Stationary distributions

A Markov chain with transition matrix $P_{x,y}$ has stationary distribution $\pi$ if,

$$\pi P = \pi$$

To interpret this if $X_0 \sim \nu$, that is $\mathbb{P}(X_0 = k) = \nu_k$, then

$$\mathbb{P}(X_t = j) = \sum_k \mathbb{P}(X_t = j \mid X_0 = k) \mathbb{P}(X_0 = k) = \sum_k \nu_k \cdot P_{kj} = \nu P_j.$$ 

So $X_t \sim \nu P$. In general $X_n \sim \nu P^n$.

So if $\pi$ is stationary then $X_n \sim \pi$ for all $n$ if $X_0 \sim \pi$.

**Example:** RW on a graph.

$$\pi_i = \frac{d_i}{2|E|} \text{ where } d_i \text{ is degree of } i.$$ 

$$\pi P_j = \sum_i \frac{d_i}{2|E|} \cdot P_{ij} = \sum_i \frac{d_i}{2|E|} \cdot \frac{1}{d_i} \mathbb{I}(\text{in}_j) = \frac{1}{2|E|} \sum_i \mathbb{I}(\text{in}_j) = \frac{d_j}{2|E|}$$

- Random to top shuffle

Let $g_k = (1,2,\ldots,k)$ and $G_n$ uniform on $\{g_k : 1 \leq k \leq n\}$.

Then $X_n = G_n X_{n-1}$ is a top to random shuffle.
Let $\pi(\sigma) = \frac{1}{n}$ be the uniform permutation.

Questions: Is $\pi$ unique? Does $X_n \rightarrow \pi$. How fast?

- Example: RW on disconnected graph

\[
\begin{pmatrix}
\frac{1}{2} & 0 \\
\frac{1}{2} & 0
\end{pmatrix}
\]

A Markov chain is irreducible if for all $i,j$ there is $n$ such that $(P^n)_{i,j} > 0$, that is $P[X_n = j | X_0 = i] > 0$.

Perron-Frobenius Theorem

If $P$ is a stochastic matrix then it has a left eigenvector $\mu$ with $\mu P = \mu$ and $\sum \mu_i = 1$. The entries of $\mu$ are positive. If $P$ is irreducible then $\mu$ is unique.

Proof: Linear Algebra.

Probabilistic Existence proof:
Let $M_n = \frac{1}{n} \sum_{i=1}^{n} M P^n_i$.

Now $M_n P - M_n = \frac{1}{n} M (P^n - P) \to 0$

$M \subset \{ v \in [0, I] : \frac{j}{n}, i = 1 \}$ compact set so

$\exists n_k$ such that $M_{n_k} \to \tilde{M}$.

Since $M_{n_k} (P - I) \to 0$, $\tilde{M} (P - I) = 0$

$\Rightarrow \tilde{M}$ is stationary

**Positivity:** We must have $M_i > 0$ for some $i$.

For any $i$, $\exists n_k$ such that $(P^n)_i > 0$.

$M_j = (M P^n)_j \Rightarrow M_i P^n_i > 0$.

**Uniqueness:** Let $S = \inf \{ n \geq 1 : X_n = i \}$.

Then $M_i = (E[S | X_0 = i])^-$.

Suppose $X_0 \sim v$ and $v$ is stationary.

Let $T_k$ be $k$-th visit to $i$. Then

$T_k - T_{k-1}$ IID

$\Rightarrow \frac{T_k}{n} \to E[S] = \frac{1}{M_i}$ a.s.

If $N_n = \# \{ 1 \leq s \leq n : X_s = i \}$ then

$\frac{N_n}{n} \to M_i$ a.s.
\[ S_0 \text{ E}(X_1|X_0) \rightarrow m; \]
But \[ \text{E}X_i = \sum_{k=1}^{n} P(X_k = i) = n \mu_i; \]
\[ \rightarrow \mu_i = \lambda_i. \]

**Periodicity:**
\[ \overset{\circ}{x} \overset{\circ}{x} \overset{\circ}{x} \]
\[ \tau(x) = \frac{1}{3} \text{ uniform} \]

\[ P(X_{3n} = 1 | X_0 = 13) = 1 \]
\[ P(X_{3n+1} = 2 | X_0 = 1) = 1 \quad \Rightarrow \quad X_n \xrightarrow{d} \tau. \]

A state \( x \) in a Markov Chain is aperiodic if \( \text{GCD}(S) = 1 \) where \( S = \{ n \geq 1 : P(X_n = x | X_0 = x) > 0 \} \).

**Claim:** Closed under Addition: \[ \text{If } n, m \in S \text{ then } n + m \in S \]
\[ P^{n+m} = \sum \limits_y P^n \overset{x}{y} P^m \overset{y}{x} \geq P^n \overset{x}{x} P^m \overset{x}{x} > 0. \]

**Fact:** If \( \text{GCD}(A) = 1 \) and \( A \) closed under addition then \( |N \setminus A| < \infty \), i.e. \( N \) such that \( \forall n \geq n', n \in A. \)

**Defn:** A Markov chain is ergodic if it is irreducible and aperiodic.

**Claim:** If \( X_n \) is ergodic then \( \exists N \) such that \( \forall x, y, n \geq N \text{ then } P^n_{x,y} > 0. \)

**Proof:** Suppose \( Z \) is aperiodic so \( \forall m \geq M \text{ } P^m_{Z,Z} > 0. \) Now for some \( k, k \)
\[ P_{x_1} > 0, \ P_{y_1} > 0. \]
\[ \forall n \geq k + l + M, \ P_{x_2} \leq P_{x_2}^k \ P_{x_2}^{n-k} \ e^{-e} P_{y_2} \geq 0. \]

**Theorem:** If \( X_n \) is ergodic with stationary distribution \( \Pi \), then \( X_n \to \Pi \) for any initial \( X_0 \).

