi^{-1/2}\phi_j = D_\pi^{-1/2} M\phi_j = \lambda_j D_\pi^{-1/2}\phi_j = \lambda_j f_j,$$

and

$$\begin{aligned}
\langle f_i, f_j \rangle_\pi &= \langle D_\pi^{-1/2}\phi_i, D_\pi^{-1/2}\phi_j \rangle_\pi \\
&= \sum_x \pi(x)[\pi(x)^{-1/2}\phi_i(x)][\pi(x)^{-1/2}\phi_j(x)] \\
&= \langle \phi_i, \phi_j \rangle.
\end{aligned}$$

Eigenbasis III

Lemma

For all $j \neq 1$, $\sum_x \pi(x) f_j(x) = 0$.

Proof: By orthonormality, $\langle f_1, f_j \rangle_\pi = 0$. Now use the fact that $f_1 \equiv 1$. ∎

Let $\delta_x(y) := \mathbb{1}_{\{x=y\}}$.

Lemma

For all x, y, $\sum_{j=1}^n f_j(x) f_j(y) = \pi(x)^{-1} \delta_x(y)$.

Proof: Using the notation of the theorem, the matrix Φ whose columns are the ϕ_js is unitary so $\Phi\Phi' = I$. That is, $\sum_{j=1}^n \phi_j(x)\phi_j(y) = \delta_x(y)$, or $\sum_{j=1}^n \sqrt{\pi(x)\pi(y)} f_j(x) f_j(y) = \delta_x(y)$. Rearranging gives the result. ∎

Eigenbasis IV

Lemma

Let $g \in \ell^2(V, \pi)$. Then $g = \sum_{j=1}^n \langle g, f_j \rangle_\pi f_j$.

Proof: By the previous lemma, for all x

$$\sum_{j=1}^n \langle g, f_j \rangle_\pi f_j(x) = \sum_{j=1}^n \sum_y \pi(y) g(y) f_j(y) f_j(x) = \sum_y \pi(y) g(y) [\pi(x)^{-1} \delta_x(y)] = g(x).$$

■

Lemma

Let $g \in \ell^2(V, \pi)$. Then $\|g\|_\pi^2 = \sum_{j=1}^n \langle g, f_j \rangle_\pi^2$.

Proof: By the previous lemma,

$$\|g\|_\pi^2 = \left\| \sum_{j=1}^n \langle g, f_j \rangle_\pi f_j \right\|_\pi^2 = \left\langle \sum_{i=1}^n \langle g, f_i \rangle_\pi f_i, \sum_{j=1}^n \langle g, f_j \rangle_\pi f_j \right\rangle_\pi = \sum_{i,j=1}^n \langle g, f_i \rangle_\pi \langle g, f_j \rangle_\pi \langle f_i, f_j \rangle_\pi,$$

Eigenvalues I

Let P be finite, irreducible and reversible.

Lemma

Any eigenvalue λ of P satisfies $|\lambda| \leq 1$.

Proof: $Pf = \lambda f \implies |\lambda| \|f\|_\infty = \|Pf\|_\infty = \max_x |\sum_y P(x,y) f(y)| \leq \|f\|_\infty$ ∎

We order the eigenvalues $1 \geq \lambda_1 \geq \cdots \geq \lambda_n \geq -1$. In fact:

Lemma

We have $\lambda_1 = 1$ and $\lambda_2 < 1$. Also we can take $f_1 \equiv 1$.

Proof: Because P is stochastic, the all-one vector is a right eigenvector with eigenvalue 1. Any eigenfunction with eigenvalue 1 is P-harmonic. By Corollary 3.22 for a finite, irreducible chain the only harmonic functions are the constant functions. So the eigenspace corresponding to 1 is one-dimensional. Since all eigenvalues are real, we must have $\lambda_2 < 1$. ∎

Eigenvalues II

Theorem (Rayleigh's quotient)

Let P be finite, irreducible and reversible with respect to π. The second largest eigenvalue is characterized by

$$\lambda_2 = \sup\left\{ \frac{\langle f, Pf\rangle_\pi}{\langle f, f\rangle_\pi} \,:\, f \in \ell^2(V,\pi),\ \sum_x \pi(x) f(x) = 0 \right\}.$$

(Similarly, $\lambda_1 = \sup_{f\in\ell^2(V,\pi)} \frac{\langle f,Pf\rangle_\pi}{\langle f,f\rangle_\pi}$.)

