Eigenvalues & Mixing Time

Xiaoyu Chen

Overview

Inner Product

Definition

Let V be a vector space over the field $\mathbb{F}$ ($\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) .
A function $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{F}$ is an inner product if for all $x, y, z \in V$ and all $c \in \mathbb{F}$,

(1)	$\langle x, x\rangle \geq 0$	*Nonnegativity*
(1a)	$\langle x, x\rangle = 0$ *iff* $x = 0$	*Positivity*
(2)	$\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$	*Additivity*
(3)	$\langle cx, y\rangle = c\langle x, y\rangle$	*Homogeneity*
(4)	$\langle x, y\rangle = \overline{\langle y, x\rangle}$	*Hermitian Property*

Inner Product

Definition

Let V be a vector space over the field $\mathbb{R}$.
A function $\langle \cdot, \cdot \rangle : V \times V \to \mathbb{F}$ is an inner product if for all $x, y, z \in V$ and all $c \in \mathbb{F}$,

(1)	$\langle x, x \rangle \geq 0$	*Nonnegativity*
(1a)	$\langle x, x \rangle = 0$ *iff* $x = 0$	*Positivity*
(2)	$\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$	*Additivity*
(3)	$\langle cx, y \rangle = c\langle x, y \rangle$	*Homogeneity*
(4)	$\langle x, y \rangle = \langle y, x \rangle$	*Hermitian Property*

Inner Product

Definition

Let V be a vector space over the field $\mathbb{R}$.
A function $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{F}$ is an inner product if for all $x, y, z \in V$ and all $c \in \mathbb{F}$,

(1)	$\langle x, x\rangle \geq 0$	*Nonnegativity*
(1a)	$\langle x, x\rangle = 0$ *iff* $x = 0$	*Positivity*
(2)	$\langle x+y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$	*Additivity*
(3)	$\langle cx, y\rangle = c\langle x, y\rangle$	*Homogeneity*
(4)	$\langle x, y\rangle = \langle y, x\rangle$	*Hermitian Property*

Property

(a) $\langle x, cy\rangle = c\langle x, y\rangle$
(b) $\langle x, y+z\rangle = \langle x, y\rangle + \langle x, z\rangle$
(c) $\langle ax+by, cw+dz\rangle = ac\langle x, w\rangle + bc\langle y, w\rangle + ad\langle x, z\rangle + bd\langle y, z\rangle$
(d) $\langle x, \langle x, y\rangle y\rangle = \langle x, y\rangle^2$
(e) $\langle x, y\rangle = 0$ *for all* $y \in V$ *iff* $x = 0$

Inner Product

Definition

Let V be a vector space over the field $\mathbb{R}$.
A function $\langle\cdot,\cdot\rangle : V \times V \to \mathbb{F}$ is an inner product if for all $x, y, z \in V$ and all $c \in \mathbb{F}$,

(1)	$\langle x, x\rangle \geq 0$	*Nonnegativity*
(1a)	$\langle x, x\rangle = 0$ *iff* $x = 0$	*Positivity*
(2)	$\langle x + y, z\rangle = \langle x, z\rangle + \langle y, z\rangle$	*Additivity*
(3)	$\langle cx, y\rangle = c\langle x, y\rangle$	*Homogeneity*
(4)	$\langle x, y\rangle = \langle y, x\rangle$	*Hermitian Property*

Property

(a) $\langle x, cy\rangle = c\langle x, y\rangle$
(b) $\langle x, y + z\rangle = \langle x, y\rangle + \langle x, z\rangle$
(c) $\langle ax + by, cw + dz\rangle = ac\langle x, w\rangle + bc\langle y, w\rangle + ad\langle x, z\rangle + bd\langle y, z\rangle$
(d) $\langle x, \langle x, y\rangle y\rangle = \langle x, y\rangle^2$
(e) $\langle x, y\rangle = 0$ *for all* $y \in V$ *iff* $x = 0$

Cauchy-Schwarz Inequality

$$\langle x, y\rangle^2 \leq \langle x, x\rangle\langle y, y\rangle \text{ for all } x, y \in V.$$

Self-Adjoint Operator

Definition

$A : V \to V$ is a self-adjoint operator of $\langle \cdot, \cdot \rangle$ if,

$$\langle Ax, y \rangle = \langle x, Ay \rangle \text{ forall } x, y \in V.$$

Self-Adjoint Operator

Definition

$A : V \to V$ is a self-adjoint operator of $\langle \cdot, \cdot \rangle$ if,

$$\langle Ax, y \rangle = \langle x, Ay \rangle \text{ forall } x, y \in V.$$

Theorem

$A : V \to V$ is a self-adjoint operator of $\langle \cdot, \cdot \rangle$ iff A admits an orthonormal eigenbasis with real eigenvalues.

Self-Adjoint Operator

Definition

$A : V \to V$ is a self-adjoint operator of $\langle \cdot, \cdot \rangle$ if,

$$\langle Ax, y \rangle = \langle x, Ay \rangle \text{ forall } x, y \in V.$$

Theorem

$A : V \to V$ is a self-adjoint operator of $\langle \cdot, \cdot \rangle$ iff A admits an orthonormal eigenbasis with real eigenvalues.

Proof of $\Rightarrow$ (high level).

Let B be the collection of all the eigenvectors of A, $\mathcal{S} := B^{\perp}$. Then for all $x \in B, y \in \mathcal{S}$, we have

$$\langle x, Ay \rangle = \langle Ax, y \rangle = \lambda \langle x, y \rangle = 0.$$

So, A is a $\mathcal{S} \to \mathcal{S}$ operator, and it should have at least one eigenfunction in $\mathcal{S}$. □

Total Variation Distance

Definition

Suppose μ and ν are two distributions on $\mathcal{X}$, then

$$\| \mu - \nu \|_{TV} := \max_{A \subseteq \mathcal{X}} |\mu(A) - \nu(A)|$$

Total Variation Distance

Definition

Suppose μ and ν are two distributions on $\mathcal{X}$, then

$$\| \mu - \nu \|_{TV} := \max_{A \subseteq \mathcal{X}} |\mu(A) - \nu(A)|$$

Fact

$$\| \mu - \nu \|_{TV} = \frac{1}{2} \sum_{x \in \mathcal{X}} |\mu(x) - \nu(x)|$$

Total Variation Distance

Definition

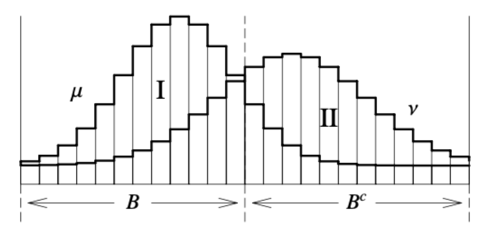
Suppose μ and ν are two distributions on $\mathcal{X}$, then

$$\| \mu - \nu \|_{TV} := \max_{A \subseteq \mathcal{X}} |\mu(A) - \nu(A)|$$

Fact

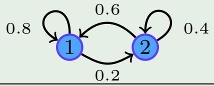
$$\| \mu - \nu \|_{TV} = \frac{1}{2} \sum_{x \in \mathcal{X}} |\mu(x) - \nu(x)|$$

Example



Mixing Time

Example (Markov Chain)



$$P = \begin{bmatrix} 0.8 & 0.2 \\ 0.6 & 0.4 \end{bmatrix}, \pi = \begin{bmatrix} 0.75 \\ 0.25 \end{bmatrix}$$

$$\pi P = \pi$$

Definition

$$d(t) := \max_{x \in \mathcal{X}} \| P^t(x, \cdot) - \pi \|_{TV}$$

Definition (Mixing Time)

$$t_{\text{mix}}(\varepsilon) := \min\{t : d(t) \leq \varepsilon\}$$

Definition ($\ell^p(\pi)$ norm)

Given a distribution π on $\mathcal{X}$ and $1 \leq p \leq \infty$, the $\ell^p(\pi)$ norm of a function $f : \mathcal{X} \to \mathbb{R}$ is defined as:

$$\| f \|_p := \begin{cases} \left[\sum_{y \in \mathcal{X}} |f(y)|^p \pi(y) \right]^{\frac{1}{p}} & 1 \leq p < \infty, \\ \max_{y \in \mathcal{X}} |f(y)| & p = \infty \end{cases}$$

Fact (non-decreasing)

For any $f : \mathcal{X} \to \mathbb{R}$, $\| f \|_p$ is non-decreasing.

Proof.

Recall Jenson's Inequality: $\mathbb{E}_\pi[g(X)] \leq g(\mathbb{E}_\pi[X])$ for all concave $g : \mathcal{X} \to \mathbb{R}$.

Suppose $p < r$, then $x \mapsto x^{p/r}$ is a concave funcion,

$$\mathbb{E}_\pi \left[|f(x)|^p \right] = \mathbb{E}_\pi \left[\left(|f(x)|^r \right)^{\frac{p}{r}} \right] \leq \left(\mathbb{E}_\pi \left[|f(x)|^r \right] \right)^{\frac{p}{r}}$$

□

Definition

For functions $f, g : \mathcal{X} \to \mathbb{R}$, define $\langle \cdot, \cdot \rangle_\pi$ as,

$$\langle f, g \rangle_\pi := \sum_{x \in \mathcal{X}} f(x) g(x) \pi(x)$$

where π is some distribution over $\mathcal{X}$.

