Finite-Time Statistics for CTMC

Thursday, November 19, 2015
2:09 PM

Homework 3 due Monday, November 23 at 2 PM.

Continuing with the Chapman-Kolmogorov equation approach on the time interval $[0, t + \Delta t]$ decomposed into $[0, \Delta t) \cup [\Delta t, t + \Delta t]$.

$$P_{ij}(t+\Delta t) = \sum_{k \in S} P_{ik}(\Delta t)\, P_{kj}(t)$$

$$P_{ik}(\Delta t) = \delta_{ik} + A_{ik}\,\Delta t + o(\Delta t)$$

$\delta_{ik} = 1$ if $i = k$, 0 if $i \neq k$

$$P_{ij}(t+\Delta t) = \sum_{k \in S} \left(\delta_{ik} + A_{ik}\,\Delta t + o(\Delta t)\right) P_{kj}(t)$$

$$= P_{ij}(t) + \sum_{k \in S} A_{ik}\,\Delta t\, P_{kj}(t) + o(\Delta t)$$

$$\lim_{\Delta t \downarrow 0} \frac{P_{ij}(t+\Delta t) - P_{ij}(t)}{\Delta t} = \lim_{\Delta t \downarrow 0} \left(\sum_{k \in S} A_{ik}\, P_{kj}(t) + \frac{o(\Delta t)}{\Delta t}\right)$$

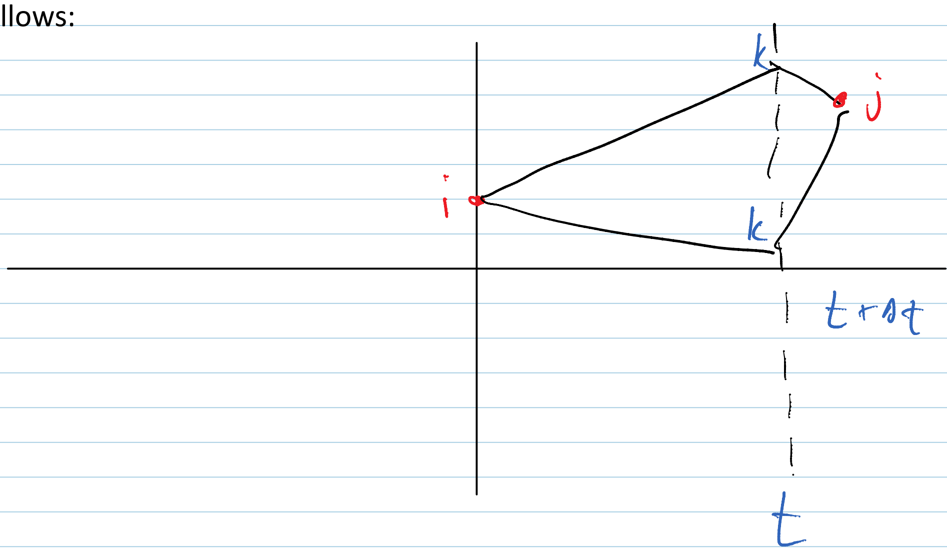
$$\boxed{\frac{dP_{ij}(t)}{dt} = \sum_{k \in S} A_{ik}\, P_{kj}(t) \qquad IC: P_{ij}(t=0) = \delta_{ij}}$$

Or in matrix form:

$$\boxed{\frac{dP(t)}{dt} = A\, P(t), \quad IC: P(t=0) = I}$$

This equation is known as the Kolmogorov backward equation.

One could also have proceeded with a similar argument but decomposing the time interval as follows:



$$P_{ij}(t+\Delta t) = \sum_{k \in S} P_{ik}(t)\, P_{kj}(\Delta t)$$

Copying the same argument to this new choice of intermediate time yields instead:

$$\boxed{\begin{aligned} \frac{dP_{ij}(t)}{dt} &= \sum_{k \in S} P_{ik}(t)\, A_{kj}, \quad IC: P_{ij}(t=0) = \delta_{ij} \\ \frac{dP(t)}{dt} &= P(t) A, \quad IC: P(0) = I \end{aligned}}$$

This equation is known as the Kolmogorov forward equation.

