Long-Time Properties of CTMCs

Monday, November 30, 2015
1:57 PM

Homework 4 due Friday, December 11 at 5 PM.

The long-run properties of CTMCs follow some variation of the rules for determining these properties for DTMCs, so we will comment on the changes.

- main difference is that CTMCs are formulated in terms of rates of change, whereas DTMCs are formulated in terms of probabilities of change
- otherwise the derivations and ideas are the same

Classification of CTMCs (communication classes, positive recurrent, null recurrent, transient)

- Communication classes are determined by the same topological procedure as for DTMCs
- Characterization of recurrence vs. transience: these concepts don't involve timing, just the sequence of states visited, so for a CTMC, we can determine transience or recurrence of a class by calculating the property for the embedded DTMC
- Positive recurrence vs. null recurrence: It is still the case that positive recurrent classes are characterized definitively by having a well-defined stationary distribution, but the stationary distribution for a CTMC is not the same as the stationary distribution for the embedded DTMC because time spent in a state matters.

Stationary distribution for a CTMC

Just derive the equation by looking for a stationary solution to the forward Kolmogorov equation for the probability distribution ϕ of a state:

$$\frac{d\vec{\phi}^{\,T}(t)}{dt} = \vec{\phi}^{\,T}(t) \cdot A$$

Let π be the stationary solution:

$$\frac{d\vec{\pi}}{dt} = 0$$

$$\boxed{\begin{gathered}\vec{\pi}^T \cdot A = 0 \\ \pi_i \geq 0 \\ \sum_{i \in S} \pi_i = 1\end{gathered}}$$

Invariant measure for a CTMC is defined similarly, but without the normalization condition.

Looking for detailed balance solutions still makes sense for CTMCs, and the formula for this is:
$\pi_i A_{ij} = \pi_j A_{ji}$ for all $i, j \in S$.

So to classify a CTMC:

- use the embedded DTMC to determine communication classes, and which classes are transient
- compute the stationary distribution using the appropriate formula for CTMCs
- positive recurrent classes have stationary distributions and vice versa.

Once we have determined the nature of the communication classes, the long-time properties of the CTMC can be analyzed in the same way as for DTMCs, with the following modifications:

1. Once a Markov chain enters a positive recurrent class, then the long-time dynamics are governed by the unique stationary distribution for that class. That stationary distribution is a limit distribution (no need to check for periodicity): $\lim_{t\to\infty} P(X(t) = j) = \pi_j$

   LLN for CTMCs: $\lim_{T\to\infty} \frac{1}{T}\int_0^T f(X(t))dt = \sum_{i\in S} \pi_i f(i)$

2. Null recurrent classes have roughly the same behavior for CTMCs and DTMCs; the probability distribution will spread out over all states with probability to be in any given state going to 0: $\lim_{t\to\infty} P(X(t) = j) = 0$. Not much more can be said in general; need special purpose techniques to analyze. Establishing a LLN in terms of the invariant measure could be possible, but much more complicated to formulate and prove.
3. For Markov chains starting in a transient class; if this class is closed, then again not much more can be said in general other than $\lim_{t\to\infty} P(X(t) = j) = 0$. But if the class is not closed (one-way links to other communication classes), then one can calculate absorption probabilities and accumulated cost/reward until absorption in one of the other (recurrent) communication classes through modified formulas from the DTMC
   a. Absorption probabilities: This is insensitive to timing, so computing absorption probabilities for the embedded DTMC will give the absorption probabilities for the CTMC. Alternatively, one can just develop a formula for absorption probabilities directly w.r.t. the transition rate matrix A.

We will derive this formula by taking a continuum limit of the DTMC formula, but it's more conventional to derive it directly in parallel to how we derived the DTMC formula, but in continuous time

- This direct derivation can be found in Goutsias and Jenkinson, "Markovian dynamics on complex reaction networks," Physics Reports 529 (2013), 199-264

But we will instead just proceed as follows: Consider the regularly sampled DTMC $X_n = X(n\Delta t)$; it has probability transition matrix. $P = \exp(A\,\Delta t) = I + A\,\Delta t + o(\Delta t)$

Recall $U_{ij} = P(X_\tau = j | X_0 = i)$ where τ is the random epoch of absorption.

