3 Absorbing Markov Chains

3.1 Discrete-time absorbing Markov chains

Let $\{X(n), n = 0, 1, \ldots\}$ be a homogeneous discrete-time Markov chain (DTMC), taking values in the state-space $\mathcal{E} := \{1, 2, \ldots, N, 1^*, 2^*, \ldots, M^*\}$. Let $\mathbf{P} = [p_{i,j}]_{i,j\in\mathcal{E}}$ be its transition probability matrix, namely, $p_{i,j} = P(X(n+1) = j \mid X(n) = i)$ for all $i, j \in \mathcal{E}$.

States $1, 2, \ldots, N$ are all transient and states $1^*, 2^*, \ldots, M^*$ are all absorbent. This means that starting at time $n = 0$ from any transient state the Markov chain will eventually reach an absorbing state and will remain forever in this absorbing state. Starting from an absorbing state the Markov chain will remain in this absorbing state.

With this definition, the transition matrix $\mathbf{P}$ takes the form

$$\mathbf{P} = \begin{bmatrix} \mathbf{A} & \mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}$$

where $\mathbf{A} = [a_{i,j}]_{1\le i,j\le N}$ and $\mathbf{R} = [r_{i,j}]_{1\le i\le N, 1^*\le j\le M^*}$ are N-by-N and N-by-M matrices, respectively, and $\mathbf{I}$ is the M-by-M identity matrix. $a_{i,j}$ is the probability of going from transient state i to transient state j in one step, and $r_{i,j}$ is the probability of going from transient state i to absorbing state j in one step.

The n-step transition probabilities are obtained by raising the matrix $\mathbf{P}$ to the power n. Namely,

$$\mathbf{P}^n = \begin{bmatrix} \mathbf{A}^n & \sum_{k=0}^{n-1} \mathbf{A}^k \mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} .$$

We know that the (i, j)-entry of the matrix $\mathbf{P}^n$ gives $P(X(n) = j \mid X(0) = i)$, the probability of going from state i to state j in n steps (see Section 2.1, Proposition 1). Therefore, the (i, j)-entry of the matrix $\mathbf{A}^n$ – denoted by $a_{i,j}^{(n)}$ – gives the probability of going from transient state i to transient state j in n steps.

Limit of $\mathbf{A}^n$

Starting from any transient state, an absorbing Markov chain will eventually reach an absorbing state. We can write that for any pair of transient states i and j

$$a_{i,j}^{(n)} = P(X(n) = j \mid X(0) = i) \to 0 \quad \text{as } n \to \infty$$

since eventually each transient state will reach an absorbing state. This implies that all entries of the matrix $\mathbf{A}^n$ will converge to 0 as $n \to \infty$, that is,

$$\mathbf{A}^n \to \mathbf{0} \quad \text{as } n \to \infty . \tag{41}$$

Mean number of visits

Let $n_{i,j}$ be the *expected* number of visits to transient state j starting from transient state i, namely,

$$n_{i,j} = E\left[\sum_{n\geq 0} \mathbb{1}_{X(n)=j} \,|\, X(0) = i\right], \quad i,j \in \{1,2,\ldots,N\}.$$

As usual $\mathbb{1}_{X(n)=j}$ is equal to 1 if $X(n) = j$ and is equal to 0 otherwise. Note that if (for instance) $X(n) = j$, $X(n+1) = j$ and $X(n+2) = j$ then the number of visits to state j in $[n, n+2]$ is three (and not one).

