Fundamental Thm of MC's

Let \( \phi_t(X_t) \) be a MC on a finite state space \( S \);

- \( P \) be its transition matrix where \( P \) is irreducible & periodic

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

ii) \( \forall x \in S \) : \( P^t_x \rightarrow \pi \)

• Proof: Two parts:
  - a) \( \exists! \pi \) s.t. \( \pi = \pi \cdot P \) & \( \pi(x) > 0, \forall x \)
  - b) assuming such a \( \pi \) exists \( P^t_x \rightarrow \pi, \forall x \)

(\( \alpha + \beta \)) \Rightarrow the uniqueness of \( \pi \) in \( \alpha \)

why? suppose exist \( \pi' \) s.t. \( \pi' = \pi' \cdot P \)

\[ \pi' = \sum_{x \in S} \pi'_x e_x \]

\[ \pi' P^t = \sum_{x \in S} \pi'_x e_x P^t \rightarrow \pi \]

on the other hand \( \pi' P^t = \pi' \)

• Implementation: postpone \( \rightarrow \) for later, do \( \rightarrow \) now

• Def. (TV distance)

\( \mu, \eta \) probability dist'n's on \( S \); their total variation distance is

\[ ||\mu - \eta||_{TV} = \frac{1}{2} \sum_{x \in S} |\mu(x) - \eta(x)| \]

Lemma:

\[ ||\mu - \eta||_{TV} \leq \max_{A \subseteq S} |\mu(A) - \eta(A)| \]

Note: If \( X, Y \) are r.v.'s on \( S \) distributed according to \( \mu, \eta \) then \( ||X - Y||_{TV} \leq ||\mu - \eta||_{TV} \).
\[\nu, \eta \text{ probability distn's over } \Omega; \text{ a distribution } w \text{ over } \Omega \times \Omega \]
called a coupling of \(\nu\) and \(\eta\) if

\[
\nu(x) = \sum_{y \in \Omega} w(x, y) \\
\eta(y) = \sum_{x \in \Omega} w(x, y)
\]

(i.e. the marginals of \(w\) w.r.t. the first resp. second coordinate are \(\nu\) and \(\eta\) respectively)

**Lemma (the Coupling Lemma)**

\(\nu, \eta\) probability distn's over \(\Omega\).

\(\triangleright\) for any coupling \(w\) of \(\nu, \eta\) it \((X, Y)\) distributed according to \(w\) then:

\[
Pr[X \neq Y] \geq \|\nu - \eta\|_{tv}
\]

\(\triangleright\) exists coupling s.t.

\[
Pr[X \neq Y] = \|\nu - \eta\|_{tv}.
\]

(this is called optimal coupling)

**Proof of the Coupling Lemma**

\(\triangleright\) \(\forall z \in \Omega: \nu(z) = Pr[X = z] = Pr[X = z, X \neq Y] + Pr[X = z, Y = z]

By symmetry:

\[
\eta(z) = \leq Pr[Y = z, X \neq Y] + \eta(z)
\]

Hence:

\[
2\|\nu - \eta\|_{tv} = \sum_{z: \nu(z) > \eta(z)} Pr[X = z, X \neq Y] + \sum_{z: \eta(z) > \nu(z)} Pr[Y = z, X \neq Y] \leq 2 \cdot Pr[X \neq Y]
\]

These events are disjoint and their union is the event \(X \neq Y\)
- Look at lower envelope (total mass below it is exactly $1 - \|\mu - \eta\|_{TV}$)
- Define coupling of $\mu, \eta$ as follows:
  - $\mu(x) = \mu(x), \eta(y) = \eta(y)$ for all $x$, w.r.t. $\min(\mu(z), \eta(z))$ set $X = Y = Z$
  - Complete the coupling in an arbitrary way

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

---

- Ready to implement step 3 of the proof of the Fundamental Theorem
- Fix arbitrary $x, y \in \Omega$
- Consider two copies of the chain $(X_t)_t$ and $(Y_t)_t$ starting at $x$ and $y$ 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 $\square \times \square$ with transition matrix:

  $$q_t((x_1, y_1), (x_2, y_2)) = \begin{cases} \Pr(x_1 \neq y_1, \eta(x_2) = \eta(y_2)), & \text{if } x_1 \neq y_1 \\ \Pr(x_1 = y_1) & \text{if } x_1 = y_1 \land y_1 = y_2, \\ 0 & \text{otherwise} \end{cases}$$

- 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 $t$:

  $$\mathbb{E} \left[ \Pr[X_t = z] - \Pr[Y_t = z] \right] \leq 2 \cdot \Pr[X_t \neq Y_t] = 2 \cdot \Pr[T > t]$$

- Hence:

  $$\|P^{(t)}_x - P^{(t)}_y\|_{TV} \xrightarrow{t \to \infty} 0, \forall x \neq y$$

- Lemma: $\|P^{(t)}_x - P^{(t)}_y\|_{TV} \leq \max_{X_2, Y_2} \|P^{(t)}_{(X_2)} - P^{(t)}_{(Y_2)}\|_{TV}$

- **Lemma:** $P^{(t)}_x \xrightarrow{t \to \infty} \Pi$ is the only place where aperiodicity is used!!
\[ \pi = \sum_{y=1}^{2} \pi_y e_y, \text{ where } \sum_{y=1}^{2} \pi_y = 1, \pi_y \geq 0 \]

\[ \pi = \pi \cdot p^t = \sum_{y=1}^{2} \pi_y e_y p^t = \sum_{y=1}^{2} \pi_y p_y^{(H)} \]

\[ \| x^{(t)} - \pi \|_V = \| p_x^{(t)} - \sum_{y=1}^{2} \pi_y p_y^{(H)} \|_V = \| \sum_{y=1}^{2} \pi_y p_x^{(H)} - \sum_{y=1}^{2} \pi_y p_y^{(H)} \|_V \leq \sum_{y=1}^{2} \pi_y \| p_x^{(H)} - p_y^{(H)} \|_V \leq \sum_{y=1}^{2} \pi_y \cdot D^{(t)} = D^{(t)} \]
pick arbitrary x

* define \( q_x(x) = 1 \) and, for \( y \neq x \), \( q_x(y) = \text{expected } \pi \text{ visits to before coming back to } x \).

* \( \pi \propto 1 \) \( \pi \propto q_x(y) \) is stationary; i.e., the normalized vector \( q_x \) is a stationary distribution of the chain.

---

(proceeding to step 2 of the proof of the fundamental theorem)