Definition: Stationary distribution.

A row vector $S = [s_1, \ldots, s_m]$, $\sum s_i = 1$. $s_i \geq 0$ is a stationary distribution for a Markov chain with transition matrix Q if

$$S Q = S. \qquad Q^T S^T = S^T$$

$P(X_n \mid X_{n-1} = i)$ still depends on ith row of Q.

* Stationary distribution is marginal, not conditional.

$P(X_n) = \sum_i P(X_n \mid X_{n-1} = i)$ is same for all n.

* stationary distribution means the distributions of X_n are all equal, not X_n themselves.

Existence and uniqueness.

Theorem For any irreducible Markov Chain, there exists a unique stationary distribution $S = [s_1, \ldots, s_n]$. $s_i > 0$ for all i. $\sum s_i = 1$.

Proof: Perron Frobenius theorem:

if all entries of a $n \times n$ matrix A are positive

then it has a unique maximal eigenvalue.
Its eigenvector has positive entries.

<u>irreducible</u> Markov chain $\Longrightarrow$ all entries of Q^T are positive

$$Q = \begin{bmatrix} \text{---} Q_1 \text{---} \\ \text{---} Q_2 \text{---} \\ \vdots \\ \text{---} Q_n \text{---} \end{bmatrix} \quad \text{obviously} \quad Q \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

Let $V = \begin{bmatrix} V_1 \\ \vdots \\ V_n \end{bmatrix}$

$$\begin{cases} (Q\,V)_k = \sum_{i=1}^{n} q_{ki} V_i \\ (Q\,V)_k = (\lambda k)_v = \lambda k_v \end{cases}$$

$$\Longrightarrow |\lambda k_v| = \left| \sum_{i=1}^{n} q_{ki} V_i \right| \leq \max_i |k_i|$$

$$|\lambda| \leq 1$$

$\Longrightarrow$ λ is the maximal eigenvalue of Q.

Q^T has the same eigenvalue of Q.

by PF theorem

$\Longrightarrow$ $\exists$ a unique maximal eigenvalue S.

$$Q^T S^T = S^T, \quad S = [S_1, \dots, S_n], \; S_i > 0. \; \forall i.$$

$$\sum S_i = 1.$$

Convergence.

Theorem 11.3.6 (Convergence to stationary distribution). Let $X_0, X_1, \ldots$ be an irreducible, aperiodic Markov chain with stationary distribution $\mathbf{s}$ and transition matrix Q. Then $P(X_n = i)$ converges to s_i as $n \to \infty$. In terms of the transition matrix, Q^n converges to a matrix in which each row is $\mathbf{s}$.

Theorem 11.3.8 (Expected time to return). Let $X_0, X_1, \ldots$ be an irreducible Markov chain with stationary distribution $\mathbf{s}$. Let r_i be the expected time it takes the chain to return to i, given that it starts at i. Then $s_i = 1/r_i$.

Proof: set H_j^n be average number of times in state j in the first n steps

$$Y_i = \begin{cases} 0, & \text{if } X_i \neq j \\ 1, & \text{if } X_i = j. \end{cases}$$

$$H_j^n = Y_0 + Y_1 + \cdots + Y_n$$

$$\Rightarrow \lim_{n\to\infty} P\left(\left|\frac{H_j^n}{n+1} - S_j\right| > \varepsilon\right) = 0. \quad \forall \varepsilon > 0.$$

average prob. $\qquad \forall j = 1, 2, \ldots$

Let r_j be the average number of steps to return to j.

$$\Rightarrow H_j^n = \frac{n+1}{r_j} \approx (n+1) S_j.$$

$$\Rightarrow \frac{1}{r_j} = S_j \quad \text{as } n \to +\infty.$$

Reversibility.

Definition 11.4.1 (Reversibility). Let $Q = (q_{ij})$ be the transition matrix of a Markov chain. Suppose there is $\mathbf{s} = (s_1, \ldots, s_M)$ with $s_i \geq 0, \sum_i s_i = 1$, such that

P(第一步i 第二步j) = P(第一步j, 第二步i)

$$s_i q_{ij} = s_j q_{ji}$$

先i后j与先j后i概率相同.

for all states i and j. This equation is called the *reversibility* or *detailed balance* condition, and we say that the chain is *reversible* with respect to $\mathbf{s}$ if it holds.

Proposition 11.4.2 (Reversible implies stationary). Suppose that $Q = (q_{ij})$ is the transition matrix of a Markov chain that is reversible with respect to a nonnegative vector $\mathbf{s} = (s_1, \ldots, s_M)$ whose components sum to 1. Then $\mathbf{s}$ is a stationary distribution of the chain.

Proof: $\sum_i s_i q_{ij} = \sum_i s_j q_{ji} = s_j$.