Proof: Recalling that $f_1 \equiv 1$, the condition $\sum_x \pi(x) f(x) = 0$ is equivalent to $\langle f_1, f\rangle_\pi = 0$. For such an f, the eigendecomposition is

$$f = \sum_{j=1}^n \langle f, f_j\rangle_\pi f_j = \sum_{j=2}^n \langle f, f_j\rangle_\pi f_j,$$

Eigenvalues III

and

$$Pf = \sum_{j=2}^{n} \langle f, f_j \rangle_\pi \lambda_j f_j,$$

so that

$$\frac{\langle f, Pf \rangle_\pi}{\langle f, f \rangle_\pi} = \frac{\sum_{i=2}^n \sum_{j=2}^n \langle f, f_i \rangle_\pi \langle f, f_j \rangle_\pi \lambda_j \langle f_i, f_j \rangle_\pi}{\sum_{j=2}^n \langle f, f_j \rangle_\pi^2} = \frac{\sum_{j=2}^n \langle f, f_j \rangle_\pi^2 \lambda_j}{\sum_{j=2}^n \langle f, f_j \rangle_\pi^2} \leq \lambda_2.$$

Taking $f = f_2$ achieves the supremum. ■

1. Review

2. Bounding the mixing time via the spectral gap

3. Applications: random walk on cycle and hypercube

4. Infinite networks

Spectral decomposition I

Theorem

Let $\{f_j\}_{j=1}^n$ be the eigenfunctions of a reversible and irreducible transition matrix P with corresponding eigenvalues $\{\lambda_j\}_{j=1}^n$, as defined previously. Assume $\lambda_1 \geq \cdots \geq \lambda_n$. We have the decomposition

$$\frac{P^t(x,y)}{\pi(y)} = 1 + \sum_{j=2}^{n} f_j(x) f_j(y) \lambda_j^t.$$

Spectral decomposition II

Proof: Let F be the matrix whose columns are the eigenvectors $\{f_j\}_{j=1}^n$ and let D_λ be the diagonal matrix with $\{\lambda_j\}_{j=1}^n$ on the diagonal. Using the notation of the eigenbasis theorem,

$$D_\pi^{1/2} P^t D_\pi^{-1/2} = M^t = (D_\pi^{1/2} F) D_\lambda^t (D_\pi^{1/2} F)',$$

which after rearranging becomes

$$P^t D_\pi^{-1} = F D_\lambda^t F'.$$

■

Example: two-state chain I

Let $V := \{0,1\}$ and, for $\alpha, \beta \in (0,1)$,

$$P := \begin{pmatrix} 1-\alpha & \alpha \\ \beta & 1-\beta \end{pmatrix}.$$

Observe that P is reversible w.r.t. to the stationary distribution

$$\pi := \left(\frac{\beta}{\alpha+\beta}, \frac{\alpha}{\alpha+\beta}\right).$$

We know that $f_1 \equiv 1$ is an eigenfunction with eigenvalue 1. As can be checked by direct computation, the other eigenfunction (in vector form) is

$$f_2 := \left(\sqrt{\frac{\alpha}{\beta}}, -\sqrt{\frac{\beta}{\alpha}}\right)',$$

with eigenvalue $\lambda_2 := 1 - \alpha - \beta$. We normalized f_2 so $\|f_2\|_\pi^2 = 1$.

Example: two-state chain II

The spectral decomposition is therefore

$$P^t D_\pi^{-1} = \begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + (1-\alpha-\beta)^t \begin{pmatrix} \frac{\alpha}{\beta} & -1 \\ -1 & \frac{\beta}{\alpha} \end{pmatrix}.$$

Put differently,

$$P^t = \begin{pmatrix} \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \\ \frac{\beta}{\alpha+\beta} & \frac{\alpha}{\alpha+\beta} \end{pmatrix} + (1-\alpha-\beta)^t \begin{pmatrix} \frac{\alpha}{\alpha+\beta} & -\frac{\alpha}{\alpha+\beta} \\ -\frac{\beta}{\alpha+\beta} & \frac{\beta}{\alpha+\beta} \end{pmatrix}.$$

(Note for instance that the case $\alpha + \beta = 1$ corresponds to a rank-one P, which immediately converges.)

