- **Fundamental Thm of MC's**

Part II of lecture (part I was slides)

Let $(X_t)_t$ be a MC on a finite state space Ω;

- P be its transition matrix where P is irreducible & aperiodic

Then: i) $\exists!\ \pi$ s.t. $(\pi = \pi \cdot P) \wedge (\pi(x) > 0, \forall x \in \Omega)$

ii) $\forall x \in \Omega$: $P_x^{(t)} \xrightarrow{t\to\infty} \pi$

- **Proof** Two parts: a) $\exists\ \pi$ s.t. $\pi = \pi \cdot P$ & $\pi(x) > 0, \forall x$

  b) assuming such a π exists $P_x^{(t)} \xrightarrow{t\to\infty} \pi, \forall x$

  $(a+b) \Rightarrow$ the uniqueness of π in a

  why? suppose exist π' s.t. $\pi' = \pi' \cdot P$

  $$\pi' = \sum_{x\in\Omega} \alpha_x \cdot e_x$$

  $$\left.\begin{array}{l} \pi' P^t = \sum_{x\in\Omega} \alpha_x e_x P^t \xrightarrow{t\to\infty} \pi \\ \text{on the othe hand } \pi' P^t = \pi' \end{array}\right\} \pi' = \pi$$

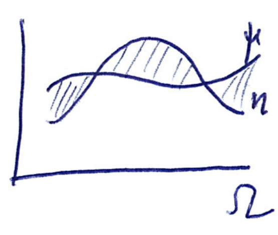
  ∎

- Implementation: postpone a) for later, do b) now

- **Def: (TV distance)**

  μ, η probability distn's on Ω; their total variation distance is

  $$\|\mu - \eta\|_{TV} = \frac{1}{2} \sum_{x\in\Omega} |\mu(x) - \eta(x)|$$



Lemma: $\|\mu - \eta\|_{TV} \equiv \max_{A \subseteq \Omega} |\mu(A) - \eta(A)|$

Note: If X, Y are r.v.'s on Ω distributed according to μ, η then $\|X - Y\|_{TV} \triangleq \|\mu - \eta\|_{TV}$.

~~Def~~ Coupling)

μ, η probability distn's over Ω; a distribution w over $\Omega \times \Omega$ ~~is called~~ is called **a coupling of μ and η** if

$$\mu(x) = \sum_{y \in \Omega} w(x,y)$$

$$\eta(y) = \sum_{x \in \Omega} w(x,y)$$

(i.e. the marginals of w w.r.t. the fist resp. second coordinate are μ and η respectively)

Lemma (the Coupling Lemma)

μ, η probability distn's over Ω.

a) for any coupling w of μ, η if (X, Y) distributed according to w then:

$$\underbrace{\mathbb{P}r}_{\text{measure defined by } w}[X \neq Y] \geq \|\mu - \eta\|_{TV}$$

b) exists coupling s.t.

$$\mathbb{P}r[X \neq Y] = \|\mu - \eta\|_{TV}.$$

(this is called **optimal coupling**)

Proof of the Coupling Lemma.

a) $\forall z \in \Omega$: $\mu(z) = \mathbb{P}r[X = z] = \mathbb{P}r[X = z, X \neq Y] + \mathbb{P}r[X = z, \cancel{Y} = z]$

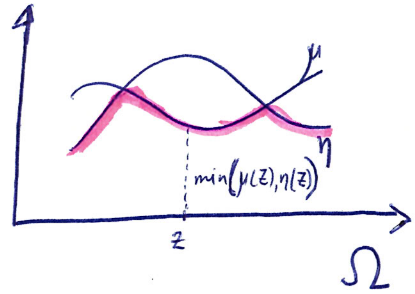
$$\leq \mathbb{P}r[X = z, X \neq Y] + \mathbb{P}r[Y = z]$$

$$\leq \mathbb{P}r[X = z, X \neq Y] + \eta(z)$$

By symmetry: $\eta(z) = \leq \mathbb{P}r[Y = z, X \neq Y] + \mu(z)$

Hence: $2 \cdot \|\mu - \eta\|_{TV} = \sum_{z: \mu(z) \geq \eta(z)} \underbrace{\mathbb{P}r[X = z, X \neq Y]}_{\text{these events are disjoint and their union is the event } X \neq Y} + \sum_{z: \eta(z) \geq \mu(z)} \underbrace{\mathbb{P}r[Y = z, X \neq Y]}_{\text{similarly}} \leq 2 \cdot \mathbb{P}r[X \neq Y]$

b> of coupling lemma



- look at lower envelope (total mass below it is exactly $1-\|\mu-\eta\|_{TV}$)
- define coupling of μ,η as follows
  - for all z, w.pr. $\min(\mu(z),\eta(z))$ set $X=Y=z$
  - complete the coupling in an arbitrary way