Define the N-by-N matrix $\mathbf{N} = [n_{i,j}]_{1\leq i,j\leq N}$. $\mathbf{N}$ is called *fundamental matrix.*

Below is the main result:

Proposition 8 (Mean number of visits)**.** *The matrix* $\mathbf{N}$ *is given by*

$$\mathbf{N} = (\mathbf{I} - \mathbf{A})^{-1}\ . \tag{42}$$

■

Proof. Let i and j be arbitrary transient states. Define the binary random variable (rv) $X_j^{(n)} = \mathbb{1}_{X(n)=j}$, which is equal to 1 if $X(n) = j$ and to 0 otherwise. We have

$$\begin{aligned}
E[X_j^{(0)} + X_j^{(1)} + \cdots + X_j^{(n)} \,|\, X(0) = i] &= \sum_{k=0}^{n} E[X_j^{(k)} \,|\, X(0) = i] \qquad (43)\\
&= \sum_{k=0}^{n} E[\mathbb{1}_{X(k)=j} \,|\, X(0) = i]\\
&= \sum_{k=0}^{n} P(X(k) = j \,|\, X(0) = i)\\
&= \sum_{k=0}^{n} a_{i,j}^{(k)} \qquad (44)
\end{aligned}$$

where we recall that $a_{i,j}^{(k)}$ is the (i,j)-entry of the matrix $\mathbf{A}^k$.

Note that the left-hand side of (43) is nothing but the expected number of visits to state j in $[0, n]$ starting from state i. On the other hand,

$$\begin{aligned}
\lim_{n\to\infty} E[X_j^{(0)} + X_j^{(1)} + \cdots + X_j^{(n)} \,|\, X(0) = i] &= E[\lim_{n\to\infty} (X_j^{(0)} + X_j^{(1)} + \cdots + X_j^{(n)}) \,|\, X(0) = i]\\
&= n_{i,j}, \qquad (45)
\end{aligned}$$

where the first equality holds because the rvs $\{X_j^{(k)}\}_k$ are all nonnegative (Beppo-Levi theorem), and the second equality is nothing but the definition of $n_{i,j}$.

Combining (44) and (45) yields

$$n_{i,j} = \sum_{k\geq 0} a_{i,j}^{(k)} \tag{46}$$

or, in matrix notation,

$$\mathbf{N} = \sum_{k \geq 0} \mathbf{A}^k. \tag{47}$$

It remains to evaluate the matrix $\sum_{k \geq 0} \mathbf{A}^k$. This can be done as follows. We have

$$(\mathbf{I} - \mathbf{A}) \sum_{k=0}^{n} \mathbf{A}^k = \mathbf{I} - \mathbf{A}^{n+1}.$$

Letting $n \to \infty$ in the above identity, then using (47) and (41) gives

$$(\mathbf{I} - \mathbf{A})\mathbf{N} = \mathbf{I}$$

which in turn implies that $\mathbf{N}$ is the inverse of the matrix $\mathbf{I} - \mathbf{A}$, i.e. $\mathbf{N} = (\mathbf{I} - \mathbf{A})^{-1}$, which concludes the proof. ★

Absorption probabilities

Let $\mathbf{B}$ be the N-by-M^* matrix whose (i, j)-entry $b_{i,j}$ gives the probability that the chain is absorbed in state $j \in \{1^*, 2^*, \ldots, M^*\}$ given that it was initially in transient state $i \in \{1, 2, \ldots, N\}$ (i.e. $X(0) = i$). By definition, $\mathbf{B}$ is a stochastic matrix.

The following holds:

Proposition 9 (Absorption probabilities)**.**

$$\mathbf{B} = \mathbf{N}\mathbf{R} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{R}\ . \tag{48}$$

■

Proof. The time before absorption can be any value in $\mathbb{N}$. When computing $b_{i,j}$, all these cases must be taken into account. We have for $i \in \{1, 2, \ldots, N\}$, $j \in \{1^*, 2^*, \ldots, M^*\}$,