The absorption probability for the regularly sampled DTMC would satisfy:

$$U_{ij} = P_{ij} + \sum_{k \in T} P_{ik} U_{kj}$$

$$\cancel{U_{ij}} = \left(\delta_{ij} + A_{ij}\,\Delta t + o(\Delta t)\right) + \sum \left(\delta_{ik} + A_{ik}\,\Delta t + o(\Delta t)\right) U_{kj}$$

$k \in T$

$$= \left(\underbrace{\delta_{ij}}_{0} + \cancel{U_{ij}}\right) + \Delta t\left(A_{ij} + \sum_{k \in T} A_{ik} U_{kj}\right) + o(\Delta t)$$

because $i \in T, j \in T^c$

$$0 = \Delta t\left(A_{ij} + \sum_{k \in T} A_{ik} U_{kj}\right) + o(\Delta t)$$

$\Delta t \downarrow 0$:

$$\boxed{A_{ij} + \sum_{k \in T} A_{ik} U_{kj} = 0 \quad \text{for } i \in T,\ j \in T^c}$$

$\Delta t \downarrow 0$:

$$\boxed{A_{ij} + \sum_{k \in T} A_{ik} U_{kj} = 0 \quad \text{for } i \in T, \; j \in T^c}$$

This is again a matrix equation or linear system, can be handled by similar techniques to the related equation for DTMCs.

3b. Accumulated cost/reward (additive functional)

$$w_i = \mathbb{E}\left(\int_0^\tau f(X(t))dt \mid X(0) = i\right)$$

where now f is a rate of cost/reward.

This relates to timing so one cannot just calculate it based on the embedded DTMC. But one can derive equations involving the transition rate matrix either from scratch in parallel with how we derived the DTMC formulas (Lawler Sec. 3.2, Karlin and Taylor Ch. 4), or one can also take a continuum limit of the regularly sampled DTMC formulas as we did for absorption probabilities.

$-A\,w = f$
$w_i = f_i = 0$ for $i \in T^c$.

where:

$$\vec{w} = \begin{pmatrix} w_1 \\ w_2 \\ \vdots \end{pmatrix} \qquad \vec{f} = \begin{pmatrix} f(1) \\ f(2) \\ \vdots \end{pmatrix}$$

Some extensions of these basic ideas:

1. One can not only calculate the probability to be absorbed in a given state/recurrent class, or the mean time until the absorption. One can calculate the full probability distribution for the time until absorption at the cost of working with a time-dependent linear equation (rather than the static linear equations above).

   If we continue to refer to τ as the random time until absorption by a recurrent class, then:

$$P(\tau > t) = P(\underline{X}(t) \in T)$$

equivalence of event formulated in terms of time vs. state

$$= \sum_{j \in T} P(\underline{X}(t) = j)$$

$$= \sum_{j \in T} \phi_j(t)$$

which can be computed by solving the FKE.

2. The formulas for CTMCs have direct analogues for SDEs, where the transition rate matrix A is replaced by a second order partial differential operator. In particular, the formula for the expected accumulated cost/reward has the form of a Poisson equation.

3. For reaction networks, we know that the transition rate matrix A is difficult to work with as a whole. The equations we have written above can be formulated in terms of the sparse matrix representation used for reaction networks.

   So for example the FKE

$$\frac{\partial \vec{\phi}}{\partial t} = \vec{\phi}(t) \cdot A$$

   would be written instead as follows:

$$\vec{\phi}_{\vec{x}}(t) = P(\vec{\underline{X}}(t) = \vec{x})$$

$$\frac{\partial \vec{\phi}_{\vec{x}}(t)}{\partial t} = \sum_{m \in \mathcal{M}} \left[\alpha_m(\vec{x} - \vec{s}_m) \vec{\phi}_{\vec{x} - \vec{s}_m}(t) - \alpha_m(\vec{x}) \vec{\phi}_{\vec{x}}(t) \right]$$

$\mathcal{M}$ is a list of reaction channels

$\alpha_m(\vec{x})$ is rate of reaction m

$\vec{s}_m$: relative state change induced by reaction m

One can do this similarly for the other equations. Note that the equations that are related to the backward Kolmogorov equation will involve the infinitesmal generator matrix in an adjoint fashion. The above equation will then refer to destination states $(x + s_m)$ rather than source states $(x - s_m)$.