$\langle f, f \rangle_\pi = \| f \|_2^2, \langle f, 1 \rangle_\pi = \| f \|_1 = \mathbb{E}_\pi[f]$

ℓ^p distance and ℓ^p mixing

Definition

$$q_t(x, y) := \frac{P^t(x, y)}{\pi(y)}$$

Definition (ℓ^p distance)

$$d^{(p)}(t) := \max_{x \in \mathcal{X}} \| \, q_t(x, \cdot) - 1 \, \|_p$$

ℓ^p distance and ℓ^p mixing

Definition

$$q_t(x,y) := \frac{P^t(x,y)}{\pi(y)}$$

Definition (ℓ^p distance)

$$d^{(p)}(t) := \max_{x\in\mathcal{X}} \| \, q_t(x,\cdot) - 1 \, \|_p$$

Definition (ℓ^p mixing time)

$$t_{\mathrm{mix}}^{(p)}(\varepsilon) := \inf\{t \geq 0 : d^{(p)}(t) \leq \varepsilon\}$$

Fact (non-decreasing)

$$d^{(p)}(t) \leq d^{(p+1)}(t)$$

$$2d(t) = d^{(1)}(t) \leq d^{(2)}(t) \leq d^{(\infty)}(t)$$

ℓ^p distance and ℓ^p mixing

Definition

$$q_t(x,y) := \frac{P^t(x,y)}{\pi(y)}$$

Definition (ℓ^p distance)

$$d^{(p)}(t) := \max_{x \in \mathcal{X}} \| \, q_t(x,\cdot) - 1 \, \|_p$$

Proposition

For a reversible Markov chain,

$$d^{(\infty)}(2t) = [d^{(2)}(t)]^2 = \max_{x \in \mathcal{X}} q_{2t}(x,x) - 1$$

ℓ^p distance and ℓ^p mixing

Definition

$$q_t(x,y) := \frac{P^t(x,y)}{\pi(y)}$$

Definition (ℓ^p distance)

$$d^{(p)}(t) := \max_{x\in\mathcal{X}} \| \, q_t(x,\cdot) - 1 \, \|_p$$

Proposition

For a reversible Markov chain,

$$d^{(\infty)}(2t) = [d^{(2)}(t)]^2 = \max_{x\in\mathcal{X}} q_{2t}(x,x) - 1$$

Proof.

$$q_{2t}(x,y) = \sum_{z\in\mathcal{X}} \frac{P^t(x,z)P^t(z,y)}{\pi(y)}$$

□

ℓ^p distance and ℓ^p mixing

Definition

$$q_t(x,y) := \frac{P^t(x,y)}{\pi(y)}$$

Definition (ℓ^p distance)

$$d^{(p)}(t) := \max_{x\in\mathcal{X}} \| \, q_t(x,\cdot) - 1 \, \|_p$$

Proposition

For a reversible Markov chain,

$$d^{(\infty)}(2t) = [d^{(2)}(t)]^2 = \max_{x\in\mathcal{X}} q_{2t}(x,x) - 1$$

Proof.

$$q_{2t}(x,y) = \sum_{z\in\mathcal{X}} \frac{P^t(x,z)}{\pi(z)} \frac{P^t(z,y)}{\pi(y)} \pi(z)$$

□

ℓ^p distance and ℓ^p mixing

Definition

$$q_t(x,y) := \frac{P^t(x,y)}{\pi(y)}$$

Definition (ℓ^p distance)

$$d^{(p)}(t) := \max_{x\in\mathcal{X}} \| \, q_t(x,\cdot) - 1 \, \|_p$$

Proposition

For a reversible Markov chain,

$$d^{(\infty)}(2t) = [d^{(2)}(t)]^2 = \max_{x\in\mathcal{X}} q_{2t}(x,x) - 1$$

Proof.

$$q_{2t}(x,y) = \sum_{z\in\mathcal{X}} \frac{P^t(x,z)}{\pi(z)} \frac{P^t(y,z)}{\pi(z)} \pi(z)$$

□

ℓ^p distance and ℓ^p mixing

Definition

$$q_t(x,y) := \frac{P^t(x,y)}{\pi(y)}$$

Definition (ℓ^p distance)

$$d^{(p)}(t) := \max_{x\in\mathcal{X}} \| \, q_t(x,\cdot) - 1 \, \|_p$$

Proposition

For a reversible Markov chain,

$$d^{(\infty)}(2t) = [d^{(2)}(t)]^2 = \max_{x\in\mathcal{X}} q_{2t}(x,x) - 1$$

Proof.

$$q_{2t}(x,y) = \sum_{z\in\mathcal{X}} \frac{P^t(x,z)}{\pi(z)} \frac{P^t(y,z)}{\pi(z)} \pi(z) = \langle q_t(x,\cdot), q_t(y,\cdot) \rangle_\pi$$

□

ℓ^p distance and ℓ^p mixing

Definition

$$q_t(x,y) := \frac{P^t(x,y)}{\pi(y)}$$

Definition (ℓ^p distance)

$$d^{(p)}(t) := \max_{x\in\mathcal{X}} \| \, q_t(x,\cdot) - 1 \, \|_p$$

Proposition

For a reversible Markov chain,

$$d^{(\infty)}(2t) = [d^{(2)}(t)]^2 = \max_{x\in\mathcal{X}} q_{2t}(x,x) - 1$$

Proof.

$$q_{2t}(x,y) = \sum_{z\in\mathcal{X}} \frac{P^t(x,z)}{\pi(z)} \frac{P^t(y,z)}{\pi(z)} \pi(z) = \langle q_t(x,\cdot), q_t(y,\cdot)\rangle_\pi$$

$$\langle q_t(x,\cdot), 1\rangle\pi = 1 \Rightarrow \langle q_t(x,\cdot) - 1, q_t(y,\cdot) - 1\rangle_\pi = q_{2t}(x,y) - 1$$

$\square$

ℓ^p distance and ℓ^p mixing

Definition

$$q_t(x,y) := \frac{P^t(x,y)}{\pi(y)}$$

Definition (ℓ^p distance)

$$d^{(p)}(t) := \max_{x\in\mathcal{X}} \| \, q_t(x,\cdot) - 1 \, \|_p$$

Proposition

For a reversible Markov chain,

$$d^{(\infty)}(2t) = [d^{(2)}(t)]^2 = \max_{x\in\mathcal{X}} q_{2t}(x,x) - 1$$

Proof.

$$q_{2t}(x,y) = \sum_{z\in\mathcal{X}} \frac{P^t(x,z)}{\pi(z)} \frac{P^t(y,z)}{\pi(z)} \pi(z) = \langle q_t(x,\cdot), q_t(y,\cdot)\rangle_\pi$$

$$\langle q_t(x,\cdot), 1\rangle\pi = 1 \Rightarrow \langle q_t(x,\cdot) - 1, q_t(y,\cdot) - 1\rangle_\pi = q_{2t}(x,y) - 1$$

$$[d^{(2)}(t)]^2 = \max_x \| \, q_t(x,\cdot) - 1 \, \|_2^2 = \max_x \langle q_t(x,\cdot) - 1, q_t(x,\cdot) - 1\rangle_\pi = \max_x q_{2t}(x,x) - 1$$

$\square$

ℓ^p distance and ℓ^p mixing

Definition

$$q_t(x,y) := \frac{P^t(x,y)}{\pi(y)}$$

Definition (ℓ^p distance)

$$d^{(p)}(t) := \max_{x\in\mathcal{X}} \| \, q_t(x,\cdot) - 1 \, \|_p$$

Proposition

For a reversible Markov chain,

$$d^{(\infty)}(2t) = [d^{(2)}(t)]^2 = \max_{x\in\mathcal{X}} q_{2t}(x,x) - 1$$

Proof.

$$q_{2t}(x,y) = \sum_{z\in\mathcal{X}} \frac{P^t(x,z)}{\pi(z)} \frac{P^t(y,z)}{\pi(z)} \pi(z) = \langle q_t(x,\cdot), q_t(y,\cdot)\rangle_\pi$$

$$\langle q_t(x,\cdot), 1\rangle \pi = 1 \Rightarrow \langle q_t(x,\cdot) - 1, q_t(y,\cdot) - 1\rangle_\pi = q_{2t}(x,y) - 1$$

$$[d^{(2)}(t)]^2 = \max_x \| \, q_t(x,\cdot) - 1 \, \|_2^2 = \max_x \langle q_t(x,\cdot) - 1, q_t(x,\cdot) - 1\rangle_\pi = \max_x q_{2t}(x,x) - 1$$

By Cauchy-Schwarz Inequality:

$$\begin{aligned} \langle q_t(x,\cdot) - 1, q_t(y,\cdot) - 1\rangle_\pi^2 &\leq \langle q_t(x,\cdot) - 1, q_t(x,\cdot) - 1\rangle_\pi \langle q_t(y,\cdot) - 1, q_t(y,\cdot) - 1\rangle_\pi \\ |q_{2t}(x,y) - 1| &\leq \sqrt{q_{2t}(x,x) - 1}\sqrt{q_{2t}(y,y) - 1} \end{aligned}$$

$$d^{(\infty)}(2t) = \max_{x,y\in\mathcal{X}} |q_{2t}(x,y) - 1| = \max_{x\in\mathcal{X}} q_{2t}(x,x) - 1 \qquad \square$$

Lemma

Let P be the transition matrix of a finite Markov chain.

1. *If λ is an eigenvalue of P, then $|\lambda| \leq 1$.*
2. *If P is irreducible, $P\mathbb{1} = \mathbb{1}$.*
3. *If P is irreducible and aperiodic, then -1 is not an eigenvalue of P.*

Lemma

Let P be the transition matrix of a finite Markov chain.

1. *If λ is an eigenvalue of P, then $|\lambda| \leq 1$.*
2. *If P is irreducible, $P\mathbb{1} = \mathbb{1}$.*
3. *If P is irreducible and aperiodic, then -1 is not an eigenvalue of P.*

Proof of (1).

$$Pf(x) = \sum_y P(x,y)f(x) \leq \max_x f(x)$$

□

Reversible Chain

Suppose we have a Markov chain $\mathcal{M}$ with transition matrix P and some distribution π.

Definition (detailed balance equation)

$$\pi(x)P(x,y) = \pi(y)P(y,x) \textit{ for all } x,y \in \mathcal{X}.$$

Reversible Chain

Suppose we have a Markov chain $\mathcal{M}$ with transition matrix P and some distribution π.

Definition (detailed balance equation)

$$\pi(x)P(x,y) = \pi(y)P(y,x) \textit{ for all } x,y \in \mathcal{X}.$$

Example (stationary distribution)

$$\pi P(x) = \sum_y \pi(y)P(y,x) = \sum_y \pi(x)P(x,y) = \pi(x)$$

Example (reversibility)

$$\begin{aligned} & \pi(x_0)P(x_0,x_1)P(x_1,x_2)P(x_2,x_3)\cdots P(x_{n-1},x_n) \\ = \quad & P(x_1,x_0)\pi(x_1)P(x_1,x_2)P(x_2,x_3)\cdots P(x_{n-1},x_n) \\ = \quad & \pi(x_n)P(x_n,x_{n-1})\cdots P(x_3,x_2)P(x_2,x_1)P(x_1,x_0) \end{aligned}$$

Reversible Chain

Suppose we have a Markov chain $\mathcal{M}$ with transition matrix P and some distribution π.

