* Basic defns

- Markov chain - (finite) state space Ω, transition prob P
  - $\mathbb{P}[X_{t+1}=y \mid X_t = x, H_{t-1}] = \mathbb{P}[X_{t+1}=y \mid X_t = x] = P(x,y)$
  - $P(x,y) \geq 0 \ \forall x,y$, $\sum_y P(x,y) = 1 \ \forall x$ (stochastic)
  - Given starting distn π_0 on Ω, $\pi_t = \pi_0 P^t$
  - Equivalent to random walk on directed, weighted graph
- Reversible MC - MC (Ω, P) with stationary distn π (ie, $\pi P = \pi$) s.t. $\pi(x)P(x,y) = \pi(y)P(y,x) \ \forall x,y$
  - Equivalent to RW on undirected graph with weights
    $w(x,y) = \pi(x)P(x,y) = \pi(y)P(y,x)$

- Irreducibility - MC (Ω, P) is irreducible if $\forall x,y \in \Omega$, $\exists\ t \geq 0$ s.t $P^t(x,y) > 0$
  - i.e., graph G is strongly connected
- Period of a state $x \in \Omega \equiv GCD(\{t \geq 1 \mid P^t(x,x) > 0\})$

  set of return times for x
- Apeiodicity - MC(Ω, P) is aperiodic if every state has period $= 1$
  - i.e. graph G is non-bipartite if undirected

- Some facts -
  1) If MC is irreducible, then every state x has the same period
  2) If MC is irreducible, and at least one x has a self-loop (i.e., $P(x,x) > 0$), then it is aperiodic
  3) For a finite state, irreducible, aperiodic MC, $\exists$ integer $r > 0$ s.t $P^r(x,y) > 0 \ \forall x, y \in \Omega$

- Hitting and First Return Times

- For any $x \in \Omega$, its hitting time is defined as

$$\tau_x = \min\{t \geq 0 \mid X_t = x\}$$
$$\tau_x^+ = \min\{t \geq 1 \mid X_t = x\}$$

- If $X_0 = x$, then τ_x^+ is called the first return time

- Lemma - For any finite-state irreducible chain, $\mathbb{E}_x[\tau_y^+] < \infty \ \forall x, y$

Pf - . We know $\exists r > 0$ s.t $P^r(x,y) > \varepsilon \ \forall x, y$ (and some $\varepsilon > 0$)
. For any $X_t \in \Omega$, $\mathbb{P}[\text{Hit } y \text{ in } \{t, t+1, \ldots, t+r\}] > \varepsilon$
. For any $k > 0$, $\mathbb{P}[\tau_y^+ > kr] < (1-\varepsilon)^k$
. $\mathbb{E}_x[\tau_y] = \sum_{t \geq 0} \mathbb{P}[\tau_y^+ > t] \leq \sum_{k \geq 0} r \, \mathbb{P}[\tau_y^+ > kr] < \frac{r}{\varepsilon}$

- The fundamental theorem of (finite) MCs

Given any MC (Ω, P) with finite Ω

existence i) $\exists$ stationary distn π s.t $\pi P = \pi$, $\pi(x) \geq 0$, $\sum_{x \in \Omega} \pi(x) = 1$

uniqueness ii) If MC is irreducible, then π is unique and moreover
$$\pi(x) = \frac{1}{\mathbb{E}_x[\tau_x^+]} > 0 \ \forall x \in \Omega$$

convergence iii) If MC is irreducible and aperiodic, then $\forall x \in \Omega$
$$\lim_{t \to \infty} \pi_t(x) = \pi(x)$$
(in fact, we can give more exact convergence rates)

Existence of π

- Let $\tilde{\pi}(y) = \mathbb{E}_z[\#\text{ of visits to } y \text{ in } \{0,1,2,\dots,\tau_z^+\}]$

