Orthogonal polynomials and mixing times

Yevgeniy Kovchegov
Oregon State University

Markov chains over discrete state space: mixing times

S - sample space

$P = \{p(i,j)\}_{i,j \in S}$ - transition probabilities

X_t distributed according to $\mu_t = \mu_0 P^t$

π - stationary distribution

Total variation distance:

$$\|\mu - \nu\|_{TV} := \frac{1}{2} \sum_{x \in S} |\mu(x) - \nu(x)| = \sup_{A \subset S} |\mu(A) - \nu(A)|$$

Mixing time:

$$t_{mix}(\varepsilon) := \inf \{t \ : \ \|\mu_t - \pi\|_{TV} \le \varepsilon, \ \text{ all } \mu_0\}$$

Note: other norms can be used.

Coupling method.

Construct process $\begin{pmatrix} X_t \\ Y_t \end{pmatrix}$ on $S \times S$ such that

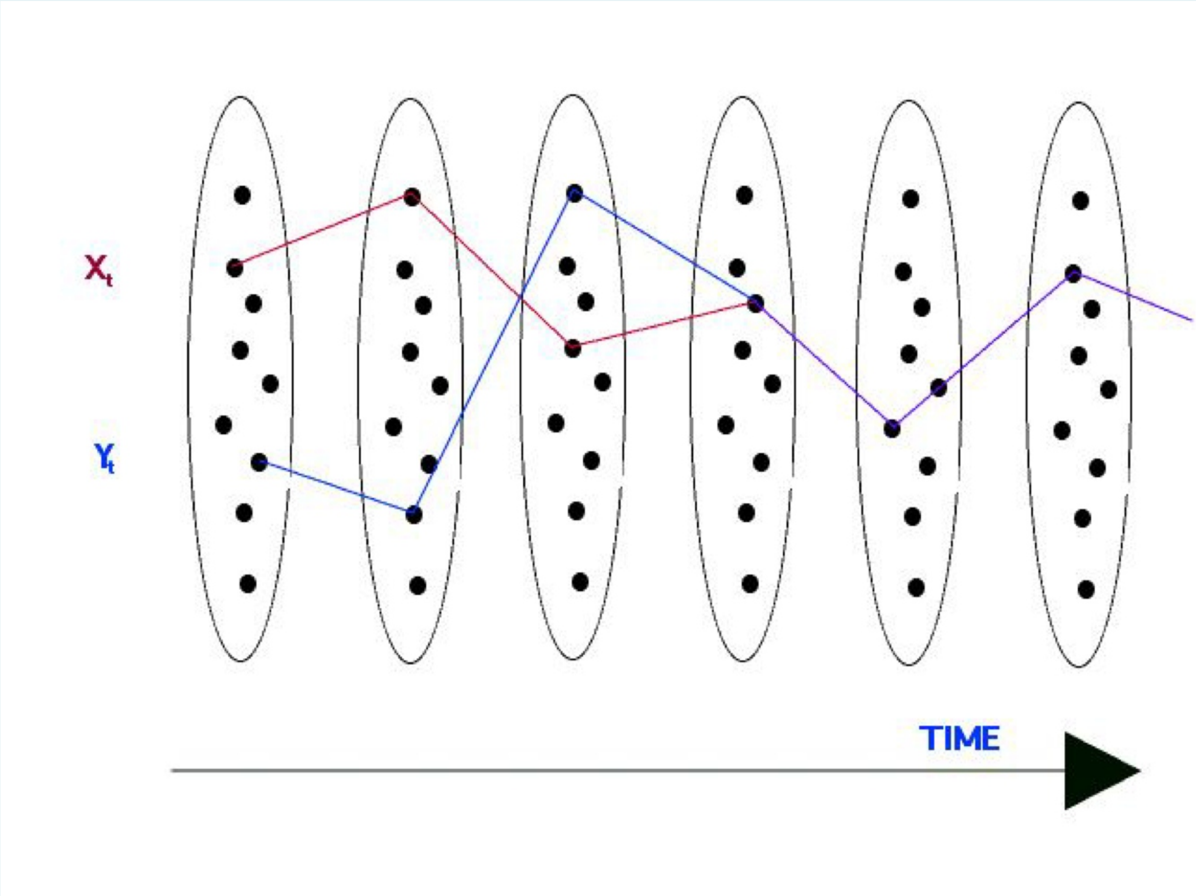
X_t is a $\{p(i,j)\}$-Markov chain
Y_t is a $\{p(i,j)\}$-Markov chain