**Coupling:** If \( X \) and \( Y \) are two R.V.

a **coupling** \((X', Y')\) is a joint distribution defined on the same probability space such that \( X \equiv X' \), \( Y \equiv Y' \).

We often define a coupling with one of two goals

a) \( X' \leq Y' \) stochastic domination

b) minimize \( \|P[X' \neq Y']\) to compare \( X \sim Y \).

**Example:** \( X \sim \text{Bin}(n, p), \ Y \sim \text{Bin}(m, p) \) for \( m > n \).

Show that \( \|P[X \geq x]\) \leq \( \|P[Y \geq x]\)

Let \( W_i \) be \IID \( \text{Ber}(p) \),

\[ X' = \sum_{i=1}^{n} W_i \equiv X, \quad Y' = \sum_{i=1}^{m} W_i \equiv Y \]

So \( Y' = X' + \sum_{i=n+1}^{m} W_i \geq X' \).

\( \|P[Y \geq x]\) = \( \|P[Y' \geq x]\) \geq \|P[X' \geq x]\) = \( \|P[X \geq x]\). \)
Let \( X_0 = x_0 \), we will prove that \( X_n \xrightarrow{d} \pi \).

Let \( Y_n \) be an independent copy of the chain, \( Y_0 \sim \pi \). Let \( T = \min \{ n \geq 0 : X_n = Y_n \} \).

Let \( Z_n = \begin{cases} X_n & T \leq n \\ Y_n & T > n \end{cases} \)

Then \( Z_n \) is a Markov chain with the same distribution as \( X_n \) and \( \Pr[X_n = x] = \Pr[Z_n = x] \).

For some large \( M \), \( \min_{x, y} P_{xy}^M = \alpha > 0 \).

We can check every \( M \) steps to see if \( T \) has happened.

Then \( \Pr[T > (l+1)M \mid T > 2M] \)

\[ \leq \max_{x \neq x'} \Pr[X_{(l+1)M} = X_{(l+1)M} \mid X_{(l+1)M} = x, Y_{(l+1)M} = y] \]

\[ \leq 1 - \min_{x \neq x'} \Pr[X_{(l+1)M} = Y_{(l+1)M} = x \mid X_{(l+1)M} = x, Y_{(l+1)M} = y] \]

\[ \leq 1 - \min_{x \neq x'} \Pr[X_{(l+1)M} = Y_{(l+1)M} = x \mid X_{(l+1)M} = x, Y_{(l+1)M} = y] \]
\[ 1 - \min_{x \neq y} P_{X,Y}(x, y) \leq 1 - \min_{x \neq y} P_{X}^{-1} P_{Y}^{-1} \leq 1 - \alpha^2 \]

So \( \Pr[T > 2M] \leq (1 - \alpha)^2 \).

\[ \Rightarrow \Pr[T > n] \to 0. \]

Now \( |\Pr[X_n = x] - \pi(x)| = |\Pr[Z_n = x] - \Pr[X_n = x]| \)

\[ \leq \Pr[Z_n \neq X_n] \]

\[ \leq \Pr[T > n] \to 0 \]

So \( X_n \xrightarrow{d} \pi \).

Total Variation Distance

\[ d_{TV}(M, \nu) = \max_A |M(A) - \nu(A)| \]

\[ = \sum_x \frac{1}{2} |M_x - \nu_x| \]

Optimal coupling of \( X \sim M, Y \sim \nu \)

\[ \Pr[X' \neq Y'] = d_{TV}(M, \nu) \]

Proof: For any coupling

\[ \Pr[X' \neq Y'] \geq \Pr[X \in A] - \Pr[Y \in A] \]

\[ = d_{TV}(M, \nu) \text{ for some } A \]
Let \( p = 1 - d_{TV}(\mathcal{M}, \mathcal{N}) \), \( Z \sim \text{Bern}(p) \)

\[
\Theta_1 = \frac{\frac{M + N}{p}}{\frac{M + N}{1 - p}}, \\
\Theta_2 = \frac{M - MN}{1 - p} \quad \text{probability measures}, \\
\Theta_3 = \frac{N - MN}{1 - p}
\]

Let \( U_i \sim \Theta_i \). Then set

\[
X' = Z U_1 + (1 - Z) U_2 \\
Y' = Z U_1 + (1 - Z) U_3
\]

\[
\Pr[X' = Y'] \geq \Pr[Z = 1] = 1 - d_{TV}(\mathcal{M}, \mathcal{N})
\]

so \( \Pr[X' \neq Y'] \leq d_{TV}(\mathcal{M}, \mathcal{N}) \).

Need to check \( X' \sim \mathcal{M}, \ Y' \sim \mathcal{N} \)

**Case 1** \( M_k \geq N_k \)

\[
\Pr[X' = k] = \Pr[Z = 1] \cdot \Pr[U_1 = k] + \Pr[Z = 0] \cdot \Pr[U_2 = k] \\
= p \cdot \frac{M_k + N_k}{p} + (1 - p) \frac{M_k - M_k + N_k}{1 - p} \geq M_k \ \checkmark
\]