(for any z) $\quad = \sum_{t=0}^{+\infty} \mathbb{P}_z[X_t = y, \tau_z^+ > t]$

- Note - $\tilde{\pi}(y) \leq \mathbb{E}_z[\tau_z^+] \ \forall y, z \Rightarrow 0 < \tilde{\pi}(y) < \infty$ if MC irreducible
- Claim - $\tilde{\pi}(y)$ is stationary. To check this -

$$\sum_x \tilde{\pi}(x) P(x,y) = \sum_x \sum_{t \geq 0}^{\infty} \mathbb{P}_z[X_t = x, \tau_z^+ > t] P(x,y)$$
$$= \sum_{t=0}^{\infty} \sum_x \mathbb{P}_z[X_t = x, X_{t+1} = y, \tau_z^+ > t]$$
$$= \sum_{t=1}^{\infty} \mathbb{P}_z[X_t = y, \tau_z^+ \geq t]$$
$$= \tilde{\pi}(y) - \underbrace{\mathbb{P}_z[X_0 = y, \tau_z^+ > 0]}_{\mathbb{1}_{\{y=z\}}} + \overbrace{\sum_{t=1}^{\infty} \mathbb{P}[X_t = y, \tau_z^+ = t]}^{= \mathbb{P}_z[X_{\tau_z^+} = y] = \mathbb{1}_{\{y=z\}}}$$
$$= \tilde{\pi}(y)$$

- Finally, to get a distn, we normalize

$$\pi(y) = \frac{\tilde{\pi}(y)}{\sum_x \tilde{\pi}(x)} = \frac{\tilde{\pi}(y)}{\mathbb{E}_z[\tau_z^+]} \Rightarrow \boxed{\pi(x) = \frac{1}{\mathbb{E}_x[\tau_x^+]}}$$

Uniqueness of π

- A function $h : \Omega \to \mathbb{R}$ is said to be harmonic at x if

$$h(x) = \sum_{y \in \Omega} P(x,y)\, h(y)$$

  — h is harmonic on Ω if $Ph = h$

- Lemma - h is harmonic on Ω iff h is constant

Pf - $\because \Omega$ is finite $\Rightarrow \exists\, x_0$ s.t $h(x_0) \geq h(x)\ \forall x \in \Omega$
(maximal)

- Let $h(x_0) = M$. If h is not constant $\Rightarrow \exists\, x \neq x_0$ s.t $h(x) < M$
- If $P(x_0, x) > 0$, then $h(x_0) = P(x_0,x)\, h(x) + \sum_{y \neq x} P(x_0,y) h(y) < M$ Contradiction
- Even otherwise, by irreducibility, $\exists$ path $x_0, x_1, \ldots, x_n = x$ s.t $P(x_i, x_{i+1}) > 0$. Repeat argument to show $h(x_{n-1}) = M$.

- By the above lemma, we know $(P - I)$ has row rank (and hence column rank) $= |\Omega| - 1$. Thus $\nu = \nu P$ has a one-dimensional space of solns $\Rightarrow \pi$ is unique

Convergence to π

- **Total Variation Distance** - For 2 dist μ, ν on Ω

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

- If Ω is finite.

$$\|\mu - \nu\|_{TV} = \frac{1}{2} \sum_{x \in \Omega} |\mu(x) - \nu(x)| = \frac{1}{2} \|\mu - \nu\|_1$$

$$= \sum_{x \in \Omega} (\mu(x) - \nu(x))^+ = 1 - \sum_{x \in \Omega} \min(\mu(x), \nu(x))$$

- $\|\mu - \nu\|_{TV} = \frac{1}{2} \sup \left\{ \sum_{x \in \Omega} \left(f(x)\mu(x) - f(x)\nu(x) \right) \middle| \max_{x \in \Omega} |f(x)| \leq 1 \right\}$

Pf - $\sum_{x \in \Omega} f(x)(\mu(x) - \nu(x)) \leq \sum_{x} |\mu(x) - \nu(x)| = 2\|\mu - \nu\|_{TV}$

for opposite direction, choose $f(x) = \begin{cases} 1 & ; \mu(x) \geq \nu(x) \\ -1 & ; \mu(x) < \nu(x) \end{cases}$

- $\|\mu - \nu\|_{TV} \leq \|\mu - \eta\|_{TV} + \|\eta - \nu\|_{TV}$ (Δ Inequality)