Once $X_t{=}Y_t$, let $X_{t+1}{=}Y_{t+1}$, $X_{t+2}{=}Y_{t+2}, \ldots$

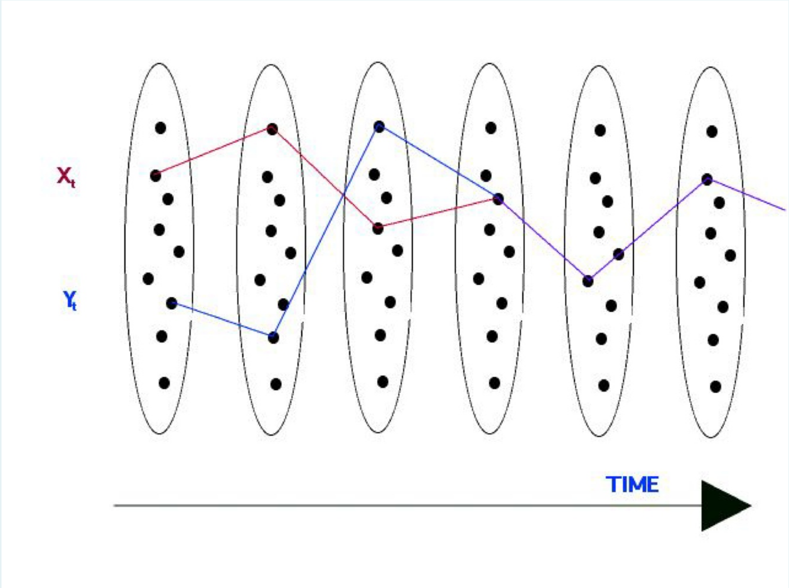
Coupling time: $T_{coupling} = \min\{t : \; X_t = Y_t\}$

Successful coupling: $\mathbf{Prob}(T_{coupling} < \infty) = 1$

Coupling method.



Coupling method.



$$O(t_{mix}) \leq O(T_{coupling})$$

by the coupling inequality. Thus constructing a coupled process that minimizes $E[T_{coupling}]$ gives an effective upper bound on mixing time.

Coupling inequality.

Given $X_0 = i$ and $Y_0 = j$. Then

$$\|P_{X_t} - P_{Y_t}\|_{TV} \leq P_{i,j}[T_{coupling} > t]$$

If $X_0 \sim \mu_0$ and $Y_0 \sim \pi$,

$$\|\mu_t - \pi\|_{TV} = \|P_{X_t} - P_{Y_t}\|_{TV} \leq \frac{\max_{i,j \in S} E_{i,j}[T_{coupling}]}{t} \leq \varepsilon$$

whenever $t \geq \frac{\max_{i,j \in S} E[T_{coupling}]}{\varepsilon}$, and

$$O(t_{mix}) \leq O(T_{coupling})$$

Example: “lazy” random walk on $\mathbb{Z}/n\mathbb{Z}$

$p(j, j+1) = 1/4$, $p(j, j-1) = 1/4$ and $p(j, j) = 1/2$

A simple coupling produces order $O(n^2)$ upper bound for mixing time, which must be correct due to CLT.

Shuffling by random transpositions.

Pick two cards at random:

1 2 3 4 **5** 6 **7** 8

Transpose them:

1 2 3 4 **7** 6 **5** 8

Iterate:

3 2 **1** 4 **7** 6 **5** 8

3 2 **6** 4 **7** **1** **5** 8

. . . etc.

Shuffling by random transpositions.

Diaconis and Shahshahani (early 80's): The mixing time for shuffling a deck of n cards by random transpositions is of order $O(n \log(n))$ with cut-off asymptotics at $\frac{1}{2} n \log(n)$.

A coupling. (in Aldous and Fill) Move card $\boxed{\mathbf{a}}$ to location i in both processes (decks), X_t and Y_t.

Shuffling by random transpositions.

