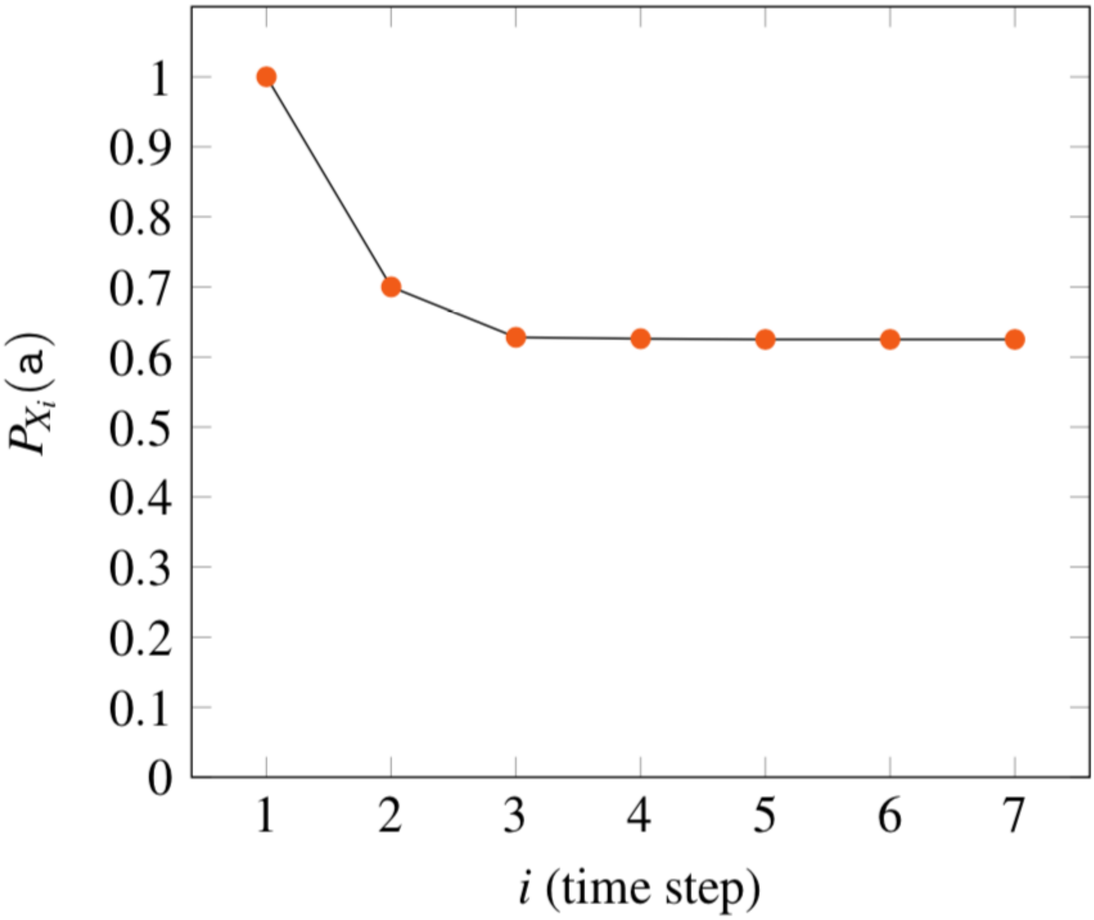
Markov Process: Stationary Distribution



Suppose that we run the process of Example 2 for a very large number of steps, and wonder what the probability will be of observing an $\mathtt{a}$ at the next step. Given the initial distribution and the state diagram, we can compute the probability distribution for every X_i. In the figure above, $P_{X_i}(\mathtt{a})$ is plotted for several values of i. The probability to observe an $\mathtt{a}$ seems to stabilize. This leads us to the following definition:

Definition: Stationary distribution

A stationary distribution for a time-invariant Markov chain is a distribution P_{X_n} such that $P_{X_{n+1}} = P_{X_n}$.

If the initial distribution of a time-invariant Markov process is stationary, then the entire process is stationary as defined previously.

Proposition

Every time-invariant finite-state Markov process has a stationary distribution.

Proof

Let $k := |\mathcal{X}|$. The $k \times k$ transition matrix R with entries $R_{ij} = P_{X_{n+1}|X_n}(j|i)$ is a stochastic matrix, as for every row i, the sum over columns is $\sum_{j=1}^{k} R_{ij} = 1$. We are interested in finding a vector $v \in \mathbb{R}^k_{\geq 0}$ such that $\|v\| = 1$ and $R^T v = v$. This vector then represents the stationary distribution. Clearly, a possible eigenvector for R is the all-1 vector $w = (1, \ldots, 1)^T$ because $Rw = w$ by definition of a stochastic matrix. Hence, 1 is an eigenvalue of R. As R and R^T have the same eigenvalues, 1 is also an eigenvalue of R^T; let $v \in \mathbb{R}^k$ be the corresponding eigenvector such that $R^T v = v$. If all coordinates of v are non-negative, one can verify that we have found a stationary distribution by renormalizing $v / \sum_{i=1}^{k} v_i$. Otherwise, let us write $v = v^+ - v^-$ with $v^+, v^- \in \mathbb{R}^k_{\geq 0}$, where we put all positive coordinates of v in v^+ and all negative coordinates of v in v^-. Note that $R^T v^+ - R^T v^- = R^T(v^+ - v^-) = R^T v = v = v^+ - v^-$. As all entries of R^T, v^+ and v^- are positive, equality must hold for both the positive and negative parts: $R^T v^+ = v^+$ and $R^T v^- = v^-$. As either $v^+ \neq 0^k$ or $v^- \neq 0^k$ (otherwise $v = 0^k$, which cannot be the case for an eigenvector), renormalizing that non-zero vector as above yields the stationary distribution.

Given the transition matrix R of a finite-state Markov process, one can find the stationary distribution μ by solving the linear equation $\mu R = \mu$ under the constraint that $\sum_i \mu_i = 1$.

Example 2: A finite-state time-invariant Markov process, continued

The matrix representation of the process above is given by

$$R = \begin{bmatrix} 0.7 & 0.3 \\ 0.5 & 0.5 \end{bmatrix}$$

. Writing out $(\mu_a, \mu_b)R = (\mu_a, \mu_b)$ results in

$$\begin{cases} 0.7\mu_a + 0.5\mu_b = \mu_a \\ 0.3\mu_a + 0.5\mu_b = \mu_b \end{cases}$$

. These are linearly dependent equations, but together with the constraint $\mu_a + \mu_b = 1$, they can be solved to $(\mu_a, \mu_b) = (5/8, 3/8)$.