Martingales

References: Karlin and Taylor Ch. 6
Lawler Sec. 5.1-5.3

Martingales are not really a stochastic modeling framework but rather a mathematical object that can simplify calculations involving stochastic models.

We need to refer to the concept of filtration which is just a mathematical encoding of information accumulation.

Filtrations are technically increasing sequences of σ −algebras on sample space:

$$\mathcal{A}_0 \subseteq \mathcal{A}_1 \subseteq \mathcal{A}_2 \subseteq \cdots \subseteq \mathcal{A}_n \subseteq \mathcal{A}_{n+1}$$

The intuition (which will suffice for practical purposes) is that $\mathcal{A}_n$ represents the information available at epoch n.

$B \in \mathcal{A}_n$ if the event B is fully determined by information available up to epoch n.

For any given stochastic model, one can define the information in multiple ways (leading to multiple filtrations), so one needs to be a little more specific. One good way to specify the filtration is to pick a set of stochastic processes whose values generate the filtration.

$\{ Y_n \}^\infty$

$\{\mathcal{A}_n\}_{n=0}^{\infty}$ is said to be generated by the stochastic processes $\{X_n\}$ and $\{Y_n\}$ provided:

$B \in \mathcal{A}_n$ if and only if

B is completely determined by $\{X_0, X_1, X_2, \ldots X_n\} \cup \{Y_0, Y_1, \ldots, Y_n\}$

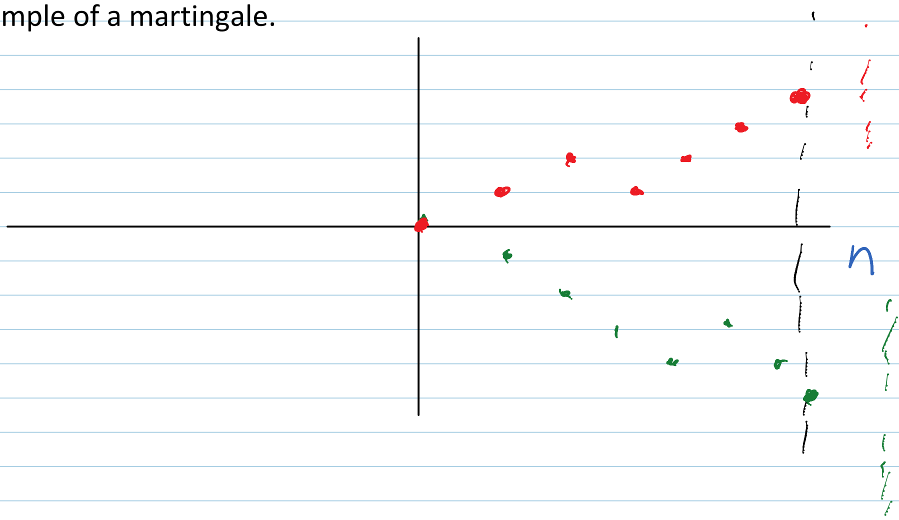
A discrete-time stochastic process $\{Y_n\}_{n=0}^{\infty}$ is said to be a martingale with respect to the filtration $\{\mathcal{A}_n\}_{n=0}^{\infty}$

provided the following two statements hold:

- $\mathbb{E}(Y_{n+1} | \mathcal{A}_n) = Y_n$ for all $n = 0,1,2,\ldots$
- $\mathbb{E}|Y_n| < \infty$ for all $n = 0,1,2,\ldots$
- Also Y_n needs to be adapted to the filtration $\mathcal{A}_n$

The last point is that the value of Y_n is part of the information "known" by $\mathcal{A}_n$.

The intuitive portion of the definition of a martingale (first point) is that the average future value is the same as the last observed value of the stochastic process. Note the classical random walk is an example of a martingale.



Markov chains and martingales are distinct concepts, possible to be one without the other.