In case of **two discrepancies** ($d = 2$) at d_1 and d_2:

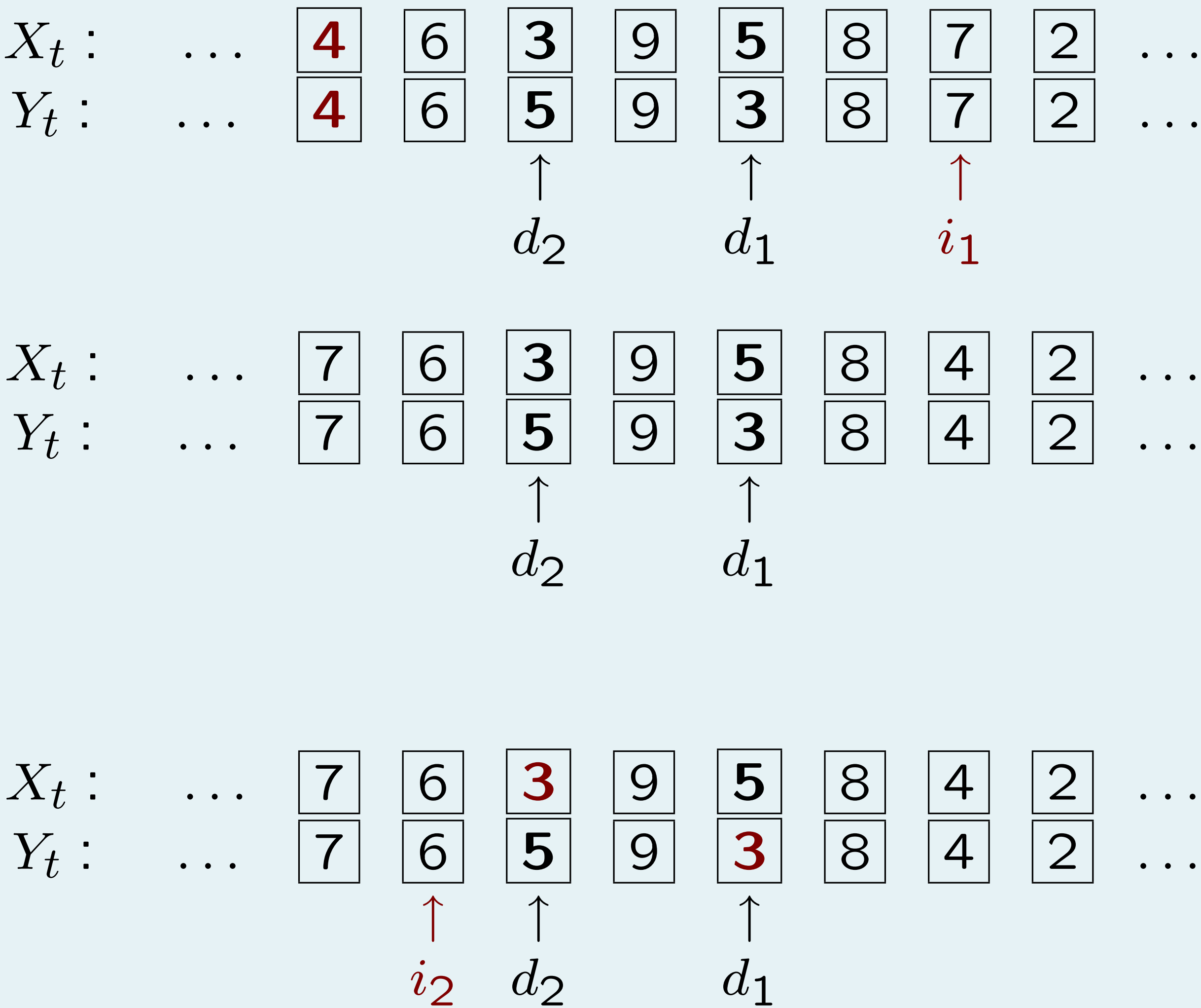
X_t :	...	4	6	**b**	9	**a**	8	7	2	...
Y_t :	...	4	6	**a**	9	**b**	8	7	2	...
				$\uparrow$		$\uparrow$				
				d_2		d_1				

Label-to-location coupling:

$$E[T_{coupling}] = \frac{n^2}{4} \qquad \text{– too large.}$$

Case of **two discrepancies** ($d = 2$):

X_t :	...	4	6	**3**	9	**5**	8	7	2	...
Y_t :	...	4	6	**5**	9	**3**	8	7	2	...
				$\uparrow$		$\uparrow$				
				d_2		d_1				



Xt : … 4 6 3 9 5 8 7 2 …
Yt : … 4 6 5 9 3 8 7 2 …
d2 d1 i1
Xt : … 7 6 3 9 5 8 4 2 …
Yt : … 7 6 5 9 3 8 4 2 …
d2 d1
Xt : … 7 6 3 9 5 8 4 2 …
Yt : … 7 6 5 9 3 8 4 2 …
i2 d2 d1

