Markov Processes

- Stochastic process
  - *movement through a series of well-defined states in a way that involves some element of randomness*
  - *for our purposes, "states" are microstates in the governing ensemble*
- Markov process
  - *stochastic process that has no memory*
  - *selection of next state depends only on current state, and not on prior states*
  - *process is fully defined by a set of <u>transition probabilities</u>* π_{ij}

    π_{ij} = probability of selecting state j next, given that presently in state i.

    Transition-probability matrix Π collects all π_{ij}

Transition-Probability Matrix

- Example
  - *system with three states*

If in state 1, will stay in state 1 with probability 0.1

If in state 1, will move to state 3 with probability 0.4

$$\Pi \equiv \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{pmatrix} = \begin{pmatrix} \boxed{0.1} & 0.5 & \boxed{0.4} \\ 0.9 & 0.1 & \boxed{0.0} \\ 0.3 & 0.3 & 0.4 \end{pmatrix}$$

Never go to state 3 from state 2

- Requirements of transition-probability matrix
  - *all probabilities non-negative, and no greater than unity*
  - *sum of each row is unity*
  - *probability of staying in present state may be non-zero*

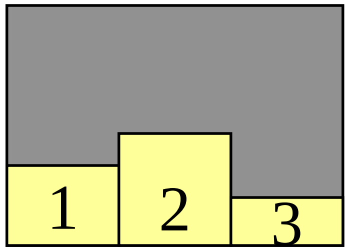
Distribution of State Occupancies

- Consider process of repeatedly moving from one state to the next, choosing each subsequent state according to Π
  - $1 \rightarrow 2 \rightarrow 2 \rightarrow 1 \rightarrow 3 \rightarrow 2 \rightarrow 2 \rightarrow 3 \rightarrow 3 \rightarrow 1 \rightarrow 2 \rightarrow 3 \rightarrow etc.$
- Histogram the occupancy number for each state
  - $n_1 = 3$
  - $n_2 = 5$
  - $n_3 = 4$

$\pi_1 = 0.33$

$\pi_2 = 0.42$

$\pi_3 = 0.25$



- After very many steps, a limiting distribution emerges
- Click here for an applet that demonstrates a Markov process and its approach to a limiting distribution

The Limiting Distribution 1.

- Consider the product of Π with itself

$$\Pi^2 \equiv \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{pmatrix} \times \begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} \\ \pi_{21} & \pi_{22} & \pi_{23} \\ \pi_{31} & \pi_{32} & \pi_{33} \end{pmatrix}$$

All ways of going from state 1 to state 2 in two steps

$$= \begin{pmatrix} \pi_{11}\pi_{11} + \pi_{12}\pi_{21} + \pi_{13}\pi_{31} & \boxed{\pi_{11}\pi_{12} + \pi_{12}\pi_{22} + \pi_{13}\pi_{32}} & etc. \\ \pi_{21}\pi_{11} + \pi_{22}\pi_{21} + \pi_{23}\pi_{31} & \pi_{21}\pi_{12} + \pi_{22}\pi_{22} + \pi_{23}\pi_{32} & etc. \\ \pi_{31}\pi_{11} + \pi_{32}\pi_{21} + \pi_{33}\pi_{31} & \boxed{\pi_{31}\pi_{12} + \pi_{32}\pi_{22} + \pi_{33}\pi_{32}} & etc. \end{pmatrix}$$

Probability of going from state 3 to state 2 in two steps

- In general Π^n is the n-step transition probability matrix
  - *probabilities of going from state* i *to* j *in exactly n steps*

$$\Pi^n \equiv \begin{pmatrix} \pi_{11}^{(n)} & \pi_{12}^{(n)} & \pi_{13}^{(n)} \\ \pi_{21}^{(n)} & \pi_{22}^{(n)} & \pi_{23}^{(n)} \\ \pi_{31}^{(n)} & \pi_{32}^{(n)} & \pi_{33}^{(n)} \end{pmatrix}$$

defines $\pi_{ij}^{(n)}$

The Limiting Distribution 2.