Example: two-state chain III

Assume $\beta \geq \alpha$. Then

$$d(t) = \max_x \frac{1}{2} \sum_y |P^t(x,y) - \pi(y)| = \frac{\beta}{\alpha+\beta} |1-\alpha-\beta|^t.$$

As a result,

$$t_{\mathrm{mix}}(\varepsilon) = \left\lceil \frac{\log\left(\varepsilon \frac{\alpha+\beta}{\beta}\right)}{\log|1-\alpha-\beta|} \right\rceil = \left\lceil \frac{\log \varepsilon^{-1} - \log\left(\frac{\alpha+\beta}{\beta}\right)}{\log|1-\alpha-\beta|^{-1}} \right\rceil.$$

Spectral decomposition: again

Recall:

Theorem

Let $\{f_j\}_{j=1}^n$ be the eigenfunctions of a reversible and irreducible transition matrix P with corresponding eigenvalues $\{\lambda_j\}_{j=1}^n$, as defined previously. Assume $\lambda_1 \geq \cdots \geq \lambda_n$. We have the decomposition

$$\frac{P^t(x,y)}{\pi(y)} = 1 + \sum_{j=2}^n f_j(x) f_j(y) \lambda_j^t.$$

Spectral gap

From the spectral decomposition, the speed of convergence of $P^t(x, y)$ to $\pi(y)$ is governed by the largest eigenvalue of P not equal to 1.

Definition (Spectral gap)

The *absolute spectral gap* is $\gamma_* := 1 - \lambda_*$ where $\lambda_* := |\lambda_2| \vee |\lambda_n|$. The *spectral gap* is $\gamma := 1 - \lambda_2$.

Note that the eigenvalues of the lazy version $\frac{1}{2}P + \frac{1}{2}I$ of P are $\{\frac{1}{2}(\lambda_j + 1)\}_{j=1}^n$ which are all nonnegative. So, there, $\gamma_* = \gamma$.

Definition (Relaxation time)

The *relaxation time* is defined as

$$\mathrm{t_{rel}} := \gamma_*^{-1}.$$

Example continued: two-state chain

There two cases:

- $\alpha + \beta \leq 1$: In that case the spectral gap is $\gamma = \gamma_* = \alpha + \beta$ and the relaxation time is $t_{rel} = 1/(\alpha + \beta)$.
- $\alpha + \beta > 1$: In that case the spectral gap is $\gamma = \gamma_* = 2 - \alpha - \beta$ and the relaxation time is $t_{rel} = 1/(2 - \alpha - \beta)$.

Mixing time v. relaxation time I

Theorem

Let P be reversible, irreducible, and aperiodic with stationary distribution π. Let $\pi_{\min} = \min_x \pi(x)$. For all $\varepsilon > 0$,

$$(\mathrm{t_{rel}} - 1)\log\left(\frac{1}{2\varepsilon}\right) \leq \mathrm{t_{mix}}(\varepsilon) \leq \log\left(\frac{1}{\varepsilon\pi_{\min}}\right)\mathrm{t_{rel}}.$$

Proof: We start with the upper bound. By the lemma, it suffices to find t such that $s(t) \leq \varepsilon$. By the spectral decomposition and Cauchy-Schwarz,

$$\left|\frac{P^t(x,y)}{\pi(y)} - 1\right| \leq \lambda_*^t \sum_{j=2}^n |f_j(x)f_j(y)| \leq \lambda_*^t \sqrt{\sum_{j=2}^n f_j(x)^2 \sum_{j=2}^n f_j(y)^2}.$$

By our previous lemma, $\sum_{j=2}^n f_j(x)^2 \leq \pi(x)^{-1}$. Plugging this back above,

$$\left|\frac{P^t(x,y)}{\pi(y)} - 1\right| \leq \lambda_*^t \sqrt{\pi(x)^{-1}\pi(y)^{-1}} \leq \frac{\lambda_*^t}{\pi_{\min}} = \frac{(1-\gamma_*)^t}{\pi_{\min}} \leq \frac{e^{-\gamma_* t}}{\pi_{\min}}.$$

Mixing time v. relaxation time II

The r.h.s. is less than ε when $t \geq \log\left(\frac{1}{\varepsilon \pi_{\min}}\right) \mathrm{t_{rel}}$.