Definition (detailed balance equation)

$$\pi(x)P(x,y) = \pi(y)P(y,x) \textit{ for all } x, y \in \mathcal{X}.$$

Fact (self adjoint)

$$\begin{aligned}\langle f, Pg\rangle_\pi &= \sum_x f(x)Pg(x)\pi(x) = \sum_x \sum_y f(x)P(x,y)g(y)\pi(x) \\ &= \sum_x \sum_y f(x)P(y,x)g(y)\pi(y) \\ &= \sum_y \sum_x g(y)P(y,x)f(x)\pi(y) = \langle Pf, g\rangle_\pi\end{aligned}$$

Eigenbasis of Reversible Chain

Lemma

Let P be reversible w.r.t π, then,

1. *The inner product space $(\mathbb{R}^{\mathcal{X}}, \langle \cdot, \cdot \rangle_\pi)$ has an orthonormal basis of real-valued eigenfunctions $\{f_j\}_{j=1}^{|\mathcal{X}|}$ corresponding to real eigenvalues $\{\lambda_j\}$.*
2. $q_t(x,y) = \dfrac{P^t(x,y)}{\pi(y)} = \displaystyle\sum_{j=1}^{|\mathcal{X}|} f_j(x) f_j(y) \lambda_j^t = 1 + \sum_{j=2}^{|\mathcal{X}|} f_j(x) f_j(y) \lambda_j^t$

Eigenbasis of Reversible Chain

Lemma

Let P be reversible w.r.t π, then,

1. *The inner product space $(\mathbb{R}^{\mathcal{X}}, \langle \cdot, \cdot \rangle_\pi)$ has an orthonormal basis of real-valued eigenfunctions $\{f_j\}_{j=1}^{|\mathcal{X}|}$ corresponding to real eigenvalues $\{\lambda_j\}$.*
2. $q_t(x,y) = \dfrac{P^t(x,y)}{\pi(y)} = \displaystyle\sum_{j=1}^{|\mathcal{X}|} f_j(x) f_j(y) \lambda_j^t = 1 + \sum_{j=2}^{|\mathcal{X}|} f_j(x) f_j(y) \lambda_j^t$

(1) is a corollary of theorem 1.

Example (how to find such basis)

Since $\pi(x)P(x,y) = \pi(y)P(y,x)$.
Let $D = \mathrm{diag}(\pi)$, then $A = D^{\frac{1}{2}} P D^{-\frac{1}{2}}$ is a real symmetric matrix $\Rightarrow \{\phi_i\}$.
$PD^{-\frac{1}{2}}\phi_i = D^{-\frac{1}{2}}(D^{\frac{1}{2}} P D^{-\frac{1}{2}})\phi_i = \lambda_i D^{-\frac{1}{2}}\phi_i \Rightarrow D^{-\frac{1}{2}}\phi_i$ is an eigenfunction of P.
For $i \neq j$: $\langle D^{-\frac{1}{2}}\phi_i, D^{-\frac{1}{2}}\phi_j \rangle_\pi = \langle \phi_i, \phi_j \rangle = 0 \Rightarrow D^{-\frac{1}{2}}\phi_i \perp_\pi D^{-\frac{1}{2}}\phi_j$.
Finally, normalize each ϕ_i to make sure that $\langle \phi_i, \phi_i \rangle_\pi = 1$.

Eigenbasis of Reversible Chain

Lemma

Let P be reversible w.r.t π, then,

1. *The inner product space $(\mathbb{R}^{\mathcal{X}}, \langle \cdot, \cdot \rangle_\pi)$ has an orthonormal basis of real-valued eigenfunctions $\{f_j\}_{j=1}^{|\mathcal{X}|}$ corresponding to real eigenvalues $\{\lambda_j\}$.*
2. $q_t(x,y) = \dfrac{P^t(x,y)}{\pi(y)} = \displaystyle\sum_{j=1}^{|\mathcal{X}|} f_j(x) f_j(y) \lambda_j^t = 1 + \sum_{j=2}^{|\mathcal{X}|} f_j(x) f_j(y) \lambda_j^t$

(1) is a corollary of theorem 1.

Example (how to find such basis)

Since $\pi(x)^{1/2} P(x,y) \pi(y)^{-1/2} = \pi(y)^{1/2} P(y,x) \pi(x)^{-1/2}$.
Let $D = \mathrm{diag}(\pi)$, then $A = D^{\frac{1}{2}} P D^{-\frac{1}{2}}$ is a real symmetric matrix $\Rightarrow \{\phi_i\}$.
$P D^{-\frac{1}{2}} \phi_i = D^{-\frac{1}{2}} (D^{\frac{1}{2}} P D^{-\frac{1}{2}}) \phi_i = \lambda_i D^{-\frac{1}{2}} \phi_i \Rightarrow D^{-\frac{1}{2}} \phi_i$ is an eigenfunction of P.
For $i \neq j$: $\langle D^{-\frac{1}{2}} \phi_i, D^{-\frac{1}{2}} \phi_j \rangle_\pi = \langle \phi_i, \phi_j \rangle = 0 \Rightarrow D^{-\frac{1}{2}} \phi_i \perp_\pi D^{-\frac{1}{2}} \phi_j$.
Finally, normalize each ϕ_i to make sure that $\langle \phi_i, \phi_i \rangle_\pi = 1$.

Eigenbasis of Reversible Chain

Lemma

Let P be reversible w.r.t π, then,

1. *The inner product space $(\mathbb{R}^{\mathcal{X}}, \langle\cdot,\cdot\rangle_\pi)$ has an orthonormal basis of real-valued eigenfunctions $\{f_j\}_{j=1}^{|\mathcal{X}|}$ corresponding to real eigenvalues $\{\lambda_j\}$.*
2. $$q_t(x,y) = \frac{P^t(x,y)}{\pi(y)} = \sum_{j=1}^{|\mathcal{X}|} f_j(x)f_j(y)\lambda_j^t = 1 + \sum_{j=2}^{|\mathcal{X}|} f_j(x)f_j(y)\lambda_j^t$$

Proof of (2).

$\delta_x(y) = \begin{cases} 1 & y = x \\ 0 & y \neq x \end{cases}$, then $P(x,y) = P\delta_y(x)$.

$$\delta_y = \sum_{j=1}^{|\mathcal{X}|} \langle \delta_y, f_j\rangle_\pi f_j = \sum_{j=1}^{|\mathcal{X}|} \pi(y)f_j(y)f_j.$$

$$P^t\delta_y = \sum_{j=1}^{|\mathcal{X}|} \pi(y)f_j(y)f_j\lambda_j^t$$

□

Eigenbasis of Reversible Chain

Lemma

Let P be reversible w.r.t π, then,

1. *The inner product space $(\mathbb{R}^{\mathcal{X}}, \langle \cdot, \cdot \rangle_\pi)$ has an orthonormal basis of real-valued eigenfunctions $\{f_j\}_{j=1}^{|\mathcal{X}|}$ corresponding to real eigenvalues $\{\lambda_j\}$.*
2. $$q_t(x,y) = \frac{P^t(x,y)}{\pi(y)} = \sum_{j=1}^{|\mathcal{X}|} f_j(x) f_j(y) \lambda_j^t = 1 + \sum_{j=2}^{|\mathcal{X}|} f_j(x) f_j(y) \lambda_j^t$$

Proof of (2).

$$\delta_x(y) = \begin{cases} 1 & y = x \\ 0 & y \neq x \end{cases}, \text{ then } P(x,y) = P\delta_y(x).$$

$$\delta_y = \sum_{j=1}^{|\mathcal{X}|} \langle \delta_y, f_j \rangle_\pi f_j = \sum_{j=1}^{|\mathcal{X}|} \pi(y) f_j(y) f_j.$$

$$P^t \delta_y(x) = \sum_{j=1}^{|\mathcal{X}|} \pi(y) f_j(y) f_j(x) \lambda_j^t$$

□

Definition

$\lambda_\star = \max\{|\lambda| : \lambda \textit{ is an eigenvalue of } P,\ \lambda \neq 1\}$
$\gamma_\star = 1 - \lambda_\star$ *is called the* ***absolute spectral gap****.*

Sort all the eigenvalues of P in order: $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{|\mathcal{X}|} \geq -1$.
$\gamma = 1 - \lambda_2$ *is called* ***spectral gap****.*

---

relaxation time: $t_{\text{tel}} = \dfrac{1}{\gamma_\star}$

Definition

$\lambda_\star = \max\{|\lambda| : \lambda \textit{ is an eigenvalue of } P, \lambda \neq 1\}$
$\gamma_\star = 1 - \lambda_\star$ *is called the* ***absolute spectral gap***.

Sort all the eigenvalues of P *in order:* $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{|\mathcal{X}|} \geq -1$.
$\gamma = 1 - \lambda_2$ *is called* ***spectral gap***.

relaxation time: $t_{\text{tel}} = \dfrac{1}{\gamma_\star}$

Example (lazy chain)

The lazy version of P is $(I + P)/2$. If the chain is lazy, then $\gamma_\star = \gamma$.

Definition

$\lambda_\star = \max\{|\lambda| : \lambda$ *is an eigenvalue of* P, $\lambda \neq 1\}$
$\gamma_\star = 1 - \lambda_\star$ *is called the* ***absolute spectral gap***.

Sort all the eigenvalues of P *in order:* $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{|\mathcal{X}|} \geq -1$.
$\gamma = 1 - \lambda_2$ *is called* ***spectral gap***.

---

relaxation time: $t_{\mathrm{tel}} = \dfrac{1}{\gamma_\star}$

Fact

$$\mathrm{Var}_\pi(P^t f) \leq (1 - \gamma_\star)^{2t} \mathrm{Var}_\pi(f)$$

Definition

$\lambda_\star = \max\{|\lambda| : \lambda$ *is an eigenvalue of* P, $\lambda \neq 1\}$
$\gamma_\star = 1 - \lambda_\star$ *is called the* ***absolute spectral gap***.