Coupling - (X, Y) said to be a coupling for a pair of distributions (μ, ν) if X and Y are defined on the same probability space, and $\mathbb{P}[X=x] = \mu(x)$, $\mathbb{P}[Y=y] = \mu(y)$

- Note - If $q \equiv$ joint distn of (X, Y). Then

$$\sum_y q(x,y) = \mu(x), \quad \sum_x q(x,y) = \nu(y), \quad q(x,y) \text{ not necessarily } = \mu(x) \cdot \nu(y)$$

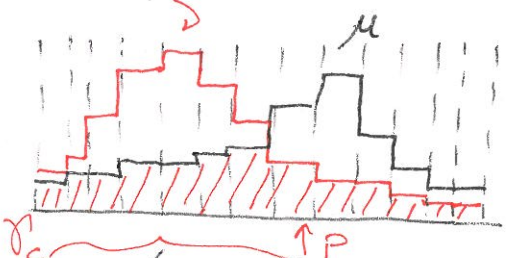
- Lemma - $\boxed{\|\mu - \nu\|_{TV} = \inf \{ \mathbb{P}[X \neq Y] \mid (X,Y) \text{ coupling of } \mu, \nu \}}$

Pf - For any coupling (X,Y) and event $A \subseteq \Omega$

$$\mu(A) - \nu(A) = \mathbb{P}[X \in A] - \mathbb{P}[Y \in A]$$
$$\leq \mathbb{P}[X \in A, Y \notin A] \leq \mathbb{P}[X \neq Y]$$

- For the other direction, consider the following coupling

Let $p = \sum_{x \in \Omega} \min(\mu(x), \nu(x)) = 1 - \|\mu - \nu\|_{TV}$



Now we construct (X,Y) as follows

1) Generate $Z \sim \mathrm{Ber}(p)$
2) If $Z = 1$, choose $w \in \Omega$ wp $\underbrace{\frac{\min(\mu(x), \nu(x))}{p}}_{\gamma_c}$. Set $X = Y = w$
3) If $Z = 0$, choose $X \sim \underbrace{\frac{(\mu(x) - \nu(x))^+}{1-p}}_{\gamma_X}$, $Y \sim \underbrace{\frac{(\nu(x) - \mu(x))^+}{1-p}}_{\gamma_Y}$

- Check -
$$\mu(x) = p\,\gamma_c(x) + (1-p)\,\gamma_X(x) = \mathbb{P}[X = x]$$
$$\nu(y) = p\,\gamma_c(y) + (1-p)\,\gamma_Y(y) = \mathbb{P}[Y = y]$$

$$\mathbb{P}[X \neq Y] \underset{\substack{\uparrow \\ \text{since } \gamma_X, \gamma_Y \text{ have non-overlapping support}}}{=} 1 - p = \|\mu - \nu\|_{TV}$$

Thm - Suppose P is irreducible, aperiodic, finite state-space, with stationary distn π. Then $\exists\ \alpha \in (0,1)$ and $C > 0$ s.t

$$d(t) = \max_{x \in \Omega} \| P^t(x,\cdot) - \pi(\cdot) \|_{TV} \leq C \alpha^t$$

Pf - Since P aperiodic, irreducible $\Rightarrow \exists\ r$ s.t $P^r(x,y) > 0\ \forall x,y$

- $\because \pi(x) > 0\ \forall x \Rightarrow \exists\ \delta$ s.t $P^r(x,y) > \delta \pi(y)\ \forall x,y$

- Let $\theta = 1-\delta \Rightarrow P^r = (1-\theta)\underline{\Pi} + \theta Q$ for stochastic mat Q $\quad \begin{pmatrix} \pi \\ \pi \\ \vdots \\ \pi \end{pmatrix}$

- Suppose $P^{kr} = (1-\theta^k)\Pi + \theta^k Q^k$ for integer $k \geq 1$

then
$$\begin{aligned} P^{(k+1)r} &= \left((1-\theta^k)\underline{\Pi} + \theta^k Q^k\right)P^r \qquad (\underline{\Pi}P = \underline{\Pi}) \\ &= (1-\theta^k)\underline{\Pi} + \theta^k Q^k\left((1-\theta)\underline{\Pi} + \theta Q\right) \\ &= (1-\theta^{k+1})\underline{\Pi} + \theta^{k+1} Q^{k+1} \qquad (Q\,\underline{\Pi} = Q \text{ for any } Q) \end{aligned}$$

