Transition Path Theory

Frank Noé (noe@math.fu-berlin.de)

January 30, 2009

1 Basic Definitions

Let $X = \{1, ..., m\}$ be a discrete state space and let $X(t)$ be a Markov chain, where t may be either discrete or continuous.

1.1 Time-Discrete Markov Chain

with propagator / transition matrix $P \in \mathbb{R}^{m\times m}$:

$$p_{ij} \geq 0 \; \forall i, j$$
$$\sum_{j=1...m} p_{ij} = 0 \; \forall i$$

Unless the system is inherently time-discrete, the meaning of p_{ij} is that it represents the transition probability of state i to state j within time τ:

$$p_{ij} = \mathbb{P}[X(t+\tau) = j \mid X(t) = i]$$

P has a single eigenvalue of 1 and otherwise eigenvalues in $(-1, ..., 1)$. When $\mu(t) \in \mathbb{R}^m$ is a probability vector, P can be used to transport this probability vector:

$$\mu(t + n\tau) = \mu(t)P^n(\tau).$$

If P is irreducible it has a stationary distribution given by

$$\pi = \pi P$$

The backward propagator is defined as:

$$\tilde{p}_{ij} = \frac{\pi_j}{\pi_i} p_{ji}$$

If P is furthermore reversible it fulfills the detailed balance equation:

$$\tilde{p}_{ij} = p_{ij} \, \forall i, j$$
$$\pi_i p_{ij} = \pi_j p_{ji} \, \forall i, j.$$

1.2 Time-continuous Markov Chain

with generator / rate matrix $L \in \mathbb{R}^{m \times m}$:

$$l_{ij} \geq 0 \, \forall i \neq j$$
$$l_{ii} = -\sum_{i \neq j} l_{ij} \, \forall i$$

l_{ij} is the rate of change from i to j in infinitesimal time and has therefore no parameter τ. L has a single eigenvalue of 0 and otherwise negative eigenvalues. When $\mu(t) \in \mathbb{R}^m$ is a probability vector, L can be used to compute the change of this probability vector:

$$\frac{d\mu(t)}{dt} = \mu(t)L$$

If L is irreducible it has a stationary distribution given by

$$0 = \pi L$$

The backward generator is defined as:

$$\tilde{l}_{ij} = \frac{\pi_j}{\pi_i} l_{ji}$$

If P is furthermore reversible it fulfills the detailed balance equation:

$$\tilde{l}_{ij} = l_{ij} \, \forall i, j.$$
$$\pi_i l_{ij} = \pi_j l_{ji} \, \forall i, j.$$

P and L are related via

$$P(\tau) = \exp(\tau L).$$

Thus for every generator we can derive a unique transition matrix, but the inverse is not always possible. The generator can be obtained *via* following limit if it exists:

$$L = \lim_{\tau \to 0^+} \frac{P(\tau) - I}{\tau}$$

The generator corresponding to the time-discrete Markov process with $\tau = 1$ is

$$L_0 = P(\tau) - I$$

also known as pseudogenerator.

2 Ingredients

2.1 Hitting probabilites for Time-Discrete Markov Chains

Hitting time of A: $H^A : \Omega \to \{0, 1, 2, ...\} \cup \{\infty\}$:

$$H^A(\omega) = \inf\{n \geq 0 : X_n(\omega) \in A\}$$

Hitting probability: The probability that starting from i we ever hit A is:

$$h_i^A = \mathbb{P}_i(H^A < \infty).$$

The vector of hitting probabilities $h^A = (h_i^A : i \in I)$ is the minimal non-negative solution to the system of linear equations:

$$h_i^A = 1 \text{for } i \in A$$
$$h_i^A = \sum_{j \in I} p_{ij} h_j^A \text{for } i \notin A.$$

(Minimality means that if $x = (x_i : i \in S)$ is another solution with $x_i \geq 0$ for all i, then $x_i \geq h_i$ for all i.)

Proof:

First we show that h^A satisfies the above equation.

1) If $X_0 = i \in A$, then $H^A = 0$, so $h_i^A = 1$.