$\Rightarrow s \cdot Q = s$

$\Rightarrow s$ is a stationary distribution.

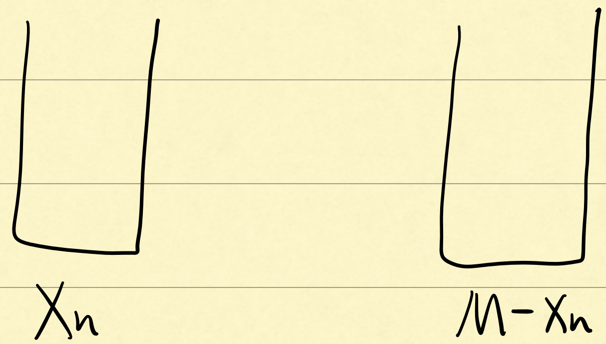
Proposition 11.4.3. If each column of the transition matrix Q sums to 1, then the uniform distribution over all states, $(1/M, 1/M, \ldots, 1/M)$, is a stationary distribution. (A nonnegative matrix such that the row sums and the column sums are all equal to 1 is called a *doubly stochastic matrix.*)

Proof: $v = [1 \cdots 1]$ satisfy $v \cdot Q = v$.

$\Rightarrow v' = [\frac{1}{M}, \cdots \frac{1}{M}]$ satisfy $v' \cdot Q = v'$.

$\Rightarrow v'$ is stationary distribution.

Example 11.4.7 (Ehrenfest)**.** There are two containers with a total of M distinguishable particles. Transitions are made by choosing a random particle and moving it from its current container into the other container. Initially, all of the particles are in the second container. Let X_n be the number of particles in the first container at time n, so $X_0 = 0$ and the transition from X_n to X_{n+1} is done as described above. This is a periodic Markov chain with state space $\{0, 1, \ldots, M\}$.



$S = (S_0, S_1, \ldots, S_M)$.

it is $Bin(M, \frac{1}{2})$ PMF.

$$S_i = \binom{M}{i} \left(\frac{1}{2}\right)^M$$

$j = i+1$

$$S_i P_{ij} = \binom{M}{i} \left(\frac{1}{2}\right)^M \cdot \frac{M-i}{M} = \binom{M-1}{i} \left(\frac{1}{2}\right)^M$$

$$S_j P_{ji} = \binom{M}{j} \left(\frac{1}{2}\right)^M \cdot \frac{j}{M} = \binom{M-1}{j-1} \left(\frac{1}{2}\right)^M = S_i P_{ij}.$$

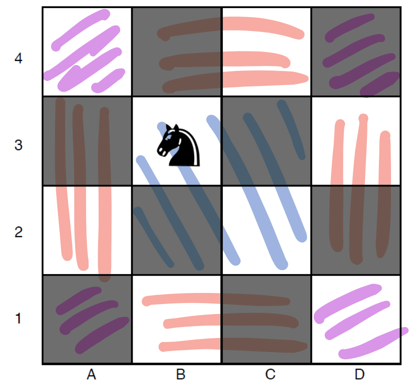
$j = i-1$, Similarly $\Rightarrow S_j P_{ji} = S_i P_{ij}$.

others. $S_j P_{ji} = S_i P_{ij} = 0.$

$\Rightarrow$ S is stationary distribution.

S_i is the long-run proportion of time that the chain spends in state i.

Example 11.4.5 (Knight on a chessboard)**.** Consider a knight randomly moving around on a 4×4 chessboard.



The 16 squares are labeled in a grid, e.g., the knight is currently at the square B3, and the upper left square is A4. Each move of the knight is an L-shaped jump: the knight moves two squares horizontally followed by one square vertically, or vice versa. For example, from B3 the knight can move to A1, C1, D2, or D4; from A4 it can move to B2 or C3. Note that from a light square, the knight always moves to a dark square and vice versa.

Suppose that at each step, the knight moves randomly, with each possibility equally likely. This creates a Markov chain where the states are the 16 squares. Compute the stationary distribution of the chain.

8 edge squares.

4 corner squares. 4 center squares.

There are only three types of squares on the board: 4 center squares, 4 corner squares (such as A4), and 8 edge squares (such as B4; exclude corner squares from being considered edge squares). We can consider the board to be an undirected network where two squares are connected by an edge if they are accessible via a single knight's move. Then a center square has degree 4, a corner square has degree 2, and an edge square has degree 3, so their stationary probabilities are $4a, 2a, 3a$ respectively for some a.

$$4a \times 4 + 2a \times 4 + 3a \times 8 = 48a = 1.$$

$$a = \frac{1}{48}.$$

$$S = [\frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{12}, \frac{1}{24}, \frac{1}{24}, \frac{1}{24}, \frac{1}{24}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}, \frac{1}{16}].$$