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* (or *irreducible*) 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.

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.

In a Markov chain with a regular transition matrix all states are *persistent*. Persistent states are returned to with certainty after a finite number of transitions. that is not the case for *transient* states.

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.

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:

Master equation constructed from[1] $P(n, s+1) = \sum_m P(m, s)W_{mn}$.

$$P(n, s+1) - P(n, s) = \sum_m \Big[W_{mn}P(m, s) - W_{nm}P(n, s)\Big].$$

Stationarity condition, $P(n, s+1) - P(n, s) = 0$, in general, requires that

$$\sum_m \Big[W_{mn}P(m, s) - W_{nm}P(n, s)\Big] = 0.$$

The detailed-balance condition postulates the existence of a stationary distribution for which each term inside the bracket vanishes separately:

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

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.

Not all stationary distributions are equilibrium distributions. Stationary nonequilibrium distributions can be found in [nex103] and [nex104].

[1] Use $P(n, s) = P(n, s) \sum_m W_{nm}$.

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 $\Lambda^{(k)}$ are real. We must distinguish between left eigenvectors $\vec{X}^{(k)}$ and right eigenvectors $\vec{Y}^{(k)}$:

$$\vec{X}^{(k)} \cdot \mathbf{W} = \Lambda^{(k)}\vec{X}^{(k)}, \quad k = 1, \ldots, N \quad \text{with} \quad \vec{X}^{(k)} \doteq \big(X_1^{(k)}, \ldots, X_N^{(k)}\big)$$

$$\mathbf{W} \cdot \vec{Y}^{(k)} = \Lambda^{(k)}\vec{Y}^{(k)}, \quad k = 1, \ldots, N \quad \text{with} \quad \vec{Y}^{(k)} = \begin{pmatrix} Y_1^{(k)} \\ \vdots \\ Y_N^{(k)} \end{pmatrix}.$$

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 satisfy the condition $|\Lambda^{(k)}| \leq 1$. There always exists at least one eigenvalue $\Lambda^{(k)} = 1$.

Right eigenvector for $\Lambda^{(k)} = 1$: $\vec{Y}^{(k)} = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$.

Left eigenvector for $\Lambda^{(k)} = 1$: stationary distribution $\vec{X}^{(k)} = \big(\pi_1, \ldots, \pi_N\big)$.

If $\mathbf{W}$ is regular then the eigenvalue $\Lambda^{(k)} = 1$ is unique and its left eigenvector is the asymptotic distribution $\vec{X}^{(k)} = \vec{\pi}$, independent of the initial condition.

The components of left eigenvectors pertaining to eigenvalues $|\Lambda^{(k)}| < 1$ are, in general, complex and add up to zero:

$$\sum_{n=1}^{N} X_n^{(k)} = 0, \quad |\Lambda^{(k)}| < 1.$$

The inequality, $|\Lambda^{(k)}| \leq 1$, holds for all eigenvalues, $k = 1, \ldots N$.

- ▷ Consider left eigenvector $\vec{X}^{(k)} \doteq \big(X_1^{(k)}, \ldots, X_N^{(k)}\big)$.
- ▷ Left eigenvalue equation: $\sum_m X_m^{(k)} W_{mn} = \Lambda^{(k)} X_n^{(k)}, \quad n = 1, \ldots, N.$
- ▷ Inequality: $\sum_m \big|X_m^{(k)}\big| W_{mn} \geq \left|\sum_m X_m^{(k)} W_{mn}\right| = \big|\Lambda^{(k)} X_n^{(k)}\big| = \big|\Lambda^{(k)}\big| \, \big|X_n^{(k)}\big|.$
- ▷ Sum: $\sum_{nm} \big|X_m^{(k)}\big| W_{mn} = \sum_m \big|X_m^{(k)}\big| \geq \big|\Lambda^{(k)}\big| \sum_n \big|X_n^{(k)}\big|.$
- ▷ Consequence: $|\Lambda^{(k)}| \leq 1$.

The structure of $\mathbf{W}$, specifically the normalization condition, $\sum_n W_{mn} = 1$, guarantee the existence of eigenvalue $\Lambda = 1$.

The associated right eigenvector has unit elements and the associated left eigenvector represents a stationary distribution as discussed earlier.

The elements of all left eigenvectors for eigenvalues $|\Lambda^{(k)}| < 1$ add up to zero.

- ▷ Eigenvalue equation: $\sum_m X_m^{(k)} W_{mn} = \Lambda^{(k)} X_n^{(k)}$.
- ▷ Use normalization: $\sum_{mn} X_m^{(k)} W_{mn} = \sum_{nm} X_n^{(k)} W_{nm} = \sum_n X_n^{(k)}$.
- ▷ Consequence: $\sum_n X_n^{(k)} = \Lambda^{(k)} \sum_n X_n^{(k)} \quad \Rightarrow \quad \big(\Lambda^{(k)} - 1\big) \sum_n X_n^{(k)} = 0.$
- ▷ Implication: Either $\Lambda^{(k)} = 1$ or $\sum_n X_n^{(k)} = 0$.

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]