Lecture 4:
Stationary distributions for CTMCs

Susana Gomes

October 15^{th}, 2021

Plan for today

1. Chapman-Kolmogorov equations, generators and the Master equation.
2. Stationary and irreducible distributions for CTMCs
3. Discussion about Assignment 1

Chapman-Kolmogorov equations for CTMCs

We can still define a transition function as before and the Chapman-Kolmogorov equations are valid.

Let $(X_t : t \geq 0)$ by a homogeneous CTMC with state space S. Then for all $t \geq 0$ the **transition function** is given by

$$p_t(x, y) := \mathbb{P}[X_t = y | X_0 = x] = \mathbb{P}[X_{t+u} = y | X_u = x] \quad \text{for all } u \geq 0.$$

Proposition: Chapman-Kolmogorov equations

The transition function is well defined and fulfills the **Chapman Kolmogorov equations**

$$p_{t+u}(x, y) = \sum_{z \in S} p_t(x, z)\, p_u(z, y) \quad \text{for all } t, u \geq 0,\ x, y \in S.$$

Generator of a CTMC

As before, we can write this in matrix notation:

We define $P_t = \big(p_t(x, y) : x, y \in S\big)$ and we can write

$$P_{t+u} = P_t \, P_u \quad \text{with} \quad P_0 = \mathbb{I}.$$

In particular,

$$\frac{P_{t+\Delta t} - P_t}{\Delta t} = P_t \frac{P_{\Delta t} - \mathbb{I}}{\Delta t} = \frac{P_{\Delta t} - \mathbb{I}}{\Delta t} P_t .$$

We now take $\Delta t \searrow 0$ and get the so-called **forward and backward equations**

$$\frac{d}{dt} P_t = P_t G = G P_t \ , \quad \text{where} \quad \boxed{G = \frac{dP_t}{dt}\Big|_{t=0}}$$

is called the **generator** of the process (sometimes also Q-matrix).

Forward and backward equations for CTMCs

The solution for these equations is given by the matrix exponential

$$P_t = \exp(tG) = \sum_{k=0}^{\infty} \frac{t^k}{k!} G^k = \mathbb{I} + tG + \frac{t^2}{2} G^2 + \ldots$$

And from this we can obtain the **distribution** π_t **at time** $t > 0$:

$$\langle \pi_t | = \langle \pi_0 | \exp(tG) \quad \text{which solves} \quad \frac{d}{dt} \langle \pi_t | = \langle \pi_t | G.$$

Note that, just like before, if S is finite, we can compute the eigenvalues of G, $\lambda_1, \ldots, \lambda_L \in \mathbb{C}$. Then, P_t **has eigenvalues** $\exp(t\lambda_i)$ with the same eigenvectors $\langle v_i |$, $| u_i \rangle$.

If the λ_i are distinct, we can still expand the initial condition in the eigenvector basis

$$\langle \pi_0 | = \alpha_1 \langle v_1 | + \ldots + \alpha_L \langle v_L |,$$

where $\alpha_i = \langle \pi_0 | u_i \rangle$. This leads to

$$\langle \pi_t | = \alpha_1 \langle v_1 | e^{\lambda_1 t} + \ldots + \alpha_L \langle v_L | e^{\lambda_L t} \, .$$

Transition rates

Using the expression for P_t we have, for $G = \big(g(x,y) : x, y \in S\big)$,

$$p_{\Delta t}(x,y) = g(x,y)\Delta t + o(\Delta t) \quad \text{for all } x \neq y \in S,$$

so the $g(x,y) \geq 0$ can be interpreted as **transition rates**.

We also have

$$p_{\Delta t}(x,x) = 1 + g(x,x)\Delta t + o(\Delta t) \quad \text{for all } x \in S.$$

P_t is stochastic

$$\sum_y p_{\Delta t}(x,y) = p_{\Delta t}(x,x) + \sum_{y \neq x} p_{\Delta t}(x,y) =$$

Since $\sum_y p_{\Delta t}(x,y) = 1$, this implies that

$$\boxed{g(x,x)} = \boxed{-\sum_{y \neq x} g(x,y) \leq 0} \quad \text{for all } x \in S.$$

$$= \cancel{1} + g(x,x)\Delta t + \sum_{y \neq x} g(x,y)\Delta t = \cancel{1} \Rightarrow g(x,x)\Delta t + \sum g(x,y)\Delta t = 0$$

