# Markov Chains [nln61]

Transitions between values of a discrete stochastic variable taking place at discrete times:

$$X = \{x_1, \ldots, x_N\}; \qquad t = s\tau, \quad s = 0, 1, 2, \ldots$$

Notation adapted to accommodate linear algebra:

$$P(x_n, t) \to P(n, s), \quad P(x_n, t_0 + s\tau | x_m, t_0) \to P(n|m; s).$$

Time evolution of initial probability distribution:

$$P(n, s) = \sum_m P(n|m; s) P(m, 0).$$

Nested Chapman-Kolmogorov equations:

$$\begin{aligned} P(n|m; s) &= \sum_i P(n|i; 1) P(i|m; s-1) \\ &= \sum_{ij} P(n|i; 1) P(i|j; 1) P(j|m; s-2) \\ &= \sum_{ijk} P(n|i; 1) P(i|j; 1) P(j|k; 1) P(k|m; s-3) = \ldots \end{aligned}$$

**Matrix representation:**

Transition matrix: $\mathbf{W}$ with elements $W_{mn} = P(n|m; 1)$.

Probability vector: $\vec{P}(s) = \Big(P(1, s), \ldots, P(N, s)\Big)$.

Time evolution via matrix multiplication: $\vec{P}(s) = \vec{P}(0) \cdot \mathbf{W}^s$.

General attributes of transition matrix:

- All elements represent probabilities: $W_{mn} \geq 0$;
  $W_{mm}$: system stays in state $m$;
  $W_{mn}$ with $m \neq n$: system undergoes a transition from $m$ to $n$.
- Normalization of probabilities: $\displaystyle\sum_n W_{mn} = 1$
- A transition $m \to n$ and its inverse $n \to m$ may occur at different rates. Hence $\mathbf{W}$ is, in general, not symmetric.

**Regularity:**

A transition matrix $\mathbf{W}$ is called *regular* if all elements of the matrix product $\mathbf{W}^s$ are nonzero (i.e. positive) for some exponent $s$.

Regularity guarantees that repeated multiplication leads to convergence:

$$\lim_{s\to\infty} \mathbf{W}^s = \mathbf{M} = \begin{pmatrix} \pi_1 & \pi_2 & \cdots & \pi_N \\ \pi_1 & \pi_2 & \cdots & \pi_N \\ \vdots & \vdots & & \vdots \\ \pi_1 & \pi_2 & \cdots & \pi_N \end{pmatrix}$$

Further multiplications have no effect:

$$\mathbf{W} \cdot \mathbf{M} = \begin{pmatrix} W_{11} & \cdots & W_{1N} \\ \vdots & & \vdots \\ W_{N1} & \cdots & W_{NN} \end{pmatrix} \cdot \begin{pmatrix} \pi_1 & \cdots & \pi_N \\ \vdots & & \vdots \\ \pi_1 & \cdots & \pi_N \end{pmatrix} = \mathbf{M}.$$

The asymptotic distribution is stationary.
The stationary distribution does not depend on initial distribution:

$$\lim_{s\to\infty} \vec{P}(s) = \vec{P}(0) \cdot \mathbf{M} = \vec{\pi} = \Big(\pi_1, \pi_2, \ldots, \pi_N\Big).$$

All elements of the stationary distribution are nonzero.

The computation of the stationary distribution $\vec{\pi}$ via repeated multiplication of the transition matrix with itself works well for regular matrices.

More generally, transition matrices may have stationary solutions that depend on the initial distribution or stationary solutions that are not asymptotic solutions of any kind.

