Lecture 4

Last time: Irreducibility, Aperiodicity $\Longrightarrow$ Convergence to unique stationary distn

Fundamental thm of MCs

- Proof of Fundamental thm via a Coupling Argument:
  - Coupled two copies of the MC $(X_t)_t$ and $(Y_t)_t$ that started at different states x, y
  - In our coupling $(X_t)_t, (Y_t)_t$ evolved independently until they met, and then sticked together (stickg coupling)
  - Argued that

$$\|P_x^{(t)} - P_y^{(t)}\|_{TV} \leq \Pr[\underset{\nwarrow \text{ meeting time}}{T_{xy}} > t]$$

This Lecture: From Art $\Rightarrow$ Technology

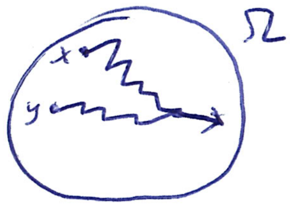
[coupling ideas to analyze MCs trace back to Doeblin in the 1930s; technology development initiated by Aldous]

Def: A coupling of a MC is a pair process $(X_t, Y_t)_t$ on $\Omega \times \Omega$ s.t.

i. $(X_t, \cdot)_t$ and $(\cdot, Y_t)_t$ are faithful copies of the MC, i.e.

$$\Pr[X_{t+1} = b \mid X_t = a] = P(a,b) = \Pr[Y_{t+1} = b \mid Y_t = a]$$

ii. if $X_t = Y_t$ then $X_{t+1} = Y_{t+1}$.



Lemma: $\overset{\max_x \|P_x^{(t)} - \pi\|_{TV}}{\Delta(t)} \leq \max_{x,y} \Pr[T_{xy} > t]$ $\leftarrow \min\{t : X_t = Y_t \mid X_0 = x, Y_0 = y\}$

Proof:

$$\begin{aligned}
\Delta(t) &\leq \max_{x,y} \|P_x^{(t)} - P_y^{(t)}\| \quad \text{(Last time)} \\
&\leq \max_{x,y} \Pr[X_t \neq Y_t \mid X_0 = x, Y_0 = y] \quad \text{(coupling lemma)} \\
&= \max_{x,y} \Pr[T_{xy} > t]. \quad \square
\end{aligned}$$

Examples

(Lazy) RW on hypercube $\{0,1\}^n$

w.pr. 1/2 do nothing
w.pr. 1/2 pick random coordinate & flip it } aperiodic & irreducible.

convergence rate: → Run to copies $\{X_t\}_t$ $(Y_t)_t$ of MC started at arbitrary x, y

→ Couple them as follows:

pick random coordinate:
- if X_t & Y_t are equal ~~at~~ that coordinate then w.pr. 1/2 keep it as is
  w.pr. 1/2 flip it in both.
- if X_t & Y_t are different at that coordinate pick 0/1 u.a.r. and set them both at that value.

→ Recall

$$\Delta(t) \le \Pr[T_{xy} > t]$$

time when chains ~~[illegible]~~ starting at x, y meet; ~~[illegible]~~ this is coupon collector time for n coupons

$$\Pr[T > n\log n + cn] \le e^{-c}$$

$$\left(\text{union bound over coupons} \quad \underbrace{\text{prob}\left[\begin{array}{l}\text{coupon } i \text{ is not collected}\\ \text{in } n\log n + cn \text{ steps}\end{array}\right]}_{\substack{\shortparallel \\ \frac{1}{n}\cdot e^{-c}}}\right)$$

→ Hence: ~~$\tau_{mix} \le O(n \cdot \log n)$~~

setting $\tau(\varepsilon) = n\cdot\log n + n\cdot\log\left(\frac{1}{\varepsilon}\right)$

$$\Delta(\tau(\varepsilon)) \le \varepsilon .$$

Part II: Rates of Convergence.

Def: $\Delta(t) = \max_x \|P_x^{(t)} - \pi\|_{TV}$ (worst distance from π after t steps)

$D(t) = \max_{x,y} \|P_x^{(t)} - P_y^{(t)}\|_{TV}$ (worst distance of two chains started at ~~x and y~~ different states)

Lemma: $\Delta(t) \underset{\text{lecture 3}}{\leq} D(t) \underset{\text{triangle inequality}}{\leq} 2\Delta(t)$

Lemma: $\Delta(t)$ is non-increasing in t.