2) If $X_0 = i \notin A$, then $H^A \geq 1$, so by the Markov property:

$$\mathbb{P}_i(H^A < \infty \mid X_1 = j) = \mathbb{P}_j(H^A < \infty) = h_j^A$$

and

$$\begin{aligned} h_i^A &= \mathbb{P}_i(H^A < \infty) \\ &= \sum_{j \in I} \mathbb{P}_i(H^A < \infty, X_i = j) \\ &= \sum_{j \in I} \mathbb{P}_i(H^A < \infty \mid X_i = j)\mathbb{P}_i(X_1 = j) \\ &= \sum_{j \in I} h_j^A p_{ij}. \end{aligned}$$

Next, we show that h^A are the minimal solution

1) Suppose that $x = (x_i : i \in I)$ is any solution to the equation. then $h_i^A = x_i = 1$ for $i \in A$.

2) Suppose $i \notin A$, then:

$$x_i = \sum_{j \in I} p_{ij} x_j = \sum_{j \in A} p_{ij} + \sum_{j \notin A} p_{ij} x_j$$

Substitute for x_j to obtain:

$$\begin{aligned} x_i = \sum_{j \in I} p_{ij} x_j &= \sum_{j \in A} p_{ij} + \sum_{j \notin A} p_{ij} \left(\sum_{k \in A} p_{jk} + \sum_{k \notin A} p_{jk} x_k \right) \\ &= \mathbb{P}_i(X_1 \in A) + \mathbb{P}_i(X_1 \notin A, X_2 \in A) + \sum_{j \notin A} \sum_{k \notin A} p_{ij} p_{jk} x_k. \end{aligned}$$

By repeated substitution for x in the final term, we obtain after n steps:

$$x_i = \mathbb{P}_i(X_i \in A) + ... + \mathbb{P}_i(X_1 \notin A, ..., X_{n-1} \notin A, X_n \in A) + \sum_{j_1 \notin A} ... \sum_{j_n \notin A} p_{ij_1} p_{j_1 j_2} ... p_{j_{n-1} j_n} x_{j_n}.$$

Now if x is non-negative, so is the last term on the right, and the remaining terms sum to $\mathbb{P}_i(H^A < n)$. So $x_i \geq \mathbb{P}_i(H^A < n)$ for all n and then:

$$x_i \geq \lim_{n \to \infty} \mathbb{P}_i(H^A < n) = \mathbb{P}_i(H^A < \infty) = h_i.$$

2.2 Hitting probabilites for time-continuous Markov Chains

Using

$$\begin{aligned} h_i^A &= 1 \text{for } i \in A \\ h_i^A &= \sum_{j \in I} p_{ij} h_j^A \text{for } i \notin A. \end{aligned}$$

with $L = P - I$ yields

$$\begin{aligned} h_i^A &= 1 \text{for } i \in A \\ h_i^A = \sum_{j \in I, j \neq i} & l_{ij} h_j^A + (l_{ii} + 1) h_i^A \text{for } i \notin A. \end{aligned}$$

and thus

$$h_i^A = 1 \text{for } i \in A$$
$$\sum_{j \in I} l_{ij} h_j^A = 0 \text{for } i \notin A.$$

2.3 Time-discrete Committor Probabilities

The committor probability, q_i^+ pertaining to two sets A, B is the probability that starting in state i, we will go to B next rather than to A:

$$q_i^+ = \mathbb{P}_i(H^B < H^A).$$

In order to compute this, we define an A-absorbing process as

$$\hat{p}_{ij} = \begin{cases} p_{ij} & i \notin A, j \in S \\ 1 & i \in A, i = j \\ 0 & i \in A, i \notin j \end{cases}$$

and then compute the hitting probability to B. Since the process is absorbing in A only, the hitting probability to B will reflect the probability to go to B next rather than to A.