X_t : … 7 3 **6** 9 **5** 8 4 2 …

Y_t : … 7 3 **5** 9 **6** 8 4 2 …

$\uparrow d_2$ (under **6** / **5**) $\uparrow d_1$ (under **5** / **6**)

X_t : … 7 3 **6** 9 **5** 8 4 2 …

Y_t : … 7 3 **5** 9 **6** 8 4 2 …

$\uparrow i_3$ (under **5** / **6**)

X_t : … 7 3 **6** 9 **5** 8 4 2 …

Y_t : … 7 3 **6** 9 **5** 8 4 2 …

Label-to-location coupling: $E[T_{coupling}] = \frac{n^2}{4}$

Here $E[T_{coupling}] \approx \sum_{d=2}^{n} \frac{n^2}{d^2} \approx \left(\frac{\pi^2}{6} - 1\right) n^2$

YK and R.Burton '07: $O(n \log(n))$ coupling.

General state space: mixing rates.

Mixing time for Markov chains over **finite state spaces:**

$$t_{mix}(\varepsilon) := \text{inf}\,\{t \;:\; \|\mu_t - \pi\|_{TV} \leq \varepsilon, \quad \text{all } \mu_0\}$$

The above definition is invalid for Markov chains over **general countably infinite state spaces**. There all depends on μ_0.

Given μ_0, $\quad t_{mix}(\varepsilon) := \text{inf}\left\{t \;:\; \|\mu_0 P^t - \pi\|_{TV} \leq \varepsilon\right\}$

Toy example.

Consider a positive recurrent nearest neighbor walk

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ q & r & p & 0 & \dots \\ 0 & q & r & p & \ddots \\ 0 & 0 & q & r & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} \qquad q > p, \quad r > 0$$

that begins at the origin: $\mu_0 = e_0^T$.

We want to find the asymptotics of $t_{mix}(\varepsilon)$ as $\varepsilon \downarrow 0$.

Method of Karlin and McGregor

Consider nearest-neighbor walk in 1D:

$$P = \begin{pmatrix} r_0 & p_0 & 0 & 0 & \cdots \\ q_1 & r_1 & p_1 & 0 & \cdots \\ 0 & q_2 & r_2 & p_2 & \ddots \\ 0 & 0 & q_3 & r_3 & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} \quad r_j = 1 - p_j - q_j$$

$\mathbf{q}(\lambda) = (Q_0, Q_1, Q_2, \ldots)^T$ solves right eigenvalue problem, i.e. $Q_0 = 1$, $\lambda \mathbf{q} = P\mathbf{q}$

$$\lambda Q_j(\lambda) = q_j Q_{j-1}(\lambda) + r_j Q_j(\lambda) + p_j Q_{j+1}(\lambda)$$

Method of Karlin and McGregor

Reversible π: $\pi_0 = 1$, $\pi_j = \frac{p_0 \cdots p_{j-1}}{q_1 \cdots q_j}$ satisfies

$$\pi_j p_j = \pi_{j+1} q_{j+1}$$

If $\rho = \sum_{k=0}^{\infty} \pi_k < \infty$, then $\nu = \frac{1}{\rho}\pi$ is the stationary probability distribution.

The operator P is self-adjoint in $\ell^2(\pi)$ given by the inner product

$$< f, g >_\pi = \sum_j f(j) g(j) \pi(j)$$

$$[P]_\pi = \begin{pmatrix} r_0 & b_0 & 0 & \dots \\ b_0 & r_1 & b_1 & \ddots \\ 0 & b_1 & r_2 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

Method of Karlin and McGregor

Spectrums of P and $[P]_\pi$ coincide $\Rightarrow$ eigenvalues are **real**.

Also $\mathbf{u}(\lambda) = (\pi_0 Q_0, \pi_1 Q_1, \ldots)^T$ solves the left eigenvalue problem, i.e.