- Define $\pi_i^{(0)}$ as a unit state vector

$$\pi_1^{(0)} = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \quad \pi_2^{(0)} = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \quad \pi_3^{(0)} = \begin{pmatrix} 0 & 0 & 1 \end{pmatrix}$$

- Then $\pi_i^{(n)} \equiv \pi_i^{(0)} \Pi^n$ is a vector of probabilities for ending at each state after n steps if beginning at state i

$$\pi_1^{(n)} = \pi_1^{(0)} \Pi^n \equiv \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} \pi_{11}^{(n)} & \pi_{12}^{(n)} & \pi_{13}^{(n)} \\ \pi_{21}^{(n)} & \pi_{22}^{(n)} & \pi_{23}^{(n)} \\ \pi_{31}^{(n)} & \pi_{32}^{(n)} & \pi_{33}^{(n)} \end{pmatrix} = \begin{pmatrix} \pi_{11}^{(n)} & \pi_{12}^{(n)} & \pi_{13}^{(n)} \end{pmatrix}$$

- The limiting distribution corresponds to $n \to \infty$
  - *independent of initial state* $\pi_1^{(\infty)} = \pi_2^{(\infty)} = \pi_3^{(\infty)} \equiv \pi$

The Limiting Distribution 3.

- Stationary property of π

$$\pi = \lim_{n\to\infty}\left[\pi_i^{(0)}\Pi^n\right]$$
$$= \left(\lim_{n\to\infty}\left[\pi_i^{(0)}\Pi^{n-1}\right]\right)\Pi$$
$$= \pi\Pi$$

- π is a left eigenvector of Π with unit eigenvalue
  - *such an eigenvector is guaranteed to exist for matrices with rows that each sum to unity*
- Equation for elements of limiting distribution π

$$\pi_i = \sum_j \pi_j \pi_{ji}$$

$$e.g.\quad \Pi = \begin{pmatrix} 0.1 & 0.5 & 0.4 \\ 0.9 & 0.1 & 0.0 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}$$

$$\pi_1 = 0.1\pi_1 + 0.9\pi_2 + 0.3\pi_3$$
$$\pi_2 = 0.5\pi_1 + 0.1\pi_2 + 0.3\pi_3$$
$$\pi_3 = 0.4\pi_1 + 0.0\pi_2 + 0.4\pi_3$$
$$\pi_1 + \pi_2 + \pi_3 = \pi_1 + \pi_2 + \pi_3$$

not independent

Detailed Balance

- Eigenvector equation for limiting distribution
  - $\pi_i = \sum_j \pi_j \pi_{ji}$
- A sufficient (but not necessary) condition for solution is
  - $\boxed{\pi_i \pi_{ij} = \pi_j \pi_{ji}}$
  - *"detailed balance" or "microscopic reversibility"*
- Thus
  - $$\pi_i = \sum_j \pi_j \pi_{ji} = \sum_j \pi_i \pi_{ij} = \pi_i \sum_j \pi_{ij} = \pi_i$$

$$\Pi = \begin{pmatrix} 0.1 & 0.5 & 0.4 \\ 0.9 & 0.1 & 0.0 \\ 0.3 & 0.3 & 0.4 \end{pmatrix}$$

For a given Π, it is not always possible to satisfy detailed balance; e.g. for this Π

$\pi_3 \pi_{32} \neq \pi_2 \underbrace{\pi_{23}}_{\text{zero}}$

Deriving Transition Probabilities

- Turn problem around...
- …given a desired π, what transition probabilities will yield this as a limiting distribution?
- *Construct transition probabilities* to satisfy detailed balance
- Many choices are possible
  - *e.g.* $\pi = \begin{pmatrix} 0.25 & 0.5 & 0.25 \end{pmatrix}$
  - *try them out*

$$\Pi = \begin{pmatrix} 0.97 & 0.02 & 0.01 \\ 0.01 & 0.98 & 0.01 \\ 0.01 & 0.02 & 0.97 \end{pmatrix}$$