For the lower bound, let f_* be an eigenfunction associated with an eigenvalue achieving $\lambda_* := |\lambda_2| \vee |\lambda_n|$. Let z be such that $|f_*(z)| = \|f_*\|_\infty$. By our previous lemma, $\sum_y \pi(y) f_*(y) = 0$. Hence

$$\lambda_*^t |f_*(z)| = |P^t f_*(z)| = \left|\sum_y [P^t(z,y) f_*(y) - \pi(y) f_*(y)]\right|$$
$$\leq \|f_*\|_\infty \sum_y |P^t(z,y) - \pi(y)| \leq \|f_*\|_\infty 2d(t),$$

so $d(t) \geq \frac{1}{2}\lambda_*^t$. When $t = \mathrm{t_{mix}}(\varepsilon)$, $\varepsilon \geq \frac{1}{2}\lambda_*^{\mathrm{t_{mix}}(\varepsilon)}$. Therefore

$$\mathrm{t_{mix}}(\varepsilon)\left(\frac{1}{\lambda_*} - 1\right) \geq \mathrm{t_{mix}}(\varepsilon) \log\left(\frac{1}{\lambda_*}\right) \geq \log\left(\frac{1}{2\varepsilon}\right).$$

The result follows from $\left(\frac{1}{\lambda_*} - 1\right)^{-1} = \left(\frac{1-\lambda_*}{\lambda_*}\right)^{-1} = \left(\frac{\gamma_*}{1-\gamma_*}\right)^{-1} = \mathrm{t_{rel}} - 1$. ■

1 Review

2 Bounding the mixing time via the spectral gap

3 Applications: random walk on cycle and hypercube

4 Infinite networks

Random walk on the cycle I

Consider simple random walk on an n-cycle. That is, $V := \{0, 1, \ldots, n-1\}$ and $P(x, y) = 1/2$ if and only if $|x - y| = 1 \mod n$.

Lemma (Eigenbasis on the cycle)

For $j = 0, \ldots, n-1$, the function

$$f_j(x) := \cos\left(\frac{2\pi j x}{n}\right), \qquad x = 0, 1, \ldots, n-1,$$

is an eigenfunction of P with eigenvalue

$$\lambda_j := \cos\left(\frac{2\pi j}{n}\right).$$

Random walk on the cycle II

Proof: Note that, for all i, x,

$$\begin{aligned}
\sum_y P(x,y) f_j(y) &= \frac{1}{2}\left[\cos\left(\frac{2\pi j(y-1)}{n}\right) + \cos\left(\frac{2\pi j(y+1)}{n}\right)\right] \\
&= \frac{1}{2}\left[\frac{e^{i\frac{2\pi j(y-1)}{n}} + e^{-i\frac{2\pi j(y-1)}{n}}}{2} + \frac{e^{i\frac{2\pi j(y+1)}{n}} + e^{-i\frac{2\pi j(y+1)}{n}}}{2}\right] \\
&= \left[\frac{e^{i\frac{2\pi jy}{n}} + e^{-i\frac{2\pi jy}{n}}}{2}\right]\left[\frac{e^{i\frac{2\pi j}{n}} + e^{-i\frac{2\pi j}{n}}}{2}\right] \\
&= \left[\cos\left(\frac{2\pi jy}{n}\right)\right]\left[\cos\left(\frac{2\pi j}{n}\right)\right] \\
&= \cos\left(\frac{2\pi j}{n}\right) f_j(y).
\end{aligned}$$

∎

Random walk on the cycle III

Theorem (Relaxation time on the cycle)

The relaxation time for lazy simple random walk on the cycle is

$$t_{\mathrm{rel}} = \frac{2}{1-\cos\left(\frac{2\pi}{n}\right)} = \Theta(n^2).$$

Proof: The eigenvalues are

$$\frac{1}{2}\left[\cos\left(\frac{2\pi j}{n}\right)+1\right].$$

The spectral gap is therefore $\frac{1}{2}(1-\cos\left(\frac{2\pi}{n}\right))$. By a Taylor expansion,

$$1-\cos\left(\frac{2\pi}{n}\right) = \frac{4\pi^2}{n^2} + O(n^{-4}).$$

■

Since $\pi_{\min} = 1/n$, we get $t_{\mathrm{mix}}(\varepsilon) = O(n^2 \log n)$ and $t_{\mathrm{mix}}(\varepsilon) = \Omega(n^2)$. We showed before that in fact $t_{\mathrm{mix}}(\varepsilon) = \Theta(n^2)$.