Using the hitting probability equations:

$$q_i^+ = 1 \text{for } i \in B$$
$$q_i^+ = \sum_{j \in I} p_{ij} q_j^+ \text{for } i \notin B.$$

with the absorbing process yields:

$$q_i^+ = 0 \text{for } i \in A$$
$$q_i^+ = 1 \text{for } i \in B$$
$$q_i^+ = \sum_{j \in I} p_{ij} q_j^+ \text{for } i \notin \{A, B\}.$$

The backward committor probability, q_i^- pertaining to two sets A, B is the probability that being in state i, we have been in A last rather than in B. In order to get the backward committor, we use the backwards propagator $\tilde{p}_{ij} = \frac{\pi_j}{\pi_i} p_{ji}$, consider a B-absorbing process for the reverse dynamics and compute the hitting probability for A:

$$q_i^- = 1 \text{for } i \in A$$
$$q_i^- = 0 \text{for } i \in B$$
$$q_i^- = \sum_{j \in I} \tilde{p}_{ij} q_j^- \text{for } i \notin \{A, B\}.$$

for reversibility / detailed balance, $p_{ij} = \frac{\pi_j}{\pi_i} p_{ji} = \tilde{p}_{ij}$ and thus:

$$q_i^- = 1 \text{for } i \in A$$
$$q_i^- = 0 \text{for } i \in B$$
$$q_i^- = \sum_{j \in I} p_{ij} q_j^- \text{for } i \notin \{A, B\}.$$

it can be easily checked that:

$$q^- = 1 - q^+$$

satisfies this equation as it transforms it to the forward committor equation.

2.4 Time-continuous Committor Probabilities

$$\hat{l}_{ij} = \begin{cases} l_{ij} & i \notin A \\ 0 & i \in A \end{cases}$$

substituted into the Dirichlet problem:

$$q_i^+ = 1 \text{for } i \in B$$
$$\sum_{j \in I} l_{ij} q_j^+ = 0 \text{for } i \notin B.$$

yields

$$q_i^+ = 0 \text{for } i \in A$$
$$q_i^+ = 1 \text{for } i \in B$$
$$\sum_{j \in I} l_{ij} q_j^+ = 0 \text{for } i \notin \{A, B\}.$$

In order to get the backward committor, we define the backwards propagator $\tilde{l}_{ij} = \frac{\pi_j}{\pi_i} l_{ji}$ and obtain:

$$q_i^- = 1 \text{for } i \in A$$
$$q_i^- = 0 \text{for } i \in B$$
$$\sum_{j \in I} \tilde{l}_{ij} q_j^- = 0 \text{for } i \notin \{A, B\}.$$

for reversibility / detailed balance, $l_{ij} = \frac{\pi_j}{\pi_i} l_{ji}$ and it can be easily checked that:

$$q^- = 1 - q^+$$

3 Reactive Flux:

3.1 Probability weight of reactive trajectories:

$$m_i^R = \pi_i q_i^- q_i^+$$

with $Z_{AB} = \sum_i m_i^R = \sum_i \pi_i q_i^- q_i^+ < 1$ it is clear that we need to normalize:

$$m_i^{AB} = Z_{AB}^{-1} \pi_i q_i^- q_i^+.$$

For detailed balance, we have:

$$m_i^{AB} = Z_{AB}^{-1} \pi_i (1 - q_i^+) q_i^+.$$

to obtain the probability distribution of reaction trajectories, i.e. the probability to be at state i and to be reactive.

3.2 Probability current of reactive trajectories:

$$f_{ij}^{AB} = \begin{cases} \pi_i q_i^- l_{ij} q_i^+ & i \neq j \\ 0 & i = j \end{cases}.$$

(for detailed balance, we have $q_i^- = 1 - q_i^+$). *The probability current is the number of jumps $i \rightarrow j$ which lie on reactive $A \rightarrow B$ trajectories.*

We have a number of nice properties:

1) **Flux conservation (Kirchhoff's 1st law)**

$$\sum_{j \in S} (f_{ij}^{AB} - f_{ji}^{AB}) = 0 \quad \forall i \notin \{A, B\}$$

Proof:

$$\sum_{j\in S}(f_{ij}^{AB}-f_{ji}^{AB}) = \pi_i q_i^- \sum_{j\neq i} l_{ij} q_j^+ - q_i^+ \sum_{j\neq i} \pi_j q_j^- l_{ji}$$
$$= \pi_i q_i^- \sum_{j\neq i} l_{ij} q_j^+ - \pi_i q_i^+ \sum_{j\neq i} q_j^- \tilde{l}_{ij}$$