$$\mathbf{u}^T P = \lambda P$$

as

$$\lambda \pi_n Q_n = q_n \pi_n Q_{n-1} + r_n \pi_n Q_n + p_n \pi_n Q_{n+1}$$

$$= p_{n-1} \pi_{n-1} Q_{n-1} + r_n \pi_n Q_n + q_{n+1} \pi_{n+1} Q_{n+1}$$

Diagonalization

If A is an $n \times n$ Hermitian matrix with simple real EVs $\lambda_1, \ldots, \lambda_n$, and if $u_1, \ldots, u_n$ and $v_1, \ldots, v_n$ are the left and right eigenvectors. Then

$$A^t = \sum_j \frac{\lambda^t v_j u_j^T}{u_j^T v_j} = \int_{\sigma(A)} \lambda^t v(\lambda) u^T(\lambda) d\psi(\lambda) \ ,$$

where $u(\lambda_j) = u_j$, $v(\lambda_j) = v_j$, and

$$\psi(\lambda) = \sum_j \frac{1}{u^T(\lambda) v(\lambda)} \delta_{\lambda_j}(\lambda) = \frac{n}{u^T(\lambda) v(\lambda)} U_{\sigma(A)}(\lambda)$$

Method of Karlin and McGregor

$$P^t = \int \lambda^t \mathbf{q}(\lambda)\mathbf{u}^T(\lambda)d\psi(\lambda)$$

$$= \int \lambda^t \begin{pmatrix} \pi_0 Q_0 Q_0 & \pi_1 Q_0 Q_1 & \cdots \\ \pi_0 Q_1 Q_0 & \pi_1 Q_1 Q_1 & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} d\psi(\lambda)$$

i.e.

$$p_t(i,j) = \pi_j \int_{-1}^{+1} \lambda^t Q_i(\lambda) Q_j(\lambda) d\psi(\lambda)$$

Applications: mixing rates

$$p_t(i,j) = \pi_j \int_{-1}^{+1} \lambda^t Q_i(\lambda) Q_j(\lambda) d\psi(\lambda)$$

Since $\rho = \sum_{k=0}^{\infty} \pi_k < \infty$ and $\nu = \frac{1}{\rho}\pi$, for an aperiodic nearest neighbor walk originating at site i

$$\|\nu - \mu_t\|_{TV} = \frac{1}{2} \sum_j \pi_j \left| \int_{(-1,1)} \lambda^t Q_i(\lambda) Q_j(\lambda) d\psi(\lambda) \right|$$

as measure ψ contains a point mass of weight $\frac{1}{\rho}$ at $\lambda = 1$. Here aperiodic means there is no point mass at $\lambda = -1$.

Applications: Mixing Times

Consider our **toy example**: Markov chain

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ q & r & p & 0 & \dots \\ 0 & q & r & p & \ddots \\ 0 & 0 & q & r & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} \qquad q > p, \quad r > 0$$

originating from $i = 0$.

We need to find the asymptotics for the "time to stationarity"

$$\|\nu - \mu_t\|_{TV} \;=\; \frac{1}{2}\sum_j \pi_j \left| \int_{(-1,1)} \lambda^t Q_j(\lambda) d\psi(\lambda) \right|$$

Applications: Mixing Times

Consider our **toy example**: Markov chain

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ q & r & p & 0 & \dots \\ 0 & q & r & p & \ddots \\ 0 & 0 & q & r & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} \qquad q > p, \quad r > 0$$

originating from $i = 0$. There is a closed form

$$\int_{(-1,1)} \lambda^t Q_n(\lambda) d\psi(\lambda) = \frac{(1+q-p)(q+r)-q}{(1+q-p)(q+r)} \cdot \left(-\frac{q}{q+r}\right)^{t+n}$$

$$+ \left(\sqrt{\frac{q}{p}}\right)^n \left(\frac{p}{q+r}\right) \frac{1}{2\pi i} \oint_{|z|=1} \frac{(\sqrt{pq}(z+z^{-1})+r)^t z^n (z-z^{-1})}{\left(z-\sqrt{\frac{p}{q}}\frac{r+(1+q-p)}{2(q+r)}\right)\left(z-\sqrt{\frac{p}{q}}\frac{r-(1+q-p)}{2(q+r)}\right)} dz$$

Theorem.(YK 2009)

$$\int_{(-1,1)} \lambda^t Q_n(\lambda) d\psi(\lambda) = \frac{(1+q-p)(q+r)-q}{(1+q-p)(q+r)} \cdot \left(-\frac{q}{q+r}\right)^{t+n}$$

$$+ \left(\sqrt{\frac{q}{p}}\right)^n \left(\frac{p}{q+r}\right) \frac{1}{2\pi i} \oint_{|z|=1} \frac{(\sqrt{pq}(z+z^{-1})+r)^t z^n (z-z^{-1})}{\left(z-\sqrt{\frac{p}{q}}\frac{r+(1+q-p)}{2(q+r)}\right)\left(z-\sqrt{\frac{p}{q}}\frac{r-(1+q-p)}{2(q+r)}\right)} dz$$