$$\begin{aligned}
b_{i,j} &= \sum_{n \geq 0} p_{i,j}^{(n+1)} = \sum_{n \geq 0} P(X(n+1) = j \,|\, X(0) = i) && \text{(use now law of total probabilities)} \\
&= \sum_{n \geq 0} \sum_{k=1}^{N} P(X(n+1) = j \,|\, X(n) = k, X(0) = i) \times P(X(n) = k \,|\, X(0) = i) \\
&= \sum_{n \geq 0} \sum_{k=1}^{N} P(X(n+1) = j \,|\, X(n) = k)\, a_{i,k}^{(n)} && \text{(Markov property)} \\
&= \sum_{n \geq 0} \sum_{k=1}^{N} r_{k,j} a_{i,k}^{(n)} = \sum_{k=1}^{N} \left(\sum_{n \geq 0} a_{i,k}^{(n)} \right) r_{k,j} \\
&= \sum_{k=1}^{N} n_{i,k} r_{k,j} && \text{(using (46))}
\end{aligned}$$

or in matrix notation $\mathbf{B} = \mathbf{N}\mathbf{R} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{R}$. The latter equality comes from (42). ★

Limit of $\mathbf{P}^n$

Combining the results above, we find

$$\mathbf{P}^n \to \begin{bmatrix} \mathbf{0} & \mathbf{B} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \quad \text{as } n \to \infty \ . \tag{49}$$

Expected absorption time

Let $T(i)$ be the *expected* time (= number of steps) before absorption starting from state i at time $n = 0$, namely,

$$T(i) = E\Big[\inf\{n \geq 0 \, : \, X(n) \in \{1^*, \ldots, M^*\} \,|\, X(0) = i\}\Big].$$

Clearly $T(i) = 0$ if i is an absorbing state. From now on we will only consider $T(i)$ when i is a transient state. Our objective is to find $T(i)$ for all $i \in \{1, 2, \ldots, N\}$. Define the column vector $\mathbf{T} = (T(i), i = 1, \ldots, N)^T$ of dimension N.

Proposition 10 (Expected absorption time)**.** *The column vector* $\mathbf{T}$ *is given by*

$$\mathbf{T} = (\mathbf{I} - \mathbf{A})^{-1}.\mathbf{1} \tag{50}$$

where $\mathbf{1}$ *is the column vector of dimension* N *which has all its entries equal to* 1*.* ■

Proof. Recall the definition of $n_{i,j}$, the expected number of visits to transient state j starting from transient state i. Therefore, we can write

$$T(i) = \sum_{j=1}^{N} n_{i,j}, \quad i \in \{1, 2, \ldots, N\} \ . \tag{51}$$

In matrix notation (51) writes

$$\mathbf{T} = \mathbf{N}.\mathbf{1} \tag{52}$$

where $\mathbf{N}$ is the fundamental matrix. The proof is readily found by combining (42) and (52). ★

Corollary 1 (Expected absorption time)**.** *The column vector* $\mathbf{T}$ *is the solution of the system*

$$\mathbf{T} = \mathbf{1} + \mathbf{A}\mathbf{T} \tag{53}$$

where $\mathbf{1}$ *is the column vector of dimension* N *which has all its entries equal to 1.*

3.2 Continuous-time absorbing Markov chains

This is the same framework as in the previous section but now we consider an absorbing, homogenous, continuous-time Markov chain (CTMC) $\{X(t), \, t \geq 0\}$ on the finite state-space

$$\mathcal{E} = \{1, 2, \ldots, N, 1^*, 2^*, \ldots, M^*\}.$$

States $1, 2, \ldots, N$ are transient and states $1^*, 2^*, \ldots, M^*$ are absorbent.

Let $q_{i,j}$ be the transition rate from state i to state j for $i \neq j$ and denote by $-q_{i,i}$ the transition rate out of state i. We know from the section on CTMC that $q_{i,i} = -\sum_{j \in \mathcal{E}, j \neq i} q_{i,j}$ for all $i \in \mathcal{E}$. We also know that $-q_{i,j}/q_{i,i}$ is the probability that the CTMC enters state $j \neq i$ when leaving state i.