Sort all the eigenvalues of P *in order:* $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{|\mathcal{X}|} \geq -1$.
$\gamma = 1 - \lambda_2$ *is called* ***spectral gap***.

relaxation time: $t_{\mathrm{tel}} = \dfrac{1}{\gamma_\star}$

Fact

$$\mathrm{Var}_\pi(P^t f) \leq (1-\gamma_\star)^{2t}\mathrm{Var}_\pi(f)$$

Example

For $A \subseteq \mathcal{X}$, let $f(x) = \begin{cases} 1 & x \in A \\ 0 & x \notin A \end{cases}$.

$$\begin{aligned}\mathrm{Var}_\pi(P^t f) &= \mathbb{E}_\pi[(P^t f(x) - \mathbb{E}_\pi[P^t f])^2] \\ &= \mathbb{E}_\pi[(P^t f(x) - \pi P^t f)^2] \\ &= \mathbb{E}_\pi[(P^t f(x) - \pi f)^2] \\ &= \mathbb{E}_\pi[(P^t(x,A) - \pi(A))^2] \geq 0\end{aligned}$$

$$\mathrm{Var}_\pi(f) \leq 1$$

So, $\mathrm{Var}_\pi(P^t f)$ could be used to measure the distance between $P^t(x,\cdot)$ and π.

Definition

$\lambda_\star = \max\{|\lambda| : \lambda$ *is an eigenvalue of* P, $\lambda \neq 1\}$
$\gamma_\star = 1 - \lambda_\star$ *is called the* ***absolute spectral gap****.*

Sort all the eigenvalues of P *in order:* $1 = \lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_{|\mathcal{X}|} \geq -1$.
$\gamma = 1 - \lambda_2$ *is called* ***spectral gap****.*

relaxation time: $t_{\mathrm{tel}} = \dfrac{1}{\gamma_\star}$

Fact

$$\mathrm{Var}_\pi(P^t f) \leq (1 - \gamma_\star)^{2t} \mathrm{Var}_\pi(f)$$

Proof.

Recall that $\mathrm{Var}_\pi(X + c) = \mathrm{Var}_\pi(X)$,

Let $a_i = \langle f, f_i \rangle_\pi$, then

$$\mathrm{Var}_\pi(P^t f) = \mathrm{Var}_\pi(\sum_{i=1}^{|\mathcal{X}|} a_i f_i \lambda_i^t) = \mathrm{Var}_\pi(\sum_{i=2}^{|\mathcal{X}|} a_i f_i \lambda_i^t).$$

$\square$

Theorem

Let P be the transition matrix of a reversible, irreducible Markov chain with state space $\mathcal{X}$, and let $\pi_{\min} := \min_{x \in \mathcal{X}} \pi(x)$. Then

$$t_{\text{mix}}^{(\infty)}(\varepsilon) \leq \lceil t_{\text{rel}} \log(\frac{1}{\varepsilon \pi_{\min}}) \rceil$$

$$t_{\text{mix}}(\varepsilon) \leq \lceil t_{\text{rel}}(\frac{1}{2} \log(\frac{1}{\pi_{\min}}) + \log(\frac{1}{2\varepsilon})) \rceil$$

Theorem

Let P be the transition matrix of a reversible, irreducible Markov chain with state space $\mathcal{X}$, and let $\pi_{\text{min}} := \min_{x \in \mathcal{X}} \pi(x)$. Then

$$t_{\text{mix}}^{(\infty)}(\varepsilon) \leq \lceil t_{\text{rel}} \log(\frac{1}{\varepsilon \pi_{\text{min}}}) \rceil$$

$$t_{\text{mix}}(\varepsilon) \leq \lceil t_{\text{rel}}(\frac{1}{2} \log(\frac{1}{\pi_{\text{min}}}) + \log(\frac{1}{2\varepsilon})) \rceil$$

Proof.

$$|q_t(x,y) - 1| = |\sum_{j=2}^{|\mathcal{X}|} f_j(x) f_j(y) \lambda_j^t|$$

□

Theorem

Let P be the transition matrix of a reversible, irreducible Markov chain with state space $\mathcal{X}$, and let $\pi_{\min} := \min_{x \in \mathcal{X}} \pi(x)$. Then

$$t_{\text{mix}}^{(\infty)}(\varepsilon) \leq \lceil t_{\text{rel}} \log(\frac{1}{\varepsilon \pi_{\min}}) \rceil$$

$$t_{\text{mix}}(\varepsilon) \leq \lceil t_{\text{rel}} (\frac{1}{2} \log(\frac{1}{\pi_{\min}}) + \log(\frac{1}{2\varepsilon})) \rceil$$

Proof.

$$|q_t(x,y) - 1| \leq \sum_{j=2}^{|\mathcal{X}|} |f_j(x) f_j(y) \lambda_j^t|$$

$\square$

Theorem

Let P be the transition matrix of a reversible, irreducible Markov chain with state space $\mathcal{X}$, and let $\pi_{\min} := \min_{x \in \mathcal{X}} \pi(x)$. Then

$$t_{\mathrm{mix}}^{(\infty)}(\varepsilon) \leq \lceil t_{\mathrm{rel}} \log(\frac{1}{\varepsilon \pi_{\min}}) \rceil$$

$$t_{\mathrm{mix}}(\varepsilon) \leq \lceil t_{\mathrm{rel}} (\frac{1}{2} \log(\frac{1}{\pi_{\min}}) + \log(\frac{1}{2\varepsilon})) \rceil$$

Proof.

$$|q_t(x,y) - 1| \leq \lambda_\star^t \sum_{j=2}^{|\mathcal{X}|} |f_j(x)||f_j(y)|$$

□

Theorem

Let P be the transition matrix of a reversible, irreducible Markov chain with state space $\mathcal{X}$, and let $\pi_{\min} := \min_{x \in \mathcal{X}} \pi(x)$. Then

$$t_{\mathrm{mix}}^{(\infty)}(\varepsilon) \le \lceil t_{\mathrm{rel}} \log(\frac{1}{\varepsilon \pi_{\min}}) \rceil$$

$$t_{\mathrm{mix}}(\varepsilon) \le \lceil t_{\mathrm{rel}} (\frac{1}{2} \log(\frac{1}{\pi_{\min}}) + \log(\frac{1}{2\varepsilon})) \rceil$$

Proof.

$$|q_t(x,y) - 1| \le \lambda_\star^t \sum_{j=2}^{|\mathcal{X}|} |f_j(x)||f_j(y)| \le \lambda_\star^t \left[\sum_{j=2}^{|\mathcal{X}|} f_j^2(x) \sum_{j=2}^{|\mathcal{X}|} f_j^2(y) \right]^{1/2}$$

□

Theorem

Let P be the transition matrix of a reversible, irreducible Markov chain with state space $\mathcal{X}$, and let $\pi_{\min} := \min_{x \in \mathcal{X}} \pi(x)$. Then

$$t_{\text{mix}}^{(\infty)}(\varepsilon) \leq \lceil t_{\text{rel}} \log(\frac{1}{\varepsilon \pi_{\min}}) \rceil$$

$$t_{\text{mix}}(\varepsilon) \leq \lceil t_{\text{rel}} (\frac{1}{2} \log(\frac{1}{\pi_{\min}}) + \log(\frac{1}{2\varepsilon})) \rceil$$

Proof.

$$|q_t(x,y) - 1| \leq \lambda_\star^t \sum_{j=2}^{|\mathcal{X}|} |f_j(x)||f_j(y)| \leq \lambda_\star^t \left[\sum_{j=2}^{|\mathcal{X}|} f_j^2(x) \sum_{j=2}^{|\mathcal{X}|} f_j^2(y) \right]^{1/2}$$

$$\pi(x) = \langle \delta_x, \delta_x \rangle_\pi = \left\langle \sum_{j=1}^{|\mathcal{X}|} f_j(x)\pi(x) f_j, \sum_{j=1}^{|\mathcal{X}|} f_j(x)\pi(x) f_j \right\rangle_\pi = \pi(x)^2 \sum_{j=1}^{|\mathcal{X}|} f_j(x)^2$$

$\square$

Theorem

Let P be the transition matrix of a reversible, irreducible Markov chain with state space $\mathcal{X}$, and let $\pi_{\min} := \min_{x \in \mathcal{X}} \pi(x)$. Then

$$t_{\text{mix}}^{(\infty)}(\varepsilon) \leq \lceil t_{\text{rel}} \log(\frac{1}{\varepsilon \pi_{\min}}) \rceil$$

$$t_{\text{mix}}(\varepsilon) \leq \lceil t_{\text{rel}} (\frac{1}{2} \log(\frac{1}{\pi_{\min}}) + \log(\frac{1}{2\varepsilon})) \rceil$$

Proof.

$$|q_t(x,y) - 1| \leq \lambda_\star^t \sum_{j=2}^{|\mathcal{X}|} |f_j(x)||f_j(y)| \leq \lambda_\star^t \left[\sum_{j=2}^{|\mathcal{X}|} f_j^2(x) \sum_{j=2}^{|\mathcal{X}|} f_j^2(y) \right]^{1/2}$$

$$\pi(x)^{-1} = \sum_{j=1}^{|\mathcal{X}|} f_j(x)^2 \qquad \square$$

Theorem

Let P be the transition matrix of a reversible, irreducible Markov chain with state space $\mathcal{X}$, and let $\pi_{\min} := \min_{x \in \mathcal{X}} \pi(x)$. Then

$$t_{\text{mix}}^{(\infty)}(\varepsilon) \leq \lceil t_{\text{rel}} \log(\frac{1}{\varepsilon \pi_{\min}}) \rceil$$

$$t_{\text{mix}}(\varepsilon) \leq \lceil t_{\text{rel}} (\frac{1}{2} \log(\frac{1}{\pi_{\min}}) + \log(\frac{1}{2\varepsilon})) \rceil$$

Proof.

