Stationary distributions

Sunday, October 15, 2017 11:39 PM

A Markov chain with transition matrix P_{xy} has stationary distribution π if,

$$\pi P = \pi$$

To interpret this if $X_0 \sim \nu$, that is $\mathbb{P}[X_0 = k] = \nu_k$

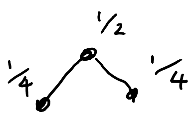
then $\mathbb{P}[X_1 = j] = \sum_k \mathbb{P}[X_1 = j \mid X_0 = k] \cdot \mathbb{P}[X_0 = k]$

$$= \sum_k \nu_k \cdot P_{kj} = (\nu P)_j$$

so $X_1 \sim \nu P$. In general $X_n \sim \nu P^n$.

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

<u>Example</u>: RW on a graph.



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

$$(\pi P)_j = \sum_i \frac{d_i}{2|E|} \cdot P_{ij} = \sum_i \frac{d_i}{2|E|} \cdot \frac{1}{d_i} I(i \sim j)$$

$$= \frac{1}{2E} \sum_i I(i \sim j) = \frac{d_j}{2E}$$

- Random to top shuffle

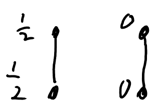
Let $g_k = (1\ 2 \ldots k)$ and G_n uniform on $\{g_k : 1 \le k \le n\}$.

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 π unique? Does $X_n \to \pi$. How fast?

- Example: RW on disconnected graph



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

Perron - Frobenius Theorem

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

Proof Linear Algebra.

Probabilistic Existence proof:

Let $\mu_n = \frac{1}{n}\sum_{i=1}^{n} \mu P^i$.

Now $\mu_n P - \mu_n = \frac{1}{n}\mu(P^{n+1} - P) \to 0$

$\mu \in \{v \in [0,1]^L : \sum_{i=1}^{L} v_i = 1\}$ compact set so

$\exists\, n_k$ such that $\mu_{n_k} \to \tilde{\mu}$.

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

$\Rightarrow \tilde{\mu}$ is stationary.

<u>Positivity</u>: We must have $\mu_i > 0$ for some i.

For any j, $\exists n$ such that $(P^n)_{ij} > 0$.

$\mu_j = (\mu P^n)_j \geq \mu_i P^n_{ij} > 0$.

<u>Uniqueness</u>: Let $S = \inf\{n \geq 1 : X_n = i\}$.

Then $\mu_i = (\mathbb{E}[S \mid X_0 = i])^{-1}$.

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

$\Rightarrow \frac{T_n}{n} \to \mathbb{E}S = \frac{1}{\mu_i}$. a.s.

If $N_n = \#\{1 \leq t \leq n : X_t = i\}$ then

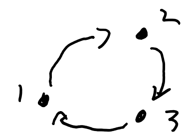
$\frac{N_n}{n} \to \mu_i$. a.s.

So $\mathbb{E}(N_n)/n \to \mu_i$

But $\mathbb{E} N_n = \sum_{t=1}^{n} \mathbb{P}[X_t = i] = n \nu_i$

$\Rightarrow \mu_i = \nu_i$

<u>Periodicity</u>:



$\pi(x) = \frac{1}{3}$ uniform

$\mathbb{P}[X_{3n} = 1 \mid X_0 = 1] = 1$

$\mathbb{P}[X_{3n+1} = 2 \mid X_0 = 1] = 1$ So $X_n \not\xrightarrow{d} \pi$.

A state x in a Markov Chain is <u>aperiodic</u> if $GCD(S) = 1$ where $S = \{n \geq 1 : \mathbb{P}[X_n = x \mid X_0 = x] > 0\}$.

<u>Claim</u> Closed under Addition: If $n, m \in S$ then $n+m \in S$

$$P^{n+m}_{xx} = \sum_y P^n_{xy} P^m_{yx} \geq P^n_{xx} P^m_{xx} > 0.$$

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

<u>Defn</u> A Markov chain is <u>ergodic</u> if it is irreducible and aperiodic.

<u>Claim</u>: If X_n is ergodic then $\exists N$ such that $\forall x, y$, $n \geq M$ then $P^n_{xy} > 0$.

<u>Proof</u>: Suppose z is aperiodic so $\forall m \geq M$ $P^m_{zz} > 0$. Now for some k, l

$p_{xz}^{k} > 0, \quad p_{zy}^{\ell} > 0.$

$$\forall\, n \geq k+\ell+M, \quad p_{xy}^{n} \geq p_{xz}^{k}\, p_{zz}^{n-k-\ell}\, p_{zy}^{\ell} > 0.$$

Theorem: If X_n is ergodic with stationary distribution π then $X_n \xrightarrow{d} \pi$ 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 $X \overset{d}{=} X'$, $Y \overset{d}{=} Y'$.

We often define a coupling with one of two goals

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

b) minimize $\mathbb{P}[X' \neq Y']$ to compare X & Y.

Example: $X \sim Bin(n, p)$, $Y \sim Bin(m, p)$ for $m > n$.
Show that $\mathbb{P}[X \geq x] \leq \mathbb{P}[Y \geq x]$

Let W_i be IID $Ber(p)$,

$$X' = \sum_{i=1}^{n} W_i \overset{d}{=} X, \qquad Y' = \sum_{i=1}^{m} W_i \overset{d}{=} Y$$

so $Y' = X' + \sum_{i=n+1}^{m} W_i \geq X'$.