The Master equation

Using the results from the previous slide, we can rewrite the equation for the distribution at time t:

$$\langle \pi_t | = \langle \pi_0 | \exp(tG) \quad \text{which solves} \quad \frac{d}{dt} \langle \pi_t | = \langle \pi_t | G,$$

as the **Master equation**

$$\frac{d}{dt} \pi_t(x) = \underbrace{\sum_{y \neq x} \pi_t(y) g(y, x)}_{\text{gain term}} - \underbrace{\sum_{y \neq x} \pi_t(x) g(x, y)}_{\text{loss term}} \quad \text{for all } x \in S\ .$$

Note that: the Gershgorin theorem now implies that either $\lambda_i = 0$ or $\mathrm{Re}(\lambda_i) < 0$ for the eigenvalues of G, so there are **no persistent oscillations for CTMCs**.

Stationary and reversible distributions

These definitions are similar to the discrete-time case...

Let $(X_t : t \geq 0)$ be a homogeneous CTMC with state space S.

The distribution $\pi(x),\ x \in S$ is called **stationary** if $\langle \pi | G = \langle 0|$, or for all $y \in S$

$$\sum_{x \in S} \pi(x) g(x,y) = \sum_{x \neq y} \big(\pi(x) g(x,y) - \pi(y) g(y,x)\big) = 0.$$

π is called **reversible** if it fulfills the **detailed balance conditions**

$$\pi(x) g(x,y) = \pi(y) g(y,x) \quad \text{for all } x, y \in S.$$

- As before, **reversibility implies stationarity**.
- Stationary distributions are left **eigenvectors** of G with **eigenvalue** 0.
- $\langle \pi | G = \langle 0|$ implies $\langle \pi | P_t = \langle \pi | \big(\mathbb{I} + \sum_{k \geq 1} t^k G^k / k!\big) = \langle \pi|$ for all $t \geq 0$

Stationary distributions (existence)

Proposition (existence):

A CTMC with **finite** state space S has **at least one** stationary distribution.

Proof. The proof is similar to the discrete-time case. Since G has row sum 0, we have $G|\mathbf{1}\rangle = |\mathbf{0}\rangle$.
So 0 is an eigenvalue of G, and its corresponding left eigenvector(s) can be shown to have non-negative entries and thus can be normalized to be stationary distributions $\langle\pi|$.

Remark:

Note that if S is countably infinite, stationary distributions may not exist.
This is the case, for example, for the SRW on $\mathbb{Z}$ or the Poisson process on $\mathbb{N}$ (which we will see later).

Stationary distributions (uniqueness)

A CTMC (or DTMC) is called **irreducible**, if for all $x, y \in S$

$$p_t(x, y) > 0 \text{ for some } t > 0.$$

Note that for continuous time irreducibility implies $p_t(x, y) > 0$ for **all** $t > 0$.

Proposition (Uniqueness):

An **irreducible** Markov chain has **at most one** stationary distribution.

Proof. Follows from the **Perron Frobenius theorem:**
Let P be a stochastic matrix ($P = P_t$ for any $t \geq 0$ for CTMCs). Then:

- We know that $\lambda_1 = 1$ is an eigenvalue of P.
- From PF, we can conclude that this eigenvalue is singular if and only if the chain is irreducible.
- As before, we can show that its corresponding left and right eigenvectors have non-negative entries and so we obtained the distribution.

More on stationary distributions

→ of G

The Perron-Frobenius theorem also implies the following:

- If the chain is continuous-time, all remaining eigenvalues $\lambda_i \in \mathbb{C}$, $i \neq 1$ satisfy $\mathrm{Re}(\lambda_i) < 0$.
- If the chain is discrete-time aperiodic (no persistent oscillations), all remaining eigenvalues $\lambda_i \in \mathbb{C}$, $i \neq 1$ satisfy $|\lambda_i| < 1$. → of P
- The second part of the Perron Frobenius theorem also implies convergence of the transition functions to the stationary distribution, since

$$p_t(x, y) = \sum_{i=1}^{|S|} \langle \delta_x | u_i \rangle \langle v_i | \, e^{\lambda_i t} \to \langle v_1 | = \langle \pi | \quad \text{as } t \to \infty.$$

This is usually called ergodicity (and we will see more of it next week).

$e^{-t} \to 0$ as $t \to \infty$