and the total variation distance $\|\nu - \mu_t\|_{TV}$ is bounded above by

$$A\left(\frac{q}{q+r}\right)^t + B(r+2\sqrt{pq})^t,$$

where $A = \frac{(1+q-p)(q+r)-q}{(1+q-p)(1-2p)}$ and $B = \frac{\left(\frac{p}{q+r}\right)\left(1+\frac{1}{\sqrt{pq}-p}\right)}{\left(1-\sqrt{\frac{p}{q}}\frac{r+(1+q-p)}{2(q+r)}\right)\left(1+\sqrt{\frac{p}{q}}\frac{r-(1+q-p)}{2(q+r)}\right)}$.

Applications: Mixing Times

Corollary. If $\frac{q}{q+r} > r + 2\sqrt{pq}$,

$$\|\nu - \mu_t\|_{TV} \geq A\left(\frac{q}{q+r}\right)^t - B(r + 2\sqrt{pq})^t$$

We can adjust the above results for any origination site $X_0 = i$.

Example. $p = \frac{1}{11}$, $r = \frac{1}{11}$ and $q = \frac{9}{11}$.
There $\frac{q}{q+r} > r + 2\sqrt{pq}$, and

$$\|\nu - \mu_t\|_{TV} = \frac{91}{171}\left(\frac{9}{10}\right)^t \pm \frac{39}{28}\left(\frac{7}{11}\right)^t$$

Comparison to geometric ergodicity

Coupling approach:

$$P(\tau_0 > t) \sim \frac{r + 2\sqrt{pq}}{p\rho(\sqrt{q} - \sqrt{p})^2}(r + 2\sqrt{pq})^t$$

where the hitting time τ_0 determines the coupling time.

Recall: if $\left(\frac{q}{q+r}\right)^t \geq (r + 2\sqrt{pq})^t$, then

$$\|\nu - \mu_t\|_{TV} = A\left(\frac{q}{q+r}\right)^t \pm B(r + 2\sqrt{pq})^t$$

Where is $A\left(\frac{q}{q+r}\right)^t$ in the above coupling?

Comparison to geometric ergodicity

Where is $A\left(\frac{q}{q+r}\right)^t$ in the above coupling?

Extreme case: $p = 0$. If $X_t = 0$ and $Y_t = 1$, the coupling time is geometric with failure probability of $\frac{q}{q+r}$.

Amend coupling rules: let X_t and Y_t be synchronized if $X_t \neq 0$ and $Y_t \neq 0$; otherwise, let X_t and Y_t move independently.

Comparison to geometric ergodicity

Where is $A\left(\frac{q}{q+r}\right)^t$ in the above coupling?

Amend coupling rules: let X_t and Y_t be synchronized if $X_t \neq 0$ and $Y_t \neq 0$; otherwise, let X_t and Y_t move independently.

Then if $\{X_t, Y_t\} = \{0, 1\}$, conditioning on $\{X_{t+1}, Y_{t+1}\} \neq \{1, 2\}$,

$$\{X_{t+1}, Y_{t+1}\} = \begin{cases} \{1\} & \text{with probability } \frac{r}{q+r}, \\ \{0, 1\} & \text{with probability } \frac{q}{q+r}, \end{cases}$$

Solving toy example

$$P = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots \\ q & r & p & 0 & \dots \\ 0 & q & r & p & \ddots \\ 0 & 0 & q & r & \ddots \\ \vdots & \vdots & \ddots & \ddots & \ddots \end{pmatrix} \qquad q > p, \quad r > 0$$

Here $\pi_0 = 1$, $\pi_n = \frac{p^{n-1}}{q^n}$ $(n \geq 1)$ and $\rho = \frac{q-p+1}{q-p}$

OPs: $\quad Q_n(\lambda) = c_1(\lambda)\rho_1^n(\lambda) + c_2(\lambda)\rho_2^n(\lambda)$,
where $\rho_{1,2}(\lambda) = \frac{\lambda - r \pm \sqrt{(\lambda-r)^2 - 4pq}}{2p}$

First we find all point masses locations, i.e. points in$[-1, 1]$ satisfying $\mathbf{q}(\lambda) = (Q_0, Q_1, Q_2, \dots)^T \in \ell^2(\pi)$

Solving toy example

First we find all point masses locations, i.e. points in $[-1,1]$ satisfying $\mathbf{q}(\lambda) = (Q_0, Q_1, Q_2, \dots)^T \in \ell^2(\pi)$ There $\|\mathbf{q}(\lambda)\|_{\ell^2(\pi)}^{-1}$ are point masses.