Random walk on the cycle IV

In this case, a sharper bound can be obtained by working directly with the spectral decomposition. By Jensen's inequality,

$$4\|P^t(x,\cdot)-\pi(\cdot)\|_{\mathrm{TV}}^2 = \left\{\sum_y \pi(y)\left|\frac{P^t(x,y)}{\pi(y)}-1\right|\right\}^2 \leq \sum_y \pi(y)\left(\frac{P^t(x,y)}{\pi(y)}-1\right)^2$$
$$= \left\|\sum_{j=2}^n \lambda_j^t f_j(x) f_j\right\|_\pi^2 = \sum_{j=2}^n \lambda_j^{2t} f_j(x)^2.$$

The last sum does not depend on x by symmetry. Summing over x and dividing by n, which is the same as multiplying by $\pi(x)$, gives

$$4\|P^t(x,\cdot)-\pi(\cdot)\|_{\mathrm{TV}}^2 \leq \sum_x \pi(x)\sum_{j=2}^n \lambda_j^{2t} f_j(x)^2 = \sum_{j=2}^n \lambda_j^{2t}\sum_x \pi(x) f_j(x)^2 = \sum_{j=2}^n \lambda_j^{2t},$$

where we used that $\|f_j\|_\pi^2 = 1$.

Random walk on the cycle V

Consider the non-lazy chain with n odd. We get

$$4d(t)^2 \leq \sum_{j=2}^{n} \cos\left(\frac{2\pi j}{n}\right)^{2t} = 2\sum_{j=1}^{(n-1)/2} \cos\left(\frac{\pi j}{n}\right)^{2t}.$$

For $x \in [0, \pi/2)$, $\cos x \leq e^{-x^2/2}$. (Indeed, let $h(x) = \log(e^{x^2/2}\cos x)$. Then $h'(x) = x - \tan x \leq 0$ since $(\tan x)' = 1 + \tan^2 x \geq 1$ for all x and $\tan 0 = 0$. So $h(x) \leq h(0) = 0$.) Then

$$4d(t)^2 \leq 2\sum_{j=1}^{(n-1)/2} \exp\left(-\frac{\pi^2 j^2}{n^2} t\right) \leq 2\exp\left(-\frac{\pi^2}{n^2} t\right) \sum_{j=1}^{\infty} \exp\left(-\frac{\pi^2 (j^2 - 1)}{n^2} t\right)$$

$$\leq 2\exp\left(-\frac{\pi^2}{n^2} t\right) \sum_{\ell=0}^{\infty} \exp\left(-\frac{3\pi^2 t}{n^2} \ell\right) = \frac{2\exp\left(-\frac{\pi^2}{n^2} t\right)}{1 - \exp\left(-\frac{3\pi^2 t}{n^2}\right)},$$

where we used that $j^2 - 1 \geq 3(j-1)$ for all $j = 1, 2, 3, \ldots$. So $t_{\mathrm{mix}}(\varepsilon) = O(n^2)$.

Random walk on the hypercube I

Consider simple random walk on the hypercube $V := \{-1,+1\}^n$ where $x \sim y$ if $\|x-y\|_1 = 1$. For $J \subseteq [n]$, we let

$$\chi_J(x) = \prod_{j \in J} x_j, \qquad x \in V.$$

These are called *parity functions*.

Lemma (Eigenbasis on the hypercube)

For all $J \subseteq [n]$, the function χ_J is an eigenfunction of P with eigenvalue

$$\lambda_J := \frac{n - 2|J|}{n}.$$

Random walk on the hypercube II

Proof: For $x \in V$ and $i \in [n]$, let $x^{[i]}$ be x where coordinate i is flipped. Note that, for all J, x,

$$\sum_y P(x,y)\chi_J(y) = \sum_{i=1}^n \frac{1}{n}\chi_J(x^{[i]}) = \frac{n-|J|}{n}\chi_J(x) - \frac{|J|}{n}\chi_J(x) = \frac{n-2|J|}{n}\chi_J(x).$$

■

Random walk on the hypercube III

Theorem (Relaxation time on the hypercube)

The relaxation time for lazy simple random walk on the hypercube is

$$\mathrm{t_{rel}} = n.$$

Proof: The eigenvalues are $\frac{n-|J|}{n}$ for $J \subseteq [n]$. The spectral gap is $\gamma_* = \gamma = 1 - \frac{n-1}{n} = \frac{1}{n}$.