$$|q_t(x,y) - 1| \leq \lambda_\star^t \sum_{j=2}^{|\mathcal{X}|} |f_j(x)||f_j(y)| \leq \lambda_\star^t \left[\sum_{j=2}^{|\mathcal{X}|} f_j^2(x) \sum_{j=2}^{|\mathcal{X}|} f_j^2(y) \right]^{1/2}$$

$$\pi(x)^{-1} \geq \sum_{j=2}^{|\mathcal{X}|} f_j(x)^2$$

$\square$

Theorem

Let P be the transition matrix of a reversible, irreducible Markov chain with state space $\mathcal{X}$, and let $\pi_{\min} := \min_{x \in \mathcal{X}} \pi(x)$. Then

$$t_{\text{mix}}^{(\infty)}(\varepsilon) \leq \lceil t_{\text{rel}} \log(\frac{1}{\varepsilon \pi_{\min}}) \rceil$$

$$t_{\text{mix}}(\varepsilon) \leq \lceil t_{\text{rel}} (\frac{1}{2} \log(\frac{1}{\pi_{\min}}) + \log(\frac{1}{2\varepsilon})) \rceil$$

Proof.

$$|q_t(x,y) - 1| \leq \lambda_\star^t \sum_{j=2}^{|\mathcal{X}|} |f_j(x)||f_j(y)| \leq \lambda_\star^t \left[\sum_{j=2}^{|\mathcal{X}|} f_j^2(x) \sum_{j=2}^{|\mathcal{X}|} f_j^2(y) \right]^{1/2}$$

$$\pi(x)^{-1} \geq \sum_{j=2}^{|\mathcal{X}|} f_j(x)^2$$

$$d^{(\infty)}(t) \leq \max_{x,y} |q_t(x,y) - 1| \leq \frac{\lambda_\star^t}{\sqrt{\pi(x)\pi(y)}} \leq \frac{\lambda_\star^t}{\pi_{\min}} = \frac{(1-\gamma_\star)^t}{\pi_{\min}} \leq \frac{e^{-\gamma_\star t}}{\pi_{\min}}$$

$\square$

Theorem

Let P be the transition matrix of a reversible, irreducible Markov chain with state space $\mathcal{X}$, and let $\pi_{\min} := \min_{x \in \mathcal{X}} \pi(x)$. Then

$$t_{\mathrm{mix}}^{(\infty)}(\varepsilon) \leq \lceil t_{\mathrm{rel}} \log(\frac{1}{\varepsilon \pi_{\min}}) \rceil$$

$$t_{\mathrm{mix}}(\varepsilon) \leq \lceil t_{\mathrm{rel}} (\frac{1}{2} \log(\frac{1}{\pi_{\min}}) + \log(\frac{1}{2\varepsilon})) \rceil$$

Proof.

$$|q_t(x,y) - 1| \leq \lambda_\star^t \sum_{j=2}^{|\mathcal{X}|} |f_j(x)||f_j(y)| \leq \lambda_\star^t \left[\sum_{j=2}^{|\mathcal{X}|} f_j^2(x) \sum_{j=2}^{|\mathcal{X}|} f_j^2(y) \right]^{1/2}$$

$$\pi(x)^{-1} \geq \sum_{j=2}^{|\mathcal{X}|} f_j(x)^2$$

$$d^{(\infty)}(t) \leq \max_{x,y} |q_t(x,y) - 1| \leq \frac{\lambda_\star^t}{\sqrt{\pi(x)\pi(y)}} \leq \frac{\lambda_\star^t}{\pi_{\min}} = \frac{(1-\gamma_\star)^t}{\pi_{\min}} \leq \frac{e^{-\gamma_\star t}}{\pi_{\min}}$$

$$d(t) \leq \frac{1}{2} \sqrt{d^{(\infty)}(2t)} \leq \frac{e^{-\gamma_\star t}}{2\sqrt{\pi_{\min}}}$$

□

Theorem

Suppose that $\lambda \neq 1$ is an eigenvalue for the transition matrix P of an irreducible and aperiodic Markov chain. Then

$$t_{\text{mix}}(\varepsilon) \geq \left(\frac{1}{1-|\lambda|} - 1\right) \log\left(\frac{1}{2\varepsilon}\right)$$

In particular, for reversible chains,

$$t_{\text{mix}}(\varepsilon) \geq (t_{\text{rel}} - 1) \log\left(\frac{1}{2\varepsilon}\right)$$

Theorem

Suppose that $\lambda \neq 1$ is an eigenvalue for the transition matrix P of an irreducible and aperiodic Markov chain. Then

$$t_{\text{mix}}(\varepsilon) \geq \left(\frac{1}{1-|\lambda|} - 1\right) \log\left(\frac{1}{2\varepsilon}\right)$$

In particular, for reversible chains,

$$t_{\text{mix}}(\varepsilon) \geq (t_{\text{rel}} - 1) \log\left(\frac{1}{2\varepsilon}\right)$$

Proof.

First note that $\mathbb{E}_\pi[f_i] = \pi f_i = \pi P f_i = \lambda_i \pi f_i$, so if $\lambda_i \neq 1$, then $\mathbb{E}_\pi[f_i] = 0$.

$$|\lambda^t f(x)| = |P^t f(x)| = \left| \sum_{y \in \mathcal{X}} [P^t(x,y) f(y) - \pi(y) f(y)] \right| \leq \| f \|_\infty 2d(t)$$

Select proper x, which makes $|f(x)| = \| f \|_\infty$. This makes $|\lambda|^t \leq 2d(t)$.

$$t_{\text{mix}}(\varepsilon)\left(\frac{1}{|\lambda|} - 1\right) \geq t_{\text{mix}}(\varepsilon) \log\left(\frac{1}{|\lambda|}\right) \geq \log\left(\frac{1}{2\varepsilon}\right)$$

$\square$

Theorem

Suppose that $\lambda \neq 1$ is an eigenvalue for the transition matrix P of an irreducible and aperiodic Markov chain. Then

$$t_{\text{mix}}(\varepsilon) \geq \left(\frac{1}{1-|\lambda|} - 1\right) \log\left(\frac{1}{2\varepsilon}\right)$$

In particular, for reversible chains,

$$t_{\text{mix}}(\varepsilon) \geq (t_{\text{rel}} - 1) \log\left(\frac{1}{2\varepsilon}\right)$$

Proof.

First note that $\mathbb{E}_\pi[f_i] = \pi f_i = \pi P f_i = \lambda_i \pi f_i$, so if $\lambda_i \neq 1$, then $\mathbb{E}_\pi[f_i] = 0$.

$$|\lambda^t f(x)| = |P^t f(x)| = \left| \sum_{y \in \mathcal{X}} [P^t(x,y) f(y) - \pi(y) f(y)] \right| \leq \| f \|_\infty 2d(t)$$

Select proper x, which makes $|f(x)| = \| f \|_\infty$. This makes $|\lambda|^{t_{\text{mix}}(\varepsilon)} \leq 2\varepsilon$.

$$t_{\text{mix}}(\varepsilon)\left(\frac{1}{|\lambda|} - 1\right) \geq t_{\text{mix}}(\varepsilon) \log\left(\frac{1}{|\lambda|}\right) \geq \log\left(\frac{1}{2\varepsilon}\right)$$

$\square$

Definition (Dirichlet Form)

$$\mathcal{E}(f,h) := \langle (I-P)f, h\rangle_\pi$$

Definition (Dirichlet Form)

$$\mathcal{E}(f,h) := \langle (I-P)f, h\rangle_\pi$$

Lemma

$$\mathcal{E}(f) := \frac{1}{2}\sum_{x,y\in\mathcal{X}} [f(x)-f(y)]^2\pi(x)P(x,y) = \mathcal{E}(f,f)$$

Definition (Dirichlet Form)

$$\mathcal{E}(f,h) := \langle (I-P)f, h\rangle_\pi$$

Lemma

$$\mathcal{E}(f) := \frac{1}{2}\sum_{x,y\in\mathcal{X}} [f(x)-f(y)]^2 \pi(x)P(x,y) = \mathcal{E}(f,f)$$

Proof.

$$\begin{aligned}\mathcal{E}(f) &= \frac{1}{2}\sum_{x,y\in\mathcal{X}} f(x)^2\pi(x)P(x,y) - \sum_{x,y\in\mathcal{X}} f(x)f(y)\pi(x)P(x,y) \\ &+\frac{1}{2}\sum_{x,y\in\mathcal{X}} f(y)^2\pi(x)P(x,y)\end{aligned}$$

$\square$

Definition (Dirichlet Form)

$$\mathcal{E}(f,h) := \langle (I-P)f, h\rangle_\pi$$

Lemma

$$\mathcal{E}(f) := \frac{1}{2}\sum_{x,y\in\mathcal{X}} [f(x)-f(y)]^2\pi(x)P(x,y) = \mathcal{E}(f,f)$$

Proof.

$$\begin{aligned}\mathcal{E}(f) &= \frac{1}{2}\sum_{x\in\mathcal{X}} f(x)^2\pi(x) - \sum_{x,y\in\mathcal{X}} f(x)f(y)\pi(x)P(x,y)\\ &+\frac{1}{2}\sum_{y\in\mathcal{X}} f(y)^2\pi(y)\end{aligned}$$

$\square$

Definition (Dirichlet Form)

$$\mathcal{E}(f,h) := \langle (I-P)f, h\rangle_\pi$$

Lemma

$$\mathcal{E}(f) := \frac{1}{2}\sum_{x,y\in\mathcal{X}} [f(x)-f(y)]^2\pi(x)P(x,y) = \mathcal{E}(f,f)$$

Proof.

$$\mathcal{E}(f) = \sum_{x\in\mathcal{X}} f(x)^2\pi(x) - \sum_{x,y\in\mathcal{X}} f(x)f(y)\pi(x)P(x,y)$$

$\square$

Definition (Dirichlet Form)

$$\mathcal{E}(f,h) := \langle (I-P)f, h\rangle_\pi$$

Lemma

$$\mathcal{E}(f) := \frac{1}{2}\sum_{x,y\in\mathcal{X}} [f(x)-f(y)]^2\pi(x)P(x,y) = \mathcal{E}(f,f)$$

Proof.

$$\begin{aligned}\mathcal{E}(f) &= \sum_{x\in\mathcal{X}} f(x)^2\pi(x) - \sum_{x,y\in\mathcal{X}} f(x)f(y)\pi(x)P(x,y)\\ &= \langle f,f\rangle_\pi - \langle f, Pf\rangle_\pi\end{aligned}$$