Here the Kolmogorov forward and Kolmogorov backward equation are for the same quantity, so they must have the same solution, and there's no need to have two equations for the same thing. But we'll see later that some other statistics we may be interested in are more naturally posed in terms of the backward or the forward equation.

Both the Kolmogorov forward and backward equation have the following formal solution:

$P(t) = \exp(At)$.

What does it mean to exponentiate a matrix, particularly if it has infinitely many states?

Matrix exponentials are defined (for square matrices):

$\exp(M) = \sum_{j=0}^{\infty} \frac{M^j}{j!}$ wherever this converges. One should check this, but if the matrix M has bounded eigenvalues (bounded in spectral norm) it will converge. This will always be true if M is finite, and otherwise the concept of matrix exponential is more of an abstraction than a useful representation.

How should one compute $\exp(At)$ in practice for finite matrices A?

One way would be to find the Jordan form (or try to diagonalize) A:

$A = C\,D\,C^{-1}$

$$A^j = (CDC^{-1})^j = C D^j C^{-1}$$

$$\exp(At) = C \exp(Dt)\, C^{-1}$$

If D is diagonal, then:

$$\exp(Dt) = \begin{pmatrix} e^{\lambda_1 t} & & & 0 \\ & e^{\lambda_2 t} & & \\ & & \ddots & \\ 0 & & & e^{\lambda_m t} \end{pmatrix}$$

if $D = \begin{pmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_m \end{pmatrix}$

Also:

"Nineteen Dubious Ways to Compute the Exponential of a Matrix, Twenty-Five Years

Later," Clive Moler, Charles van Loan, SIAM Review 45, p. 3 (2003).

But these matrix exponential approaches are really only useful for CTMCs with reasonably small number of states. For larger CTMCs, especially if A has some good structure, one is better off to try to solve the differential equations analytically or numerically.

One technical remark: From a mathematical standpoint (for both CTMCs and SDEs) the Kolmogorov backward equation is easier to establish from a technical standpoint (fewer conditions on the CTMC/SDE is necessary for it to make sense) than for the Kolmogorov forward equation. It's because the row sums of A are under better control than the column sums of A.

Kolmogorov forward equation for probability distribution of future states

Suppose we want to calculate the evolution of the probability distribution for the state, as function of time: $\phi_j(t) = P(X(t) = j)$.

If we know the probability transition function, then of course:

$$\phi_j(t) = \sum_{i \in S} \phi_i(0) P_{ij}(t)$$

law of total prob

$$\vec{\phi}(t) = \vec{\phi}(0) \cdot P(t)$$

But this approach requires knowing $P(t)$ which involves solving a matrix equation involving $|S| \times |S|$ variables.

Let's formulate an alternative approach that does not require us to solve for $P(t)$. Let's take the relationship we just developed, and differentiate it w.r.t. time:

$$\frac{d\vec{\phi}}{dt} = \vec{\phi}(0) \cdot \frac{dP}{dt}$$

KFE

$$\frac{d\vec{\phi}}{dt} = \vec{\phi}(0) \cdot P(t) A$$

$$\boxed{\frac{d\vec{\phi}}{dt} = \vec{\phi}(t) \cdot A}$$

$$\frac{d}{dt}$$

$$\vec{\phi}(t=0) = \vec{\phi}^{(0)}$$

Kolmogorov forward equation for the probability distribution of the state. Note this is a system of $|S|$ linear differential equations. Even if $|S| = \infty$ this can be written as an appropriate sum.