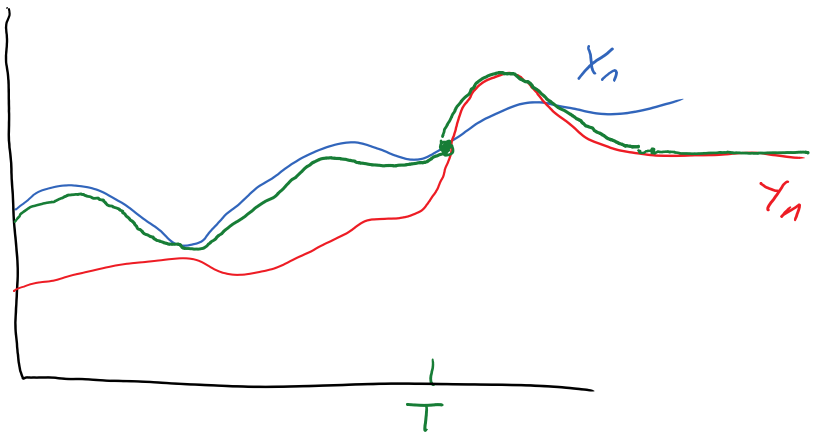
$$\mathbb{P}[Y \geq x] = \mathbb{P}[Y' \geq x] \geq \mathbb{P}[X' \geq x] = \mathbb{P}[X \geq x].$$

$$\mathbb{P}[Y \geq x] = \mathbb{P}[Y' \geq x] \geq \mathbb{P}[X' \geq x] = \mathbb{P}[X \geq x].$$

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 $\mathbb{P}[X_n = x] = \mathbb{P}[Z_n = x]$.

For some large M, $\min_{x,y} P^M_{xy} = \alpha > 0$.

We can check every M steps to see if T has happened

Then $\mathbb{P}[T > (\ell+1)M \mid T > \ell M]$

$$\leq \max_{x \neq x'} \mathbb{P}[X_{(\ell+1)M} \neq Y_{(\ell+1)M} \mid X_{\ell M} = x, Y_{\ell M} = y]$$

$$\leq 1 - \min_{x \neq x'} \mathbb{P}[X_{(\ell+1)m} = Y_{(\ell+1)m} = x \mid X_{\ell m} = x, Y_{\ell m} = y]$$

$$\leq 1 - \min_{x \neq x'} \mathbb{P}[X_{(\ell+1)m} = Y_{(\ell+1)m} = x \mid X_{\ell m} = x, Y_{\ell m} = y]$$

$$\leq 1 - \min_{x \neq x'} P^M_{xx} P^M_{yx} \leq 1 - \alpha^2$$

So $\mathbb{P}[T > \ell M] \leq (1-\alpha)^\ell$.

$\Rightarrow \mathbb{P}[T > n] \to 0$.

Now $|\mathbb{P}[X_n = x] - \pi(x)| = |\mathbb{P}[Z_n = x] - \mathbb{P}[Y_n = x]|$

$$\leq \mathbb{P}[Z_n \neq Y_n]$$

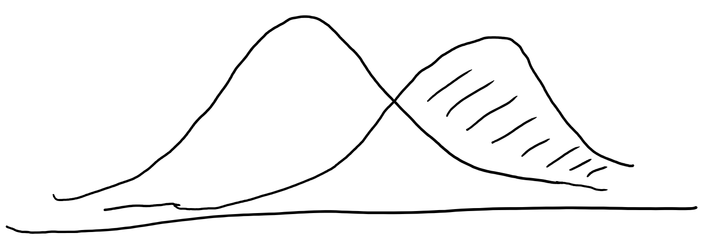
$$\leq \mathbb{P}[T > n] \to 0$$

So $X_n \xrightarrow{d} \pi$.

Total variation Distance

$$d_{TV}(\mu, \nu) = \max_A |\mu(A) - \nu(A)|$$

$$= \sum_x \tfrac{1}{2} |\mu_x - \nu_x|$$



Optimal coupling of $X \sim \mu$, $Y \sim \nu$

$\mathbb{P}[X' \neq Y'] = d_{TV}(\mu, \nu)$

Proof: For any coupling

$$\mathbb{P}[X' \neq Y'] \geq \mathbb{P}[X \in A] - \mathbb{P}[Y \in A]$$

$$= d_{TV}(\mu, \nu) \text{ for some } A$$

Let $p = 1 - d_{TV}(\mu, \nu)$, $Z \sim \mathrm{Ber}(p)$

$$\Theta_1 = \frac{\mu \wedge \nu}{p},$$

$$\Theta_2 = \frac{\mu - \mu \wedge \nu}{1-p} \qquad \text{probability measures.}$$

$$\Theta_3 = \frac{\nu - \mu \wedge \nu}{1-p}$$

Let $W_i \sim \Theta_i$. Then set

$$X' = Z W_1 + (1-Z) W_2$$

$$Y' = Z W_1 + (1-Z) W_3$$

$$\mathbb{P}[X' = Y'] \geq \mathbb{P}[Z = 1] = 1 - d_{TV}(\mu, \nu)$$

so $\mathbb{P}[X' \neq Y'] \leq d_{TV}(\mu, \nu)$.

Need to check $X' \sim \mu$, $Y' \sim \nu$

<u>Case 1</u> $\mu_k \geq \nu_k$

$$\mathbb{P}[X' = k] = \mathbb{P}[Z=1] \cdot \mathbb{P}[W_1 = k] + \mathbb{P}[Z=0] \cdot \mathbb{P}[W_2 = k]$$

$$= p \, \frac{\mu_k \wedge \nu_k}{p} + (1-p) \frac{\mu_k - \mu_k \wedge \nu_k}{1-p} = \mu_k \quad \checkmark$$