□

Definition (Dirichlet Form)

$$\mathcal{E}(f,h) := \langle (I-P)f, h\rangle_\pi$$

Lemma

$$\mathcal{E}(f) := \frac{1}{2}\sum_{x,y\in\mathcal{X}} [f(x)-f(y)]^2\pi(x)P(x,y) = \mathcal{E}(f,f)$$

Lemma

Let P be the transition matrix for a reversible Makrov chain. The spectral gap $\gamma = 1-\lambda_2$ satisfies

$$\gamma = \min_{\substack{f:\mathcal{X}\to\mathbb{R}\\ f\perp_\pi \mathbb{1}, \|f\|_2=1}} \mathcal{E}(f) = \min_{\substack{f:\mathcal{X}\to\mathbb{R}\\ f\perp_\pi \mathbb{1}, f\neq 0}} \frac{\mathcal{E}(f)}{\| f \|_2^2}$$

Definition (Dirichlet Form)

$$\mathcal{E}(f,h) := \langle (I-P)f, h\rangle_\pi$$

Lemma

$$\mathcal{E}(f) := \frac{1}{2}\sum_{x,y\in\mathcal{X}} [f(x)-f(y)]^2\pi(x)P(x,y) = \mathcal{E}(f,f)$$

Lemma

Let P be the transition matrix for a reversible Makrov chain. The spectral gap $\gamma = 1-\lambda_2$ satisfies

$$\gamma = \min_{\substack{f:\mathcal{X}\to\mathbb{R}\\ f\perp_\pi \mathbb{1}, \|f\|_2=1}} \mathcal{E}(f) = \min_{\substack{f:\mathcal{X}\to\mathbb{R}\\ f\perp_\pi \mathbb{1}, f\neq 0}} \frac{\mathcal{E}(f)}{\| f \|_2^2}$$

Proof.

Let $a_j := \langle f, f_j\rangle_\pi$. Since $f \perp_\pi \mathbb{1}$, $f = \sum_{j=2}^{|\mathcal{X}|} a_j f_j$. Assume $\| f \|_2^2 = \sum_{j=2}^{|\mathcal{X}|} a_j^2 = 1$.

$$\langle (I-P)f, f\rangle_\pi = \sum_{j=2}^{|\mathcal{X}|} a_j^2(1-\lambda_j) \geq 1-\lambda_2$$

□

Definition (Dirichlet Form)

$$\mathcal{E}(f,h) := \langle (I-P)f, h\rangle_\pi$$

Lemma

$$\mathcal{E}(f) := \frac{1}{2}\sum_{x,y\in\mathcal{X}} [f(x)-f(y)]^2 \pi(x)P(x,y) = \mathcal{E}(f,f)$$

Lemma

Let P be the transition matrix for a reversible Makrov chain. The spectral gap $\gamma = 1-\lambda_2$ satisfies

$$\gamma = \min_{\substack{f:\mathcal{X}\to\mathbb{R}\\ f\perp_\pi \mathbb{1}, \|f\|_2=1}} \mathcal{E}(f) = \min_{\substack{f:\mathcal{X}\to\mathbb{R}\\ f\perp_\pi \mathbb{1}, f\neq 0}} \frac{\mathcal{E}(f)}{\| f \|_2^2}$$

Example

Observe that $\mathcal{E}(f) = \mathcal{E}(f+c)$. So if $f : \mathcal{X} \to \mathbb{R}$ is a non-constant function, then

$$\gamma = \min_{\substack{f:\mathcal{X}\to\mathbb{R}\\ f\perp_\pi \mathbb{1}, f\neq 0}} \frac{\mathcal{E}(f-\mathbb{E}_\pi[f])}{\| f-\mathbb{E}_\pi[f] \|_2^2} = \min_{\substack{f:\mathcal{X}\to\mathbb{R}\\ f\perp_\pi \mathbb{1}, f\neq 0}} \frac{\mathcal{E}(f)}{\mathrm{Var}_\pi(f)}$$

Definition (Dirichlet Form)

$$\mathcal{E}(f,h) := \langle (I-P)f, h\rangle_\pi$$

Lemma

$$\mathcal{E}(f) := \frac{1}{2}\sum_{x,y\in\mathcal{X}} [f(x)-f(y)]^2\pi(x)P(x,y) = \mathcal{E}(f,f)$$

Lemma

Let P be the transition matrix for a reversible Makrov chain. The spectral gap $\gamma = 1-\lambda_2$ satisfies

$$\gamma = \min_{\substack{f:\mathcal{X}\to\mathbb{R} \\ f\perp_\pi \mathbb{1}, \|f\|_2=1}} \mathcal{E}(f) = \min_{\substack{f:\mathcal{X}\to\mathbb{R} \\ f\perp_\pi \mathbb{1}, f\not\equiv 0}} \frac{\mathcal{E}(f)}{\| f \|_2^2}$$

Example

Observe that $\mathcal{E}(f) = \mathcal{E}(f+c)$. So if $f:\mathcal{X}\to\mathbb{R}$ is a non-constant function, then

$$\gamma = \min_{\substack{f:\mathcal{X}\to\mathbb{R} \\ f\perp_\pi \mathbb{1}, f\not\equiv 0}} \frac{\mathcal{E}(f-\mathbb{E}_\pi[f])}{\| f-\mathbb{E}_\pi[f] \|_2^2} = \min_{\substack{f:\mathcal{X}\to\mathbb{R} \\ \mathbb{E}_\pi[f]=0, f\not\equiv 0}} \frac{\mathcal{E}(f)}{\mathrm{Var}_\pi(f)}$$

Definition (edge measure)

$$Q(x,y) := \pi(x)P(x,y), \qquad Q(A,B) := \sum_{x\in A, y\in B} Q(x,y)$$

Definition (conductance)

$$\Phi(S) := \frac{Q(S,S^c)}{\pi(S)}, \qquad \Phi_\star := \min_{S:\pi(S)\le\frac{1}{2}} \Phi(S)$$

Example

Conductance

Theorem

Let λ_2 *be the second largest eigenvalue of a reversible transition matrix* P*, and let* $\gamma = 1 - \lambda_2$*. Then*

$$\frac{\Phi_\star^2}{2} \leq \gamma \leq 2\Phi_\star$$

Conductance

Theorem

Let λ_2 be the second largest eigenvalue of a reversible transition matrix P, and let $\gamma = 1 - \lambda_2$. Then

$$\frac{\Phi_\star^2}{2} \leq \gamma \leq 2\Phi_\star$$

Proof of $\gamma \leq 2\Phi_\star$.

$$\gamma = \min_{\mathbb{E}_\pi[f]=0, f\neq 0} \frac{\mathcal{E}(f)}{\mathrm{Var}_\pi(f)} = \min_{\mathbb{E}_\pi[f]=0, f\neq 0} \frac{\sum_{x,y\in\mathcal{X}} \pi(x)P(x,y)[f(x)-f(y)]^2}{\sum_{x,y\in\mathcal{X}} \pi(x)\pi(y)[f(x)-f(y)]^2}$$

For any S with $\pi(S) \leq \frac{1}{2}$ define the function

$$f_S(x) = \begin{cases} -\pi(S^c) & x \in S \\ \pi(S) & x \notin S \end{cases}$$

Then

$$\gamma \leq \frac{2Q(S,S^c)}{2\pi(S)\pi(S^c)} \leq \frac{2Q(S,S^c)}{\pi(S)} \leq 2\Phi(S)$$

□

Conductance

Theorem

Let λ_2 be the second largest eigenvalue of a reversible transition matrix P, and let $\gamma = 1 - \lambda_2$. Then

$$\frac{\Phi_\star^2}{2} \leq \gamma \leq 2\Phi_\star$$

Proof of $\Phi_\star^2 \leq 2\gamma$.

$\square$

Product Chain

Consider d Markov chains, the j-th of them is $(P_j, \pi_j, \mathcal{X}_j)$. How to merge them?

$$(f^{(1)} \otimes f^{(2)} \otimes \cdots \otimes f^{(d)})(x_1, \cdots, x_d) := f^{(1)}(x_1) f^{(2)}(x_2) \cdots f^{(d)}(x_d)$$

Consider this transition function:

$$\tilde{P}(\mathbb{x}, \mathbb{y}) = \sum_{j=1}^{d} w_j P_j(x_j, y_j) \prod_{i: i \neq j} [x_i = y_i], \qquad \mathbb{x} = (x_1, x_2, \cdots, x_d)$$

This product chain has the following properties:

Suppose that for each $j = 1, 2, \cdots, d$, the transition matrix P_j on state space $\mathcal{X}_j$ has eigenfunction $\varphi^{(j)}$ with eigenvalue $\lambda^{(j)}$. Let w be a probability distribution on $\{1, \cdots, d\}$, then:

Product Chain

Consider d Markov chains, the j-th of them is $(P_j, \pi_j, \mathcal{X}_j)$. How to merge them?

$$(f^{(1)} \otimes f^{(2)} \otimes \cdots \otimes f^{(d)})(x_1, \cdots, x_d) := f^{(1)}(x_1) f^{(2)}(x_2) \cdots f^{(d)}(x_d)$$

Consider this transition function:

$$\tilde{P}(\mathbb{x}, \mathbb{y}) = \sum_{j=1}^{d} w_j P_j(x_j, y_j) \prod_{i:i\neq j} [x_i = y_i], \qquad \mathbb{x} = (x_1, x_2, \cdots, x_d)$$

This product chain has the following properties:

Suppose that for each $j = 1, 2, \cdots, d$, the transition matrix P_j on state space $\mathcal{X}_j$ has eigenfunction $\varphi^{(j)}$ with eigenvalue $\lambda^{(j)}$. Let w be a probability distribution on $\{1, \cdots, d\}$, then:

$\tilde{\varphi} := \varphi^{(1)} \otimes \cdots \otimes \varphi^{(d)}$ is an eigenfunction with eigenvalue $\displaystyle\sum_{j=1}^{d} w_j \lambda^{(j)}$

Proof.