Least efficient

$$\Pi = \begin{pmatrix} 0 & 1 & 0 \\ 0.5 & 0 & 0.5 \\ 0 & 1 & 0 \end{pmatrix}$$

Most efficient

$$\Pi = \begin{pmatrix} 0.42 & 0.33 & 0.25 \\ 0.17 & 0.66 & 0.17 \\ 0.25 & 0.33 & 0.42 \end{pmatrix}$$

Barker

$$\Pi = \begin{pmatrix} 0.0 & 0.5 & 0.5 \\ 0.25 & 0.5 & 0.25 \\ 0.5 & 0.5 & 0.0 \end{pmatrix}$$

Metropolis

Metropolis Algorithm 1.

- Prescribes transition probabilities to satisfy detailed balance, given desired limiting distribution
- Recipe:
  From a state *i*…
  - *with probability* τ_{ij}, choose a trial state j for the move (note: $\tau_{ij} = \tau_{ji}$)
  - *if* $\pi_j > \pi_i$, *accept* j *as the new state*
  - *otherwise, accept state* j *with probability* π_j/π_i

    generate a random number R on (0,1); accept if $R < \pi_j/\pi_i$
  - *if not accepting* j *as the new state, take the present state as the next one in the Markov chain* $(\pi_{ii} \neq 0)$

Metropolis, Rosenbluth, Rosenbluth, Teller and Teller, J. Chem. Phys., ***21*** *1087 (1953)*

Metropolis Algorithm 2.

- What are the transition probabilities for this algorithm?
  - *Without loss of generality, define* i *as the state of greater probability*

$$\pi_i > \pi_j$$

$$\pi_{ij} = \tau_{ij} \times \frac{\pi_j}{\pi_i}$$

$$\left(\text{in general: } \pi_{ij} = \tau_{ij} \min\left(\frac{\pi_j}{\pi_i}, 1\right) \right)$$

$$\pi_{ji} = \tau_{ji}$$

$$\pi_{ii} = 1 - \sum_{j \neq i} \pi_{ij}$$

- Do they obey detailed balance?

$$\pi_i \pi_{ij} \stackrel{?}{=} \pi_j \pi_{ji}$$

$$\pi_i \tau_{ij} \frac{\pi_j}{\pi_i} \stackrel{?}{=} \pi_j \tau_{ji}$$

$$\tau_{ij} = \tau_{ji}$$

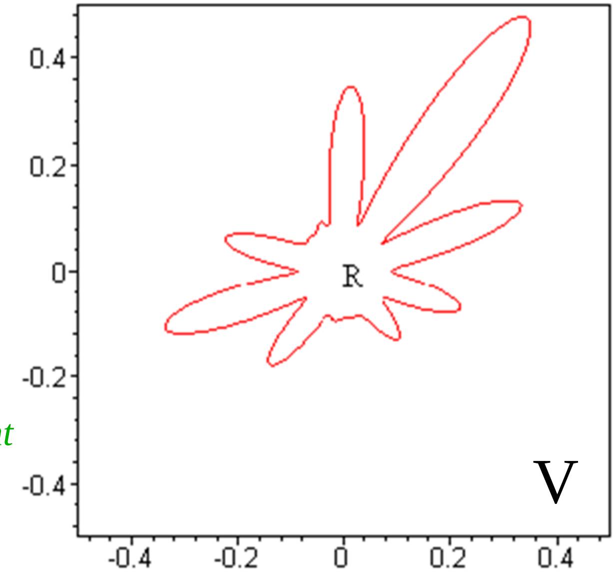
- Yes, as long as the *underlying matrix* T of the Markov chain is symmetric
  - *this can be violated, but acceptance probabilities must be modified*

Markov Chains and Importance Sampling 1.