So,

$$\psi(x) = \frac{q-p}{1+q-p} \cdot \delta_1(x) + \frac{(1+q-p)(q+r)-q}{(1+q-p)(q+r)} \cdot \delta_{-\frac{q}{q+r}}(x)$$

$$+ \text{ continous spectrum}$$

Solving toy example

For the continuous part , we use $(P-sI)^{-1}e_0 \in \ell^2(\mathbb{C},\pi)$

to find $(e_0,(P-sI)^{-1}e_0) = \frac{\chi_{|\rho_1(s)|<\sqrt{\frac{q}{p}}}}{\rho_1(s)-s} + \frac{\chi_{|\rho_2(s)|<\sqrt{\frac{q}{p}}}}{\rho_2(s)-s}$

Next use $C_+ - C_- = I$ property of Cauchy transform to obtain

$$\psi(x) = \frac{q-p}{1+q-p}\cdot\delta_1(x) + \frac{(1+q-p)(q+r)-q}{(1+q-p)(q+r)}\cdot\delta_{-\frac{q}{q+r}}(x)$$

$$+\frac{\sqrt{4pq-(x-r)^2}}{2\pi((r+q)x+q)(1-x)}\cdot\chi_{(r-2\sqrt{pq},r+2\sqrt{pq})}(x)$$

Reversible Markov chains over general state space

If P is a reversible M.Ch. Then there exists a spectral prob. meas. $d\psi$ and $Q_j(\lambda)$ - orthogonal polynomials w.r.t. $d\psi$ s.t.

$$P^t = F \left(\int_{-1}^{1} s^t Q_i(s) Q_j(s) d\psi(s) \right) F^T,$$

where $F = \begin{bmatrix} | & | & \\ Q_0(P)e_0 & Q_1(P)e_0 & \cdots \\ | & | & \end{bmatrix}$ and $F^T = F^{-1}$

How it works

- P is **reversible**: $\exists \pi$ s.t.

$$\pi(i)p(i,j) = \pi(j)p(j,i)$$

$\Rightarrow$ P is self-adjoint in $\ell^2(\pi)$

$\Leftrightarrow$

$$\begin{pmatrix} \sqrt{\pi_0} & 0 & \cdots \\ 0 & \sqrt{\pi_1} & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix} P \begin{pmatrix} \frac{1}{\sqrt{\pi_0}} & 0 & \cdots \\ 0 & \frac{1}{\sqrt{\pi_1}} & \ddots \\ \vdots & \ddots & \ddots \end{pmatrix}$$

is symmetric.

How it works

- There is a probability measure $d\psi$ s.t.

$$(e_0, (P - zI)^{-1} e_0) = \int \frac{d\psi(s)}{s - z}, \quad Im(z) > 0$$

- Next find a Jacobi operator $P_\triangle$ s. t.

$$(e_0, P^k e_0) = \int s^k d\psi(s) = (e_0, P_\triangle^k e_0)$$

- $Q_j(\lambda)$ are orthogonal polynomials w.r.t. $d\psi$ associated with $P_\triangle$

Example: Pentadiagonal Chebyshev operator

Consider

$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & \cdots \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \cdots \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \cdots \\ 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & \ddots \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \ddots \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \ddots \\ \cdots & \cdots & \cdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

Pentadiagonal Chebyshev operator

Classical Fourier analysis $\Rightarrow$

$$\Big(e_0, (P - zI)^{-1} e_0\Big) = \tfrac{1}{2\pi} \int_0^{2\pi} \frac{d\theta}{\frac{1}{2}[\cos(\theta)+\cos(2\theta)]-z}$$

$$= \int_{-\frac{9}{16}}^{1} \frac{d\psi(s)}{s - z} \ ,$$

where

$$d\psi(s) = \frac{1}{2\pi\sqrt{s+\frac{9}{16}}} \left(\frac{\chi_{[-\frac{9}{16},1]}(s)}{\sqrt{1-\left(\sqrt{s+\frac{9}{16}}-\frac{1}{4}\right)^2}} + \frac{\chi_{[-\frac{9}{16},0)}(s)}{\sqrt{1-\left(\sqrt{s+\frac{9}{16}}+\frac{1}{4}\right)^2}} \right) ds$$