Let $\tilde{P}_j(\mathbb{x}, \mathbb{y}) = P_j(x_j, y_j) \displaystyle\prod_{i:i\neq j} [x_i = y_i]$, then

$$\begin{aligned}
\tilde{P}_j \tilde{\varphi}(\mathbb{x}) &= \sum_{\mathbb{y}} \tilde{P}_j(\mathbb{x}, \mathbb{y}) \tilde{\varphi}(\mathbb{y}) = \sum_{\mathbb{x}'_j} \tilde{P}_j(\mathbb{x}, \mathbb{x}'_j) \tilde{\varphi}(\mathbb{x}'_j) \\
&= \sum_{x'_j} P_j(x_j, x'_j) \varphi^{(j)}(x'_j) \prod_{i\neq j} \varphi^{(i)}(x_i) = \lambda^{(j)} \varphi^{(j)}(x_j) \sum_{i\neq j} \varphi^{(i)}(x_i) \\
&= \lambda^{(j)} \tilde{\varphi}(\mathbb{x})
\end{aligned}$$

$\square$

Product Chain

Consider d Markov chains, the j-th of them is $(P_j, \pi_j, \mathcal{X}_j)$. How to merge them?

$$(f^{(1)} \otimes f^{(2)} \otimes \cdots \otimes f^{(d)})(x_1, \cdots, x_d) := f^{(1)}(x_1) f^{(2)}(x_2) \cdots f^{(d)}(x_d)$$

Consider this transition function:

$$\tilde{P}(\mathbb{x}, \mathbb{y}) = \sum_{j=1}^{d} w_j P_j(x_j, y_j) \prod_{i: i \neq j} [x_i = y_i], \qquad \mathbb{x} = (x_1, x_2, \cdots, x_d)$$

This product chain has the following properties:
Suppose that for each $j = 1, 2, \cdots, d$, the transition matrix P_j on state space $\mathcal{X}_j$ has eigenfunction $\varphi^{(j)}$ with eigenvalue $\lambda^{(j)}$. Let w be a probability distribution on $\{1, \cdots, d\}$, then:

If P_j has an eigenbasis for all j, then $\tilde{P}$ has an eigenbasis.

Proof.

Suppose $a_1, a_2 \in \mathbb{R}^{\mathcal{X}_a}, b_1, b_2 \in \mathbb{R}^{\mathcal{X}_b}$. Then

$$\begin{aligned}
\langle a_1 \otimes b_1, a_2 \otimes b_2 \rangle_{\pi_a \otimes \pi_b} &= \sum_{x \in \mathcal{X}_a, y \in \mathcal{X}_b} a_1(x) a_2(x) \pi_a(x) b_1(y) b_2(y) \pi_b(y) \\
&= \left(\sum_{x \in \mathcal{X}_a} a_1(x) a_2(x) \pi_a(x) \right) \left(\sum_{y \in \mathcal{X}_b} b_1(y) b_2(y) \pi_b(y) \right) \\
&= \langle a_1, a_2 \rangle_{\pi_a} \cdot \langle b_1, b_2 \rangle_{\pi_b}
\end{aligned}$$

$\square$

Induced Chain

Theorem

Let (X_t) be a reversible Markov chain on $\mathcal{X}$ with stationary measure π and spectral gap γ. Let $A \subset \mathcal{X}$ be non-empty and let γ_A be the spectral gap for the chain induced on A. Then $\gamma_A \geq \gamma$. ($P_A(x,y) = \Pr_x[X_{\tau_A^+} = y], \pi_A = \dfrac{\pi}{\pi(A)}$)

Induced Chain

Theorem

Let (X_t) be a reversible Markov chain on $\mathcal{X}$ with stationary measure π and spectral gap γ. Let $A \subset \mathcal{X}$ be non-empty and let γ_A be the spectral gap for the chain induced on A. Then $\gamma_A \geq \gamma$. ($P_A(x,y) = \Pr_x[X_{\tau_A^+} = y], \pi_A = \dfrac{\pi}{\pi(A)}$)

Proof.

$\pi(x)P_A(x,y) = \pi(y)P_A(y,x)$, so P_A is reversible.

So, $\exists \varphi : A \to \mathbb{R}$, such that $\gamma_A = \dfrac{\mathcal{E}(\varphi)}{\mathrm{Var}_{\pi_A}(\varphi)}$ and $\mathbb{E}_{\pi_A}(\varphi) = 0$.

Extend φ to $\phi : \mathcal{X} \to \mathbb{R}$ by $\phi(x) := \mathbb{E}_x[\varphi(X_{\tau_A})]$.

For $x \in A$: $P\phi(x) = \sum_{y \in \mathcal{X}} P(x,y)\mathbb{E}_y[\varphi(X_{\tau_A})] = \mathbb{E}_x[\varphi(X_{\tau_A^+})] = P_A\varphi(x)$

For $x \notin A$: $(I-P)\phi(x) = 0$. Since $\mathbb{E}_x[\varphi(X_{\tau_A})] = \mathbb{E}_x[\varphi(X_{\tau_A^+})], \qquad \forall x \notin A$

$$\mathcal{E}(\phi) = \langle (I-P)\phi, \phi\rangle_\pi = \sum_{x \in A} [(I-P)\phi(x)]\phi(x)\pi(x) = \pi(A)\mathcal{E}_A(\varphi)$$

$$\mathrm{Var}_\pi(\phi) \geq \sum_{x \in A} [\varphi(x) - \mathbb{E}_\pi[\phi]]^2 \pi(x) \geq \pi(A) \sum_{x \in A} \varphi^2(x)\pi_A(x) = \pi(A)\mathrm{Var}_{\pi_A}(\varphi)$$

$$\gamma \leq \frac{\mathcal{E}(\phi)}{\mathrm{Var}_\pi(\phi)} \leq \frac{\pi(A)\mathcal{E}_A(\varphi)}{\pi(A)\mathrm{Var}_{\pi_A}(\varphi)} \leq \gamma_A \qquad \square$$

Induced Chain

Theorem

Let (X_t) be a reversible Markov chain on $\mathcal{X}$ with stationary measure π and spectral gap γ. Let $A \subset \mathcal{X}$ be non-empty and let γ_A be the spectral gap for the chain induced on A. Then $\gamma_A \geq \gamma$. $(P_A(x,y) = \Pr_x[X_{\tau_A^+} = y], \pi_A = \dfrac{\pi}{\pi(A)})$

Proof.

$\pi(x)P_A(x,y) = \pi(y)P_A(y,x)$, so P_A is reversible.

So, $\exists \varphi : A \to \mathbb{R}$, such that $\gamma_A = \dfrac{\mathcal{E}(\varphi)}{\mathrm{Var}_{\pi_A}(\varphi)}$ and $\mathbb{E}_{\pi_A}(\varphi) = 0$.

Extend φ to $\phi : \mathcal{X} \to \mathbb{R}$ by $\phi(x) := \mathbb{E}_x[\varphi(X_{\tau_A})]$.

For $x \in A$: $P\phi(x) = \sum_{y \in \mathcal{X}} P(x,y)\mathbb{E}_y[\varphi(X_{\tau_A})] = \mathbb{E}_x[\varphi(X_{\tau_A^+})] = P_A\varphi(x)$

For $x \notin A$: $(I-P)\phi(x) = 0$. Since $\mathbb{E}_x[\varphi(X_{\tau_A})] = \mathbb{E}_x[\varphi(X_{\tau_A^+})]$, $\quad \forall x \notin A$

$$\mathcal{E}(\phi) = \langle (I-P)\phi, \phi\rangle_\pi = \sum_{x \in A} [(I-P_A)\varphi(x)]\phi(x)\pi(x) = \pi(A)\mathcal{E}_A(\varphi)$$

$$\mathrm{Var}_\pi(\phi) \geq \sum_{x \in A} [\varphi(x) - \mathbb{E}_\pi[\phi]]^2 \pi(x) \geq \pi(A) \sum_{x \in A} \varphi^2(x)\pi_A(x) = \pi(A)\mathrm{Var}_{\pi_A}(\varphi)$$

$$\gamma \leq \frac{\mathcal{E}(\phi)}{\mathrm{Var}_\pi(\phi)} \leq \frac{\pi(A)\mathcal{E}_A(\varphi)}{\pi(A)\mathrm{Var}_{\pi_A}(\varphi)} \leq \gamma_A$$

$\square$

Induced Chain

Theorem

Let (X_t) be a reversible Markov chain on $\mathcal{X}$ with stationary measure π and spectral gap γ. Let $A \subset \mathcal{X}$ be non-empty and let γ_A be the spectral gap for the chain induced on A. Then $\gamma_A \geq \gamma$. $(P_A(x,y) = \Pr_x[X_{\tau_A^+} = y], \pi_A = \dfrac{\pi}{\pi(A)})$

Proof.

$\pi(x)P_A(x,y) = \pi(y)P_A(y,x)$, so P_A is reversible.

So, $\exists \varphi : A \to \mathbb{R}$, such that $\gamma_A = \dfrac{\mathcal{E}(\varphi)}{\mathrm{Var}_{\pi_A}(\varphi)}$ and $\mathbb{E}_{\pi_A}(\varphi) = 0$.

Extend φ to $\phi : \mathcal{X} \to \mathbb{R}$ by $\phi(x) := \mathbb{E}_x[\varphi(X_{\tau_A})]$.

For $x \in A$: $P\phi(x) = \sum_{y \in \mathcal{X}} P(x,y)\mathbb{E}_y[\varphi(X_{\tau_A})] = \mathbb{E}_x[\varphi(X_{\tau_A^+})] = P_A\varphi(x)$

For $x \notin A$: $(I-P)\phi(x) = 0$. Since $\mathbb{E}_x[\varphi(X_{\tau_A})] = \mathbb{E}_x[\varphi(X_{\tau_A^+})], \quad \forall x \notin A$

$$\mathcal{E}(\phi) = \langle (I-P)\phi, \phi\rangle_\pi = \sum_{x \in A} [(I-P_A)\varphi(x)]\varphi(x)\pi(x) = \pi(A)\mathcal{E}_A(\varphi)$$

$$\mathrm{Var}_\pi(\phi) \geq \sum_{x \in A} [\varphi(x) - \mathbb{E}_\pi[\phi]]^2 \pi(x) \geq \pi(A) \sum_{x \in A} \varphi^2(x)\pi_A(x) = \pi(A)\mathrm{Var}_{\pi_A}(\varphi)$$

$$\gamma \leq \frac{\mathcal{E}(\phi)}{\mathrm{Var}_\pi(\phi)} \leq \frac{\pi(A)\mathcal{E}_A(\varphi)}{\pi(A)\mathrm{Var}_{\pi_A}(\varphi)} \leq \gamma_A \qquad \square$$

Comapre Dirichlet Form

Theorem

Let P and $\tilde{P}$ be reversible transition matrices with stationary distributions π and $\tilde{\pi}$, respectively. If $\tilde{\mathcal{E}}(f) \le \alpha \mathcal{E}(f)$ for all f, then

$$\tilde{\gamma} \le \left[\max_{x \in \mathcal{X}} \frac{\pi(x)}{\tilde{\pi}(x)} \right] \alpha \gamma$$

Comapre Dirichlet Form

Theorem

Let P and $\tilde{P}$ be reversible transition matrices with stationary distributions π and $\tilde{\pi}$, respectively. If $\tilde{\mathcal{E}}(f) \leq \alpha \mathcal{E}(f)$ for all f, then

$$\tilde{\gamma} \leq \left[\max_{x \in \mathcal{X}} \frac{\pi(x)}{\tilde{\pi}(x)} \right] \alpha \gamma$$

Proof.