The formal solution to this equation is:

$$\vec{\phi}(t) = \vec{\phi}(0) \cdot e^{At}$$

$$\left(\text{CTMC version of } \vec{\phi}^{(n)} = \vec{\phi}^{(0)} P^n \right)$$

$$P = e^{A\Delta t} \quad \text{for regularly sampled DTMC}$$

Also, the DTMC analog for the Kolmogorov forward equation would be:

$$\phi^{(n+1)} = \phi^{(n)} P$$

Kolmogorov backward equation describes the evolution of future expected values

$$u_i(t) = \mathbb{E}[f(X(t))|X(0) = i]$$

Here f is a deterministic function on the state space of the CTMC, which we can think of as a "value" function.

- finance
- ecology (fitness)

We can represent this expected future value using the definition of conditional expectation:

$$u_i(t) = \sum_{j \in S} f(j)\, P(X(t) = j \mid X(0) = i)$$

$$= \sum_{j \in S} f(j)\, P_{ij}(t)$$

$$\vec{u} = \{u_i\}_{i \in S}$$

$$\vec{f} = \{f(i)\}_{i \in S}$$

$$\vec{u}(t) = P(t) \cdot \vec{f}$$

$$\frac{d\vec{u}}{dt} = \frac{dP}{dt} \cdot \vec{f}$$

$\downarrow$ KBE

$$\frac{d\vec{u}}{dt} = A\, P(t) \cdot \vec{f}$$

$$\boxed{\begin{aligned} \frac{d\vec{u}}{dt} &= A \vec{u} \\ \vec{u}(t=0) &= \vec{f} \end{aligned}}$$

Kolmogorov backward equation for expected future values. This is related to the computational implementation of backward stochastic programming.

We could have also solved this problem for DTMCs

$$u_i^{(n)} = \mathbb{E}\left[f(X_n) \mid X_0 = i \right]$$

$$u_i^{(n)} = \sum_{j \in S} f(j)\, P(X_n = j \mid X_0 = i)$$

$$u_i^{(n)} = \sum_{j \in S} f(j) \left(P^n\right)_{ij}$$

$$\vec{u}^{(n)} = P^n \vec{u}^{(0)}$$

$$\vec{u}^{(n)} = P \vec{u}^{(n-1)}$$

Note also that in most applications (finance and ecology) the expected future value involves a fixed end time and a variable initial time so we would equivalently write:

$$u_i(t) = \mathbb{E}[f(X(T))|X(T-t) = i]$$

Example: Poisson counting process

State space $S = \mathbb{Z}_{\geq 0}$ with transition rate matrix:

$A_{i,i+1} = \lambda, A_{i,i} = -\lambda$; otherwise $A_{ij} = 0$.

$$A = \begin{array}{c|c} & \begin{array}{ccccc} 0 & 1 & 2 & & \end{array} \\ \hline \begin{array}{c} 0 \\ 1 \\ 2 \end{array} & \begin{pmatrix} -\lambda & \lambda & 0 & 0 & \cdots & 0 \\ 0 & -\lambda & \lambda & 0 & 0 & \\ 0 & 0 & -\lambda & \lambda & \ddots & \\ & & & \ddots & \ddots & \\ & 0 & & & \ddots & \ddots \end{pmatrix} \end{array}$$

Let's solve for the probability distribution of the state at future times, starting from initial condition $X_0 = 0$.

$\phi_i^{(0)} = \delta_{i0}$

$$\vec{\phi}^{(0)} = \begin{pmatrix} 1 \\ 0 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

Kolmogorov forward equation for the probability distribution of state:

$$\frac{d\vec{\phi}^T}{dt} = \vec{\phi}^T \cdot A$$

$$\frac{d\phi_0}{dt} = -\lambda \phi_0$$

$$\frac{d\phi_j}{dt} = -\lambda \phi_j + \lambda \phi_{j-1} \quad \text{for } j \geq 1$$

This infinite systems of constant coefficient linear differential equations can be solved by Laplace transform.