Due to the committor equations, we have $\sum_{j\notin\{A,B\}} l_{ij} q_j^+ = 0$ and $\sum_{j\notin\{A,B\}} \tilde{l}_{ji} q_j^- = 0$ and thus:

$$\sum_{j\in S}(f_{ij}^{AB}-f_{ji}^{AB}) = -\pi_i q_i^- l_{ii} q_i^+ + \pi_i q_i^+ \tilde{l}_{ii} q_i^- = 0.$$

From $q_i^+ = 1 \forall i \in A$ and $q_i^- = 0 \forall i \in B$ we see that

$$f_{ij}^{AB} = 0 \forall j \in A$$
$$f_{ij}^{AB} = 0 \forall i \in B$$

thus flux is not conserved at A and B, but throughout the network such that:

$$\sum_{i\in A, j\notin A} f_{ij}^{AB} = \sum_{j\notin B, i\in B} f_{ji}^{AB}.$$

Remarks:

- It is worth noting that by setting q_i^+ as negative potential, f_{ij}^+ as current and $\pi_i l_{ij}$ as conductance provides an electric network theory with Ohm's law and Kirchhoff's laws being valid.

- All TPT is valid when substituting p_{ij} in l_{ij}.

3.3 Effective current

is defined as

$$f_{ij}^+ = \max\{f_{ij}^{AB} - f_{ji}^{AB}, 0\}$$

and gives the *net average number of reactive trajectories per time unit making a transition from i to j on their way from A to B.*

3.4 Total rate

The total number of reactive $A \to B$ trajectories per time unit is simply given by the effective current flowing out of A and into B:

$$K = \sum_{i\in A, j\notin A} f_{ij}^{AB} = \sum_{i\in A, j\notin A} \pi_i l_{ij} q_i^+ = \sum_{j\notin B, i\in B} f_{ji}^{AB} = \sum_{i\in B, j\notin B} \pi_j q_j^- l_{ji}$$

If the system is ergodic, every trajectory must go back from B to A in order to be able to transit to B again. Thus, K is also equal to the number of reactive $B \to A$ trajectories and equal to the number of $A \to B \to A$ cycles.

4 Pathways

- A reaction pathway $w = (i_0, i_1, ..., i_n)$ from A to B is a simple pathway, such that

$$i_0 \in A,\ i_n \in B,\ i_1...i_n \notin \{A, B\}.$$

- the capacity of a pathway w is the minimal effective current:

$$c(w) = \min_{(i,j)\in w} \{f_{ij}^+\}$$

- the bottleneck of a reaction pathway w is the edge with the minimal effective current:

$$(b_1, b_2) = \arg \min_{(i,j)\in w} \{f_{ij}^+\}$$

- The best pathway is one that maximizes the minimal current. This is only necessarily unique at the bottleneck. However, following algorithm is a rational way to find a unique best pathway in graph G:

BestPath(G,A,B)

1. Determine bottleneck (b_1, b_2) in G
2. Decompose G into L and R, which are the parts of G left of b_1 and right of b_2
3. $w_L = \begin{cases} b_1 & if\ b_1 \in A \\ BestPath(L, A, \{b_1\} & else \end{cases}$
4. $w_R = \begin{cases} b_2 & if\ b_2 \in B \\ BestPath(L, \{b_2\}, B & else \end{cases}$
5. return $(w_L$,$w_R)$

In order to decompose the network into individual pathways, let $w = BestPath(G, A, B)$ and subtract that pathway from the network:

$$(f_{ij}^+)' = f_{ij}^+ - c(w) if\ (i,j) \in w$$
$$(f_{ij}^+)' = f_{ij}^+ else.$$

It directly follows from the flux conservation laws that in the new pathways we still have flux conservation and

$$K' = K - c(w).$$

We will have $K' = 0$ when all $A \to B$ pathways have been subtracted and the decomposition is finished. This results in a set of $A \to B$ pathways whose statistical contribution to the $A \to B$ is given by the capacity of each, $c(w)$.

Literature

1. J. Norris: "Markov Chains". Cambride University Press.
   Parts available from: http://www.statslab.cam.ac.uk/~james/Markov/

2. Ph. Metzner, Ch. Schütte, and E. Vanden-Eijnden: "Transition Path Theory for Markov Jump Processes". Mult. Mod. Sim. (2007)