Proof:

- Imagine two copies of the mc, $(X_t)_t$ and $(Y_t)_t$, where ~~Let~~ $X_0 = x$ and Y_0 ~~is~~ drawn from π; [in particular, for all t, Y_t follows π]
- ~~By the coupling~~ Using the coupling lemma couple the evolution of X_t and Y_t so that at time τ

$$\|X_\tau - Y_\tau\|_{TV} \equiv \mathbb{P}r[X_\tau \neq Y_\tau] \quad \{$$

- For the next step, set $X_{\tau+1} = Y_{\tau+1}$, if $X_\tau = Y_\tau$
  o.w. let the two chains take independent steps.
- $\|X_{\tau+1} - Y_{\tau+1}\|_{TV} \leq \mathbb{P}r[X_{\tau+1} \neq Y_{\tau+1}] \leq \mathbb{P}r[X_\tau \neq Y_\tau] \equiv \underbrace{\|X_\tau - Y_\tau\|_{TV}}_{\|P_x^{(\tau)} - \pi\|_{TV}}$

  $\|X_{\tau+1} - Y_{\tau+1}\|_{TV} \equiv \|P_x^{(\tau+1)} - \pi\|_{TV}$

☒

④

Def. (Mixing Time)

$$\tau_{mix} = \min\left\{ t : \Delta(t) \leq \frac{1}{2e} \right\}$$

generally
$$\tau(\varepsilon) = \min\{ t : \Delta(t) \leq \varepsilon \}$$

[$\exists$ and is finite ~~by~~ when the Fundamental theorem applies]

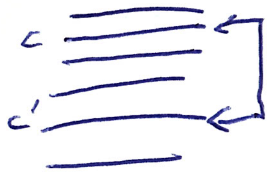
Lemma: $\Delta(t) \leq \exp\left(-\left\lfloor \frac{t}{\tau_{mix}} \right\rfloor\right)$

Proof:
- Consider two copies of the chain $(X_t)_t$ and $(Y_t)_t$ starting at x, y respecti
- Couple chains so that
$$\| p_x^{(t)} - p_y^{(t)} \|_{TV} = \mathbb{P}r[X_t \neq Y_t]$$
- Conditioning on first ~~t~~ t steps, construct a coupling of $(X_{t+s})_{s=1}^{t'}, (Y_{t+s})_{s=1}^{t'}$ as follows:
  - → if $X_t = Y_t$, set $X_{t+s} = Y_{t+s}$, $\forall s = 1, \dots, t'$
  - → o.w. suppose $X_t = x'$ and $Y_t = y'$; use coupling lemma to couple the chains so that
$$\mathbb{P}r[X_{t+t'} \neq Y_{t+t'} \mid X_t = x', Y_t = y'] \quad x' \neq y'$$
$$\|$$
$$\| P_{x'}^{(t')} - P_{y'}^{(t')} \|_{TV}$$
- Hence:
$$\| p_x^{(t+t')} - p_y^{(t+t')} \|_{TV} \leq \mathbb{P}r[X_{t+t'} \neq Y_{t+t'}] \leq \underbrace{\mathbb{P}r[X_t \neq Y_t]} \cdot \mathbb{P}r[X_{t+t'} \neq Y_{t+t'} \mid X_t \neq$$
$$\| \qquad\qquad \leq$$
$$\| P_x^{(t)} - P_y^{(t)} \|_{TV} \cdot \mathbf{D}(t')$$
$$\leq \mathbf{D}(t) \cdot \mathbf{D}(t')$$
- Hence: $\mathbf{D}(t+t') \leq D(t) \cdot D(t')$ (since above is true for all x, y)
- Hence $D(k \cdot t) \leq D(t)^k$, for all integers $k \geq 1$
- Hence: $\Delta(k \cdot \tau_{mix}) \leq D(k \cdot \tau_{mix}) \leq D(\tau_{mix})^k \leq (2 \cdot \Delta(\tau_{mix}))^k \leq e^{-k}$

Random Transposition Shuffle

Part III: Applications to Shuffling

- pick two cards c and c' u.a.r.
- switch them;
- repeat.



- Equivalently:
  - pick card c and position p u.a.r.
  - exchange card c with whatever card is at position p
- Coupling of $(X_t)_t$, $(Y_t)_t$: pick same c and p at all steps t
- Let $d(X_t, Y_t) =$ # positions in two decks that differ
- If card c is at same position in X_t, Y_t then: $d_{t+1} = d_t$
- If -//- c is at different -//- -//- -//- :
  - if card at position p is the same, then $d_{t+1} = d_t$
  - if -//- p different, then $d_{t+1} \leq d_t - 1$

Hence $\mathbb{P}r[\, d_{t+1} < d_t \,] = \left(\frac{d_t}{n}\right)^2$ ⇝ expected time to decrease is $\left(\frac{n}{d_t}\right)^2$

So $\mathbb{E}[T_{xy}] \leq \sum_{d=1}^{n} \left(\frac{n}{d}\right)^2 \leq c \cdot n^2$

Hence, Markov's Inequality: $\mathbb{P}r[T_{xy} > c' \cdot n^2] < \frac{c}{c'} = \frac{1}{2e}$, for appropriate choice of c'

$\Rightarrow \quad \tau_{mix} = O(n^2)$

exercise: Design better coupling giving $\tau_{mix} = O(n \cdot \log n)$?
(2 pt)