- Importance sampling specifies the desired limiting distribution
- We can use a Markov chain to generate quadrature points according to this distribution
- Example

$$s = \begin{cases} 1 & \text{inside R} \\ 0 & \text{outside R} \end{cases}$$

$$\langle r^2 \rangle = \frac{\int_{-0.5}^{+0.5} dx \int_{-0.5}^{+0.5} dy (x^2 + y^2) s(x,y)}{\int_{-0.5}^{+0.5} dx \int_{-0.5}^{+0.5} dy s(x,y)} = \frac{\langle r^2 s \rangle_V}{\langle s \rangle_V}$$



q = normalization constant

- *Method 1*: let $\pi_1(x,y) = s(x,y)/q_1$
- *then*

$$\langle r^2 \rangle = \frac{\left\langle \frac{r^2 s}{\pi_1} \right\rangle_{\pi_1}}{\left\langle \frac{s}{\pi_1} \right\rangle_{\pi_1}} = \frac{\langle q_1 r^2 \rangle_{\pi_1}}{\langle q_1 \rangle_{\pi_1}} = \frac{q_1 \langle r^2 \rangle_{\pi_1}}{q_1} = \langle r^2 \rangle_{\pi_1}$$

Simply sum r^2 with points given by Metropolis sampling

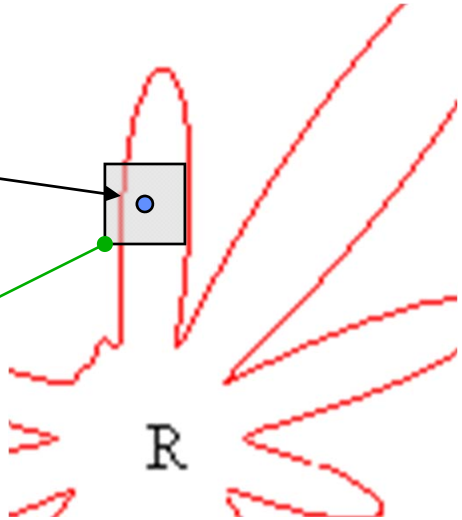
Markov Chains and Importance Sampling 2.

Example (cont'd)

- *Method 2: let* $\pi(x,y) = r^2 s / q_2$
- *then*

$$\left\langle r^2 \right\rangle = \frac{\left\langle \frac{r^2 s}{\pi_2} \right\rangle_{\pi_2}}{\left\langle \frac{s}{\pi_2} \right\rangle_{\pi_2}} = \frac{\left\langle q_2 \right\rangle_{\pi_2}}{\left\langle q_2 / r^2 \right\rangle_{\pi_2}} = \frac{q_2}{q_2 \left\langle 1/r^2 \right\rangle_{\pi_2}} = \frac{1}{\left\langle r^{-2} \right\rangle_{\pi_2}}$$

Algorithm and transition probabilities



- *given a point in the region* R
- *generate a new point in the vicinity of given point*