■

Because $|V| = 2^n$, $\pi_{\min} = 1/2^n$. Hence we have $\mathrm{t_{mix}}(\varepsilon) = O(n^2)$ and $\mathrm{t_{mix}}(\varepsilon) = \Omega(n)$. We have shown before that in fact $\mathrm{t_{mix}}(\varepsilon) = \Theta(n \log n)$.

Random walk on the hypercube IV

As we did for the cycle, we obtain a sharper bound by working directly with the spectral decomposition. By the same argument,

$$4d(t)^2 \leq \sum_{J \neq \emptyset} \lambda_J^{2t}.$$

Consider the lazy chain again. Then

$$4d(t)^2 \leq \sum_{J \neq \emptyset} \left(\frac{n-|J|}{n}\right)^{2t} = \sum_{\ell=1}^{n} \binom{n}{\ell} \left(1 - \frac{\ell}{n}\right)^{2t} \leq \sum_{\ell=1}^{n} \binom{n}{\ell} \exp\left(-\frac{2t\ell}{n}\right)$$
$$= \left(1 + \exp\left(-\frac{2t}{n}\right)\right)^n - 1.$$

So $t_{\mathrm{mix}}(\varepsilon) \leq \frac{1}{2} n \log n + O(n)$.

1 Review

2 Bounding the mixing time via the spectral gap

3 Applications: random walk on cycle and hypercube

4 Infinite networks

Some remarks about infinite networks I

Remark (Recurrent case)

The previous results cannot in general be extended to infinite networks. Suppose P is irreducible, aperiodic and positive recurrent. Then it can be shown that, if π is the stationary distribution, then for all x

$$\|P^t(x,\cdot) - \pi(\cdot)\|_{\mathrm{TV}} \to 0,$$

as $t \to +\infty$. However, one needs stronger conditions on P than reversibility for the spectral theorem to apply, e.g., compactness (that is, P maps bounded sets to relatively compact sets (i.e. whose closure is compact)).

Some remarks about infinite networks II

Example (A positive recurrent chain whose P is not compact)

For $p < 1/2$, let (X_t) be the birth-death chain with $V := \{0, 1, 2, \ldots\}$, $P(0,0) := 1 - p$, $P(0,1) = p$, $P(x, x+1) := p$ and $P(x, x-1) := 1 - p$ for all $x \geq 1$, and $P(x, y) := 0$ if $|x - y| > 1$. As can be checked by direct computation, P is reversible with respect to the stationary distribution $\pi(x) = (1-\gamma)\gamma^x$ for $x \geq 0$ where $\gamma := \frac{p}{1-p}$. For $j \geq 1$, define $g_j(x) := \pi(j)^{-1/2}\mathbb{1}_{\{x=j\}}$. Then $\|g_j\|_\pi^2 = 1$ for all j so $\{g_j\}_j$ is bounded in $\ell^2(V, \pi)$. On the other hand,

$$Pg_j(x) = p\pi(j)^{-1/2}\mathbb{1}_{\{x=j-1\}} + (1-p)\pi(j)^{-1/2}\mathbb{1}_{\{x=j+1\}}.$$

Some remarks about infinite networks III

Example (Continued)

So

$$\|Pg_j\|_\pi^2 = p^2\pi(j)^{-1}\pi(j-1) + (1-p)^2\pi(j)^{-1}\pi(j+1) = 2p(1-p).$$

Hence $\{Pg_j\}_j$ is also bounded. However, for $j > \ell$

$$\begin{aligned}\|Pg_j - Pg_\ell\|_\pi^2 &\geq (1-p)^2\pi(j)^{-1}\pi(j+1) + p^2\pi(\ell)^{-1}\pi(\ell-1)\\ &= 2p(1-p).\end{aligned}$$

So $\{Pg_j\}_j$ does not have a converging subsequence and therefore is not relatively compact.

Some remarks about infinite networks IV

Most random walks on infinite networks we have encountered so far were transient or null recurrent. In such cases, there is no stationary distribution to converge to. In fact:

Theorem

If P is an irreducible chain which is either transient or null recurrent, we have for all x, y

$$\lim_t P^t(x,y) = 0.$$

Proof: In the transient case, since $\sum_t \mathbb{1}_{X_t=y} < +\infty$ a.s. under $\mathbb{P}_x$, we have $\sum_t P^t(x,y) = \mathbb{E}_x[\sum_t \mathbb{1}_{X_t=y}] < +\infty$ so $P^t(x,y) \to 0$.