Obtaining Jacobi operator

$$P_\triangle = \begin{pmatrix} a_0 & b_0 & 0 & \cdots \\ b_0 & a_1 & b_1 & \ddots \\ 0 & b_1 & a_2 & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix} \qquad b_j > 0$$

$$(e_0, Pe_0) = (e_0, P_\triangle e_0) = a_0,$$

$$(e_0, P^2 e_0) = (e_0, P_\triangle^2 e_0) = a_0^2 + b_0^2$$

$$(e_0, P^3 e_0) = (e_0, P_\triangle^3 e_0) = (a_0^2 + b_0^2) a_0 + (a_0 + a_1) b_0^2$$

$$(e_0, P^4 e_0) = (e_0, P_\triangle^4 e_0)$$

$$= (a_0^2 + b_0^2)^2 + (a_0 + a_1)^2 b_0^2 + b_0^2 b_1^2$$

... obtaining a_0, b_0, a_1, $b_1, \ldots$

Pentadiagonal Chebyshev operator

$$P = \begin{pmatrix} 0 & \frac{1}{2} & \frac{1}{2} & 0 & 0 & 0 & \cdots \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & \frac{1}{4} & 0 & 0 & \cdots \\ \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \cdots \\ 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \frac{1}{4} & \ddots \\ 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \frac{1}{4} & \ddots \\ 0 & 0 & 0 & \frac{1}{4} & \frac{1}{4} & 0 & \ddots \\ \cdots & \cdots & \cdots & \ddots & \ddots & \ddots & \ddots \end{pmatrix}$$

$$(e_0, Pe_0) = 0, \quad (e_0, P^2 e_0) = \frac{1}{4}, \quad (e_0, P^3 e_0) = \frac{3}{32}$$

$(e_0, P^4 e_0) = \frac{9}{64}, \quad \ldots$

Pentadiagonal Chebyshev operator

Thus

$$a_0 = 0, \quad b_0 = \frac{1}{2}, \quad a_1 = \frac{3}{8}, \quad b_1 = \frac{\sqrt{11}}{8}, \text{ etc.}$$

and

$$P_\triangle = \begin{pmatrix} 0 & \frac{1}{2} & 0 & \cdots \\ \frac{1}{2} & \frac{3}{8} & \frac{\sqrt{11}}{8} & \ddots \\ 0 & \frac{\sqrt{11}}{8} & \ddots & \ddots \\ \vdots & \ddots & \ddots & \ddots \end{pmatrix}$$

and $Q_0(\lambda) = 1$, $Q_1(\lambda) = 2\lambda$,

$$Q_2(\lambda) = \frac{32}{\sqrt{11}}\lambda^2 - \frac{6}{\sqrt{11}}\lambda - \frac{4}{\sqrt{11}}, \quad \ldots$$

Reversible Markov chains over general state space

$$P^t = F \left(\int_{-1}^{1} s^t Q_i(s) Q_j(s) d\psi(s) \right) F^T,$$

where $F = \begin{bmatrix} | & | & \\ Q_0(P)e_0 & Q_1(P)e_0 & \cdots \\ | & | & \end{bmatrix}$

Thus, taken a norm $\|\cdot\|$,

$$t_{mix}(\epsilon) = \min \left\{ t : \left\| \mu_0 F \left(\int_{(-1,1)} s^t Q_j(s) Q_k(s) d\psi(s) \right) F^T \right\| \le \epsilon \right\}$$

Reversible Markov chains over general state space

Take an ℓ^2 norm $\|\cdot\|_2$ in

$$t_{mix}(\epsilon) = \min \left\{ t : \left\| \mu_0 F \left(\int_{(-1,1)} s^t Q_j(s) Q_k(s) d\psi(s) \right) F^T \right\| \leq \epsilon \right\}$$

There, for $\mu_0 = e_0$, the distance to stationarity is

$$\left\| \int_{(-1,1)} s^t \sum_{k=0}^{\infty} Q_k(s) Q_k(P) d\psi(s) \; e_0 \right\|_2^2$$

$$= \sum_{k=0}^{\infty} \left[\int_{(-1,1)} s^t Q_k(s) d\psi(s) \right]^2$$

as $\|Q_k\|^2_{L^2(\psi)} = 1$

Papers

- Y.K. *Electronic Communications In Probability* **14** (2009), 90-103

- Y.K. *Electronic Communications In Probability* **15** (2010), 59-67