$x^{new} = x + r(-1,+1)\delta x \quad y^{new} = y + r(-1,+1)\delta y$

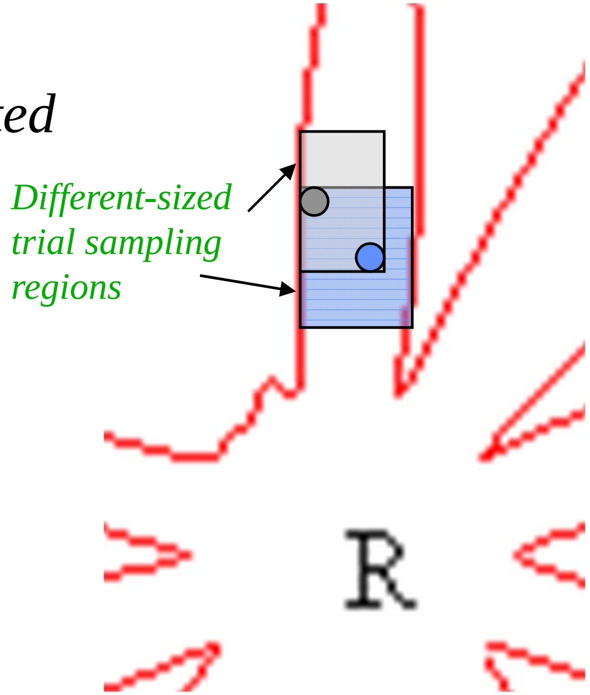
- *accept with probability* $\min(1, \pi^{new} / \pi^{old})$
- *note* $\frac{\pi_1^{new}}{\pi_1^{old}} = \frac{s^{new}/q_1}{s^{old}/q_1} = \frac{s^{new}}{s^{old}}$

Normalization constants cancel!

- *Method 1: accept all moves that stay in* R
- *Method 2: if in* R*, accept with probability* $\left(r^2\right)^{new} \big/ \left(r^2\right)^{old}$

Markov Chains and Importance Sampling 3.

- Subtle but important point
  - *Underlying matrix* T *is set by the trial-move algorithm (select new point uniformly in vicinity of present point)*
  - *It is important that new points are selected in a volume that is independent of the present position*
  - *If we reject configurations outside* R*, without taking the original point as the "new" one, then the underlying matrix becomes asymmetric*

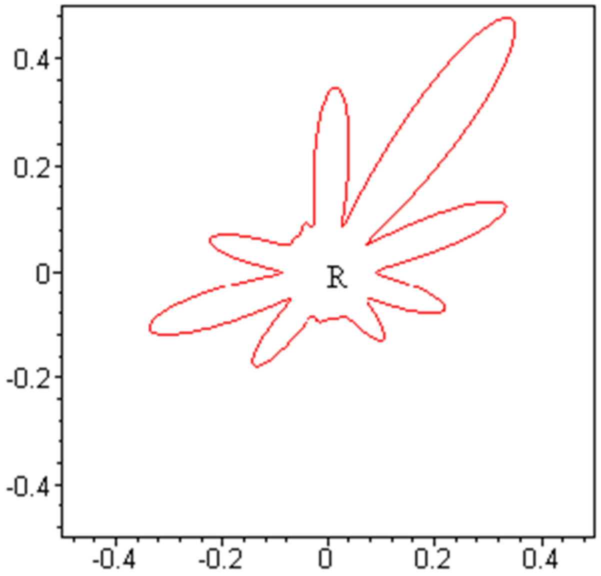


Evaluating Areas with Metropolis Sampling

- What if we want the absolute area of the region R, not an average over it?

$$A = \int_{-0.5}^{+0.5} dx \int_{-0.5}^{+0.5} dy s(x,y) = \langle s \rangle_V$$

  - *Let* $\pi_1(x,y) = s(x,y)/q_1$
  - *then* $A = \left\langle \frac{s}{\pi_1} \right\rangle_{\pi_1} = \langle q_1 \rangle_{\pi_1} = q_1$
  - *We need to know the normalization constant* q_1
  - *but this is exactly the integral that we are trying to solve!*



- Absolute integrals difficult by MC
  - *relates to free-energy evaluation*

Summary

- Markov process is a stochastic process with no memory
- Full specification of process is given by a matrix of transition probabilities Π
- A distribution of states are generated by repeatedly stepping from one state to another according to Π
- A desired limiting distribution can be used to construct transition probabilities using detailed balance
  - *Many different* Π *matrices can be constructed to satisfy detailed balance*
  - *Metropolis algorithm is one such choice, widely used in MC simulation*
- Markov Monte Carlo is good for evaluating averages, but not absolute integrals