- $$P^{kr+j} - \underline{\Pi} = \theta^k\left(Q^k P^j - \underline{\Pi}\right)$$

$$\Rightarrow \| P^{kr+j}(x,\cdot) - \pi \|_{TV} = \frac{\theta^k}{2} \| (Q^k P^j - \underline{\Pi})_x \|_1 \leq \theta^k$$

$$\leq \| Q^k P^j_{(x)} \|_1 + \| \underline{\Pi} \|_1 = 2$$

- For more refined estimates of convergence, define
  - $d_x(t) = \|e_x P^t - \pi\|_{TV}$, $\bar{d}(t) = \max_{x\in\Omega} d_x(t)$
  - $d_{xy}(t) = \|e_x P^t - e_y P^t\|_{TV}$, $\bar{d}(t) = \max_{x,y\in\Omega} d_{x,y}(t)$

- Lemma - i) $\bar{d}(t+s) \le \bar{d}(t)\,\bar{d}(s)$ (submultiplicative)

  ii) $d(t+s) \le 2\,d(t)\,d(s)$

  iii) $d(t) \le \bar{d}(t) \le 2d(t)$

Pf - iii) $d_{x,y}(t) \le \|e_x P^t - \pi\|_{TV} + \|\pi - e_y P^t\|_{TV} = d_x(t) + d_y(t)$

$\Rightarrow \bar{d}(t) \le 2\,d(t)$

Also
$$
\begin{aligned}
d_x(t) &= \max_{A\subseteq\Omega} |P^t(x,A) - \pi(A)| \\
&= \max_{A\subseteq\Omega} \Big| \sum_{y\in\Omega} \pi(y)\big(P^t(x,A) - P^t(y,A)\big) \Big| \\
&\le \sum_{y\in\Omega} \pi(y) \max_{A\subseteq\Omega} |P^t(x,A) - P^t(y,A)| = \sum \pi(y)\, d_{xy}(t) \\
&\le \bar{d}(t)
\end{aligned}
$$

$\Rightarrow d(t) \le \bar{d}(t)$

(Alt - check that $\|\mu - \nu\|_{TV}$ is convex in μ for fixed ν)

i) Let (X_s, Y_s) be optimal coupling for $e_x P^s$ and $e_y P^s$ ⑨

i.e., $\mathbb{P}[X_s \neq Y_s] = d_{x,y}(s) \leq \bar{d}(s)$

- $$e_x P^{s+t} - e_y P^{s+t} = \sum_{z \in \Omega} \mathbb{P}[X_s = z]\, e_z P^t - \mathbb{P}[Y_s = z]\, e_z P^t$$

$$= \mathbb{E}_{X_s, Y_s}\left[(e_{X_s} - e_{Y_s}) P^t \right]$$

$$\leq \mathbb{P}[X_s \neq Y_s]\, \bar{d}(t) \leq \bar{d}(s)\, \bar{d}(t)$$

ii) $d(t+s) \leq \bar{d}(t)\, d(s)$ (or $\bar{d}(s)\, d(t)$)

$\leq 2\, d(t)\, d(s)$

- Combining above, we have $d(ct) \leq \bar{d}(ct) \leq (\bar{d}(t))^c$
- Mixing Time $\quad t_{mix}(\varepsilon) = \min\{t \mid d(t) \leq \varepsilon\}$

(Usually, $t_{mix}(\underbrace{1/4}_{\text{alt, } 1/2e}) \triangleq t_{mix}$)

- $$d(\ell\, t_{mix}) \leq (\bar{d}(t_{mix}))^{\ell} \leq (2 d(t_{mix}))^{\ell} \leq 2^{-\ell}$$

$$\Rightarrow t_{mix}(\varepsilon) \leq \lceil \log_2 (1/\varepsilon) \rceil\, t_{mix}$$