Target: $\dfrac{1}{\mathrm{Var}_{\tilde{\pi}}(f)} \leq c \cdot \dfrac{1}{\mathrm{Var}_{\pi}(f)}$

$$\mathrm{Var}_{\pi}(f) \leq \mathbb{E}_{\pi}[(f - \mathbb{E}_{\tilde{\pi}}(f))^2] = \sum_{x \in \mathcal{X}} [f(x) - \mathbb{E}_{\tilde{\pi}}(f)]^2 \pi(x) \leq c \cdot \underbrace{\sum_{x \in \mathcal{X}} [f(x) - \mathbb{E}_{\tilde{\pi}}(f)]^2 \tilde{\pi}(x)}_{\mathrm{Var}_{\tilde{\pi}}(f)}$$

□

Path Method

Let $(P,\pi),(\tilde{P},\tilde{\pi})$ be two reversible chain. Let $E := \{(x,y) : P(x,y) > 0\}$.
For each $(x,y) \in \tilde{E}$, choose a path $\Gamma_{xy} = \{(x,x_1),(x_1,x_2),\cdots,(x_{k-1},y)\}$ from E.

Definition (Congestion ratio)

$$B := \max_{e\in E}\left(\frac{1}{Q(e)}\sum_{\substack{x,y\\ \Gamma_{xy}\ni e}}\tilde{Q}(x,y)|\Gamma_{xy}|\right)$$

Path Method

Let $(P, \pi), (\tilde{P}, \tilde{\pi})$ be two reversible chain. Let $E := \{(x, y) : P(x, y) > 0\}$.
For each $(x, y) \in \tilde{E}$, choose a path $\Gamma_{xy} = \{(x, x_1), (x_1, x_2), \cdots, (x_{k-1}, y)\}$ from E.

Definition (Congestion ratio)

$$B := \max_{e \in E} \left(\frac{1}{Q(e)} \sum_{\substack{x,y \\ \Gamma_{xy} \ni e}} \tilde{Q}(x, y) |\Gamma_{xy}| \right)$$

Theorem

$$\tilde{\mathcal{E}}(f) \leq B \mathcal{E}(f)$$

Path Method

Let $(P, \pi), (\tilde{P}, \tilde{\pi})$ be two reversible chain. Let $E := \{(x, y) : P(x, y) > 0\}$.
For each $(x, y) \in \tilde{E}$, choose a path $\Gamma_{xy} = \{(x, x_1), (x_1, x_2), \cdots, (x_{k-1}, y)\}$ from E.

Definition (Congestion ratio)

$$B := \max_{e \in E} \left(\frac{1}{Q(e)} \sum_{\substack{x,y \\ \Gamma_{xy} \ni e}} \tilde{Q}(x, y) |\Gamma_{xy}| \right)$$

Theorem

$$\tilde{\mathcal{E}}(f) \leq B \mathcal{E}(f)$$

Proof.

Let $\nabla f(e = \{z, w\}) := f(w) - f(z)$ (directed)

$$\begin{aligned} 2\tilde{\mathcal{E}}(f) &= \sum_{(x,y) \in \tilde{E}} \tilde{Q}(x, y)[f(x) - f(y)]^2 = \sum_{(x,y) \in \tilde{E}} \tilde{Q}(x, y) \left[\sum_{e \in \Gamma_{xy}} \nabla f(e) \right]^2 \\ &\leq \sum_{x,y} \tilde{Q}(x, y) |\Gamma_{xy}| \sum_{e \in \Gamma_{xy}} [\nabla f(e)]^2 \qquad \text{(Cauchy-Schwarz)} \\ &= \sum_{e \in E} \left[\sum_{\Gamma_{xy} \ni e} \tilde{Q}(x, y) |\Gamma_{xy}| \right] [\nabla f(e)]^2 \end{aligned}$$

$\square$

Path Method

Let $(P,\pi), (\tilde{P},\tilde{\pi})$ be two reversible chain. Let $E := \{(x,y) : P(x,y) > 0\}$.
For each $(x,y) \in \tilde{E}$, choose a path $\Gamma_{xy} = \{(x,x_1),(x_1,x_2),\cdots,(x_{k-1},y)\}$ from E.

Definition (Congestion ratio)

$$B := \max_{e\in E}\left(\frac{1}{Q(e)}\sum_{\substack{x,y\\ \Gamma_{xy}\ni e}}\tilde{Q}(x,y)|\Gamma_{xy}|\right)$$

Theorem

$$\tilde{\mathcal{E}}(f) \leq B\mathcal{E}(f)$$

Proof.

Let $\nabla f(e=\{z,w\}) := f(w) - f(z)$ (directed)

$$\begin{aligned}
2\tilde{\mathcal{E}}(f) &= \sum_{(x,y)\in\tilde{E}}\tilde{Q}(x,y)[f(x)-f(y)]^2 = \sum_{(x,y)\in\tilde{E}}\tilde{Q}(x,y)\left[\sum_{e\in\Gamma_{xy}}\nabla f(e)\right]^2 \\
&\leq \sum_{x,y}\tilde{Q}(x,y)|\Gamma_{xy}|\sum_{e\in\Gamma_{xy}}[\nabla f(e)]^2 \qquad \text{(Cauchy-Schwarz)} \\
&= \sum_{e\in E}\left[\frac{1}{Q(e)}\sum_{\Gamma_{xy}\ni e}\tilde{Q}(x,y)|\Gamma_{xy}|\right]Q(e)[\nabla f(e)]^2
\end{aligned}$$

$\square$

Path Method

Let $(P,\pi),(\tilde{P},\tilde{\pi})$ be two reversible chain. Let $E := \{(x,y) : P(x,y) > 0\}$.
For each $(x,y) \in \tilde{E}$, choose a path $\Gamma_{xy} = \{(x,x_1),(x_1,x_2),\cdots,(x_{k-1},y)\}$ from E.

Definition (Congestion ratio)

$$B := \max_{e \in E} \left(\frac{1}{Q(e)} \sum_{\substack{x,y \\ \Gamma_{xy} \ni e}} \tilde{Q}(x,y)|\Gamma_{xy}| \right)$$

Theorem

$$\tilde{\mathcal{E}}(f) \leq B\mathcal{E}(f)$$

Proof.

Let $\nabla f(e = \{z,w\}) := f(w) - f(z)$ (directed)

$$\begin{aligned} 2\tilde{\mathcal{E}}(f) &= \sum_{(x,y)\in\tilde{E}} \tilde{Q}(x,y)[f(x)-f(y)]^2 = \sum_{(x,y)\in\tilde{E}} \tilde{Q}(x,y)\left[\sum_{e\in\Gamma_{xy}} \nabla f(e)\right]^2 \\ &\leq \sum_{x,y} \tilde{Q}(x,y)|\Gamma_{xy}| \sum_{e\in\Gamma_{xy}} [\nabla f(e)]^2 \qquad \text{(Cauchy-Schwarz)} \\ &= \sum_{e\in E}\left[\frac{1}{Q(e)} \sum_{\Gamma_{xy}\ni e} \tilde{Q}(x,y)|\Gamma_{xy}|\right] Q(e)[\nabla f(e)]^2 \leq 2B\mathcal{E}(f) \end{aligned}$$

$\square$

Path Method

Let $(P,\pi), (\tilde{P},\tilde{\pi})$ be two reversible chain. Let $E := \{(x,y) : P(x,y) > 0\}$.
For each $(x,y) \in \tilde{E}$, choose a path $\Gamma_{xy} = \{(x,x_1),(x_1,x_2),\cdots,(x_{k-1},y)\}$ from E.

Definition (Congestion ratio)

$$B := \max_{e\in E}\left(\frac{1}{Q(e)}\sum_{\substack{x,y\\ \Gamma_{xy}\ni e}}\tilde{Q}(x,y)\sum_{\Gamma:e\in\Gamma\in\mathcal{P}_{xy}}\nu_{xy}(\Gamma)|\Gamma|\right)$$

Theorem

$$\tilde{\mathcal{E}}(f) \le B\mathcal{E}(f)$$

Proof.

Let $\nabla f(e=\{z,w\}) := f(w) - f(z)$ (directed)

$$\begin{aligned}
2\tilde{\mathcal{E}}(f) &= \sum_{(x,y)\in\tilde{E}}\tilde{Q}(x,y)[f(x)-f(y)]^2 = \sum_{(x,y)\in\tilde{E}}\tilde{Q}(x,y)\left[\sum_{\Gamma\in\mathcal{P}_{xy}}\nu_{xy}(\Gamma)(f(x)-f(y))\right]^2 \\
&\le \sum_{(x,y)\in\tilde{E}}\tilde{Q}(x,y)\sum_{\Gamma\in\mathcal{P}_{xy}}\nu_{xy}(\Gamma)(f(x)-f(y))^2 \\
&= \sum_{(x,y)\in\tilde{E}}\tilde{Q}(x,y)\sum_{\Gamma\in\mathcal{P}_{xy}}\nu_{xy}(\Gamma)\left[\sum_{e\in\Gamma}\nabla f(e)\right]^2
\end{aligned}$$

$\square$