**Eigenvalue problem:**

The eigenvalues $\Lambda_1, \ldots, \Lambda_N$ of $\mathbf{W}$ are the solutions of the secular equation:

$$\det\big(\mathbf{W} - \Lambda\mathbf{E}\big) = 0, \quad E_{ij} = \delta_{ij}.$$

For an asymmetric $\mathbf{W}$ not all eigenvalues $\Lambda_n$ are real. We must distinguish between left eigenvectors $\vec{X}_n$ and right eigenvectors $\vec{Y}_n$:

$$\vec{X}_n \cdot \mathbf{W} = \Lambda_n \vec{X}_n, \quad n = 1, \ldots, N \quad \text{with} \quad \vec{X}_n \doteq \big(X_{n1}, \ldots, X_{nN}\big)$$

$$\mathbf{W} \cdot \vec{Y}_n = \Lambda_n \vec{Y}_n, \quad n = 1, \ldots, N \quad \text{with} \quad \vec{Y}_n = \begin{pmatrix} Y_{1n} \\ \vdots \\ Y_{Nn} \end{pmatrix}.$$

The two eigenvector matrices are orthonormal to one another:

$$\mathbf{X} \cdot \mathbf{Y} = \mathbf{E}, \quad \text{where} \quad \mathbf{X} \doteq \begin{pmatrix} \vec{X}_1 \\ \vdots \\ \vec{X}_N \end{pmatrix}, \quad \mathbf{Y} \doteq \big(\vec{Y}_1, \ldots, \vec{Y}_N\big).$$

All eigenvalues $\Lambda_n$ of the transition matrix $\mathbf{W}$ satisfy the condition $|\Lambda_n| \leq 1$.

There always exists at least one eigenvalue $\Lambda_n = 1$.

The right eigenvector for $\Lambda_n = 1$ is $\vec{Y}_n = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

The left eigenvector for $\Lambda_n = 1$ is a stationary distribution $\vec{X}_n = \big(\pi_1, \ldots, \pi_N\big)$.

If $\mathbf{W}$ is regular then the eigenvalue $\Lambda_n = 1$ is unique and its left eigenvector is the asymptotic distribution $\vec{X}_n = \vec{\pi}$, independent of the inital condition.

**Ergodicity:**

In an *ergodic* transition matrix $\mathbf{W}$ any two states are connected, directly or indirectly, by allowed transitions. Regularity implies ergodicity but not vice versa.

A block-diagonal transition matrix,

$$\mathbf{W} = \begin{pmatrix} W_{1,1} & \cdots & W_{1,n} & 0 & \cdots & 0 \\ \vdots & & \vdots & \vdots & & \vdots \\ W_{n,1} & \cdots & W_{n,n} & 0 & \cdots & 0 \\ 0 & \cdots & 0 & W_{n+1,n+1} & \cdots & W_{n+1,N} \\ \vdots & & \vdots & \vdots & & \vdots \\ 0 & \cdots & 0 & W_{N,n+1} & \cdots & W_{N,N} \end{pmatrix}$$

implies non-ergodicity because inter-block transitions are prohibited.

**Absorbing states:**

If there exists a state $n$ that allows only transitions into it but not out of it then row $n$ of the transition matrix has diagonal element $W_{nn} = 1$ and off-diagonal elements $W_{nn'} = 0$ $(n' \neq n)$.

For an ergodic system we then have

$$\lim_{s\to\infty} \vec{P}(0) \cdot \mathbf{W}^s = \vec{\pi} = (0, \ldots, 0, 1, 0, \ldots, 0),$$

with the 1 at position $n$.

**Detailed balance:**

The detailed balance condition postulates the existence of a stationary distribution $\vec{\pi}$ satisfying the relations

$$W_{mn}\pi_m = W_{nm}\pi_n, \quad n, m = 1, \ldots, N.$$

Detailed balance requires that $W_{mn} = 0$ if $W_{nm} = 0$. Microscopic (quantum or classical) dynamics guarantees that this requirement is fulfilled.

The detailed balance condition, if indeed satisfied, can be used to determine the stationary distribution.

**Applications:**

- ▷ House of the mouse: two-way doors only [nex102]
- ▷ House of the mouse: some one-way doors [nex103]
- ▷ House of the mouse: one-way doors only [nex104]
- ▷ House of the mouse: mouse with inertia [nex105]
- ▷ House of the mouse: mouse with memory [nex43]
- ▷ Mixing marbles red and white [nex42]
- ▷ Random traffic around city block [nex86]
- ▷ Modeling a Markov chain [nex87]