Clearly: $\Pr[X=Y] = 1-\|\mu-\eta\|_{TV}$ □

- Ready to implement step b of the proof of the Fundamental thm
- fix arbitrary $x,y\in\Omega$
- Consider two copies of the chain $(X_t)_t$ and $(Y_t)_t$ starting at x & y respectively
- Let (X_t,Y_t) be an independent coupling of the chains modified as follows:

  If at some time s $X_s=Y_s$ then $X_t=Y_t$, for all $t\geq s$ (sticky chain coupling)

- [In other words $(X_t,Y_t)_t$ is a chain on $\Omega\times\Omega$ w/ transition matrix:

$$q\big((x_1,y_1),(x_2,y_2)\big)=\begin{cases}P(x_1,x_2)\cdot P(y_1,y_2) & ,\text{ if } x_1\neq y_1\\ P(x_1,x_2) & ,\text{ if } x_1=y_1 \ \&\ x_2=y_2\\ 0 & ,\text{ o.w.}\end{cases}$$
]

- Let T be the (random) first time that the chains meet, i.e. $T=\min\{t: X_t=Y_t\}$
- By the coupling lemma, for all t:

$$\underbrace{\sum_{z\in\Omega}\left|\Pr[X_t=z]-\Pr[Y_t=z]\right|}_{\|P_x^{(t)}-P_y^{(t)}\|_{TV}}\leq 2\cdot\Pr[X_t\neq Y_t]=2\cdot\Pr[T>t] \quad (**)$$

Lemma: $\Pr[T>t]\xrightarrow{t\to+\infty}0$ [the only place where aperiodicity is used!!]

Proof: - By exercise last time IRR + AP $\Rightarrow\exists$ time τ s.t. $P^{\tau}(z_1,z_2)>0,\ \forall z_1,z_2$ call $C=\min_{z_1,z_2}P^{\tau}(z_1,z_2)$

- hence: $P^{\tau}(x_1,z)\cdot P^{\tau}(y_1,z)\geq C^2>0,\ \forall x_1,y_1,z$
- $\Pr[T>k\cdot\tau]\leq(1-C^2)^k\xrightarrow{k\to+\infty}0$ □

- Hence

$$\|P_x^{(t)}-P_y^{(t)}\|_{TV}\xrightarrow{t\to\infty}0,\ \forall x,y$$

- Lemma (exercise): $\forall z$: $\|P_z^{(t)}-\pi\|_{TV}\leq\underbrace{\max_{x,y}\|P_x^{(t)}-P_y^{(t)}\|_{TV}}_{D(t)}$

- ∴ $\forall z$: $P_z^{(t)}\xrightarrow{t\to\infty}\pi$ □ (end of step b of Fundamental theorem)

(proof of exercise from previous page)

$$\pi = \sum_{y \in \Omega} \alpha_y \cdot e_y \quad , \text{ where } \sum_y \alpha_y = 1 \ , \ \alpha_y \geq 0$$

$$\pi = \pi \cdot P^t = \sum_y \alpha_y \, e_y P^t = \sum_y \alpha_y \cdot P_y^{(t)}$$

$$\|P_x^{(t)} - \pi\|_{TV} = \| P_x^{(t)} - \sum_y \alpha_y P_y^{(t)} \|_{TV} = \| \sum_y \alpha_y P_x^{(t)} - \sum_y \alpha_y P_y^{(t)} \|_{TV} \leq \sum_y \alpha_y \| P_x^{(t)} - P_y^{(t)} \|_{TV}$$

$$\leq \sum_y \alpha_y \cdot D(t) = D(t) \quad \boxtimes$$

(proceeding to step α of the proof of the fundamental thm)

- pick arbitrary x
- define $q_x(x) = 1$ and, for $y \neq x$, $q_x(y)$ = expected # that the MC ~~visits~~ started at x visits y before coming back to x
- $\mathcal{E}x$ (1pt) : Show that $\pi(y) \propto q_x(y)$ is stationary ; i.e. the normalized vector q_x is a stationary distribution of the chain.

$\boxtimes \boxtimes$

(end of the proof of the fundamental theorem)