Markov chains

- in general, it is a sequence of possible events

or random quantities $S^{(k)}$, $k = 1, \dots, \infty$

which are chosen from a certain set of states (configurations) $\{A_i\}$, $i = 1, \dots, M$ (could be infinite, but we will consider M to be finite for simplicity)

but occurrences of states in the chain are not independent, they depend on the previous state:

if in „time" k state A_i appears with probability $\pi_i^{(k)}$

then in „time" k+1 state A_j appears with prob. $\pi_j^{(k+1)}$

given by $$\pi_j^{(k+1)} = \sum_{i=1}^{M} \pi_i^{(k)} W_{i \to j}$$

or written as vectors and using the transition matrix W with elements $W_{i \to j} = P(A_i \to A_j)$ = probability of transition from A_i to A_j

$$\pi^{(k+1)} = \pi^{(k)} \cdot W$$

W must satisfy condition

$$\sum_{j=1}^{M} W_{i \to j} = 1 \quad \text{for all } i$$

(sum of probabilities of going from A_i to anywhere)

- examples in Mathematica

1) Problematic network - two-state system

see Nezbeda, Kolafa, Kotrla, page 46

2) Random flipping of a die over an edge

- in both examples above, we can see that after a while the propability π_j of A_j is independent of the previous states

$$\pi^{(k+1)} \simeq \pi^{(k)} \quad \text{for } k \text{ large} \qquad \left(\text{"the system is losing memory"}\right)$$

or $$\lim_{n\to\infty} \pi^{(n)} = \pi$$

where π satisfies the equation for left eigenvalues and eigenvectors of the transition matrix

$$\lambda \pi = \pi \cdot W$$

and π is the eigenvector corresponding to $\lambda = 1$:

$$\pi = \pi \cdot W$$

[Note: It can be shown that for transition matrices the other eigenvalues are less than 1 (in absolute value)]

- under which conditions this happens?

It can be shown that (see e.g. W. Feller: An Introduction to Probability Theory and Its Applications)
if 1) all states (configurations) can be obtained (reached) from an arbitrary state in a finite time with non-zero probability

2) and no state is periodic

(A_i is periodic, if there is a period m such that for $\pi_i^{(k)} = 0$ will be $\pi_i^{(k+m)} = 0$ and for $\pi_i^{(k)} \neq 0$ will be $\pi_i^{(k+m)} \neq 0$)

then for arbitrary initial distribution of probabilities $\pi^{(1)}$ there exists a limit

$$\pi = \lim_{k\to\infty} \pi^{(k)}$$

and π satisfies the equation

$$\pi = \pi \cdot W$$

Such a system and a set of states are called ergodic.

- in Monte Carlo simulations we do not have the transition matrix, but only the limiting distribution of probabilities give usually as

$$\pi_j = g(A_j) = \frac{e^{-\beta H(A_j)}}{Z}, \quad Z = \sum_j e^{-\beta H(A_j)}$$

but it can be more general $g(A_j)$, of course.

How to construct the transition matrix W ?

- we have 3 conditions:

$$W_{i\to j} \geq 0 \quad \text{for all } i,j = 1,\dots,M$$

$$\sum_{j=1}^{M} W_{i\to j} = 1 \quad \text{for all } i = 1,\dots,M \qquad (*)$$

$$\pi . W = \pi$$

- the last condition is called the detailed balance and it is a necessary condition to get the limit of stationary propability distribution. It will be valid if we suppose the stronger condition of microscopic reversibility

$$\pi_i W_{i\to j} = \pi_j W_{j\to i} \qquad (**)$$

because

$$\pi . W \Leftarrow \sum_{i=1}^{M} \pi_i W_{i\to j} = \pi_j \underbrace{\sum_{i=1}^{M} W_{j\to i}}_{1} = \pi_j \Leftrightarrow \pi$$

- conditions (*) do not determine W uniquely (we have 2M equations for M×M unknowns)

from (**) we have either

$$\frac{W_{j\to i}}{W_{i\to j}} = \frac{\pi_i}{\pi_j} = \frac{g(A_i)}{g(A_j)}$$

(it is independent of normalization (Z)!)

or $W_{j\to i} = W_{i\to j} = 0$ (if $g(A_i) > 0$ for all A_i as in the case when $g(A_i) = e^{-\beta H(A_i)}/Z$)

but still there is many posibilities how to choose elements $W_{i\to j}$

a) Metropolis method, the most common choice of W

$$W_{i\to j} = \begin{cases} \alpha_{i\to j} & \text{for } i\neq j,\ \pi_j \geq \pi_i \\ \alpha_{i\to j}\dfrac{\pi_j}{\pi_i} & \text{for } i\neq j,\ \pi_j < \pi_i \\ 1 - \sum\limits_{k\neq i} W_{i\to k} & \text{for } i=j \end{cases}$$

where $\alpha_{i\to j}$ is an arbitrary symmetric stochastic matrix satisfying $\alpha_{i\to j} \geq 0$

- the first two conditions of (*) are satisfied thanks to $\sum_{j=1}^{n} \alpha_{i\to j} = 1$ (necessary for $W_{i\to i} \geq 0$)

and moreover we have for $\pi_j \geq \pi_i$

$$\pi_i W_{i\to j} = \pi_i \alpha_{i\to j} \overset{\text{symmetric } \alpha}{=} \pi_j \alpha_{j\to i} \frac{\pi_i}{\pi_j} = \pi_j W_{j\to i} \quad \checkmark$$

and for $\pi_j < \pi_i$

$$\pi_i W_{i\to j} = \pi_i \alpha_{i\to j} \frac{\pi_j}{\pi_i} = \pi_j \alpha_{j\to i} = \pi_j W_{j\to i} \quad \checkmark$$

b) Barker (Glauber) transition matrix:

$$W_{i\to j} = \begin{cases} \alpha_{i\to j}\dfrac{\pi_j}{\pi_i+\pi_j} & \text{for } i\neq j \\ 1 - \sum\limits_{k\neq i} W_{i\to k} & \text{for } i=j \end{cases}$$

where again $\alpha_{i\to j}$ is an arbitrary symmetric stochastic matrix

- now

$$\pi_i W_{i\to j} = \pi_i \alpha_{i\to j} \frac{\pi_j}{\pi_i+\pi_j} \overset{\text{symmetric } \alpha}{=} \pi_j \alpha_{j\to i} \frac{\pi_i}{\pi_i+\pi_j} = \pi_j W_{j\to i} \quad \checkmark$$

- there are other methods, e.g. heat-bath method (see Creutz et al., Phys. Rev. Lett. 42 (1979) 1390)

- how to choose $\alpha_{i\to j}$?
  - that depends on a particular system we have
  - these are probabilities of generating A_j from the configuration A_i and there are different strategies (e.g. changing just one site in the lattice or moving one molecule randomly etc.)