Note that $q_{i,j} = 0$ for any state $j \in \mathcal{E}$, if i is an absorbing state. The infinitesimal generator of the CTMC can be written as follows

$$\mathbf{Q} = \begin{bmatrix} \tilde{\mathbf{Q}} & \tilde{\mathbf{R}} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

where we have used the N-by-N matrix $\tilde{\mathbf{Q}}$ defined by

$$\tilde{\mathbf{Q}} = [q_{i,j}]_{1 \leq i,j \leq N} \ , \tag{54}$$

and the N-by-M matrix $\tilde{\mathbf{R}} = [\tilde{r}_{i,j}]_{1 \leq i \leq N, 1^* \leq j \leq M^*}$, where $\tilde{r}_{i,j}$ is the transition rate from transient state i to absorbing state j.

Embedded Markov chain at jump times

Let us observe the Markov chain $\{X(t),\, t \geq 0\}$ at *jump times*. We obtain an *embedded* Markov chain that is a homogeneous, absorbing, DTMC, whose transition probability matrix is

$$\mathbf{P} = \Big[p_{i,j}\Big]_{i,j,\in\mathcal{E}} \ , \qquad \text{with} \quad p_{i,j} = \begin{cases} -q_{i,j}/q_{i,i} & \text{if } i \in \{1, \ldots, N\}, j \in \mathcal{E}, j \neq i \\ 1 & \text{if } j = i \in \{1^*, \ldots, M^*\} \text{ (by convention),} \\ 0 & \text{otherwise.} \end{cases}$$

We can write, as in Section 3.1,

$$\mathbf{P} = \begin{bmatrix} \mathbf{A} & \mathbf{R} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \ .$$

In this setting, the (i,j)-entry of the matrix $\mathbf{A} = [a_{i,j}]_{1 \leq i,j \leq N}$ is

$$a_{i,j} = \begin{cases} -\frac{q_{i,j}}{q_{i,i}} & \text{if } j \neq i, \\ 0 & \text{if } j = i. \end{cases} \tag{55}$$

The (i,j)-entry of the matrix $\mathbf{R} = [r_{i,j}]_{1 \leq i \leq N, 1^* \leq j \leq M^*}$ is

$$r_{i,j} = -\frac{q_{i,j}}{q_{i,i}} \tag{56}$$

for $i \in \{1, \ldots, N\}$ and $j \in \{1^*, \ldots, M^*\}$.

Mean number of visits and absorption probabilities

Given matrix $\mathbf{P}$ of the embedded Markov chain, we know by Proposition 8 that $n_{i,j}$, the expected number of visits to transient state j starting from transient state i, is given by the (i,j)-entry of the inverse of the matrix $\mathbf{I} - \mathbf{A}$. In other words, $\mathbf{N} = [n_{i,j}]_{1 \leq i,j \leq N} = (\mathbf{I} - \mathbf{A})^{-1}$.

By Proposition 9, the probability that the chain is absorbed in state $j \in \{1^*, \ldots, M^*\}$ given that it initiated in state $i \in \{1, \ldots, N\}$ is

$$b_{i,j} = \sum_{k=1}^{N} n_{i,k} r_{k,j}.$$

In other words, $\mathbf{B} = [b_{i,j}]_{1 \leq i \leq N, 1^* \leq j \leq M^*} = \mathbf{NR} = (\mathbf{I} - \mathbf{A})^{-1}\mathbf{R}$.

Expected absorption time

The embedded Markov chain spends in a transient state $i \in \{1, \ldots, N\}$ a time that is not constant but random having an exponential distribution with rate $-q_{i,i}$. Unlike the case of absorbing discrete-time Markov chains, the number of visits to a transient state and the expected time spent in that state are two different things.

Let $t_{i,j}$ be the expected time spent in transient state j before absorption starting from transient state i at time $t = 0$. Similarly to what was done for an absorbing DTMC, let $T(i)$ be the expected time before absorption starting from state i.

Define the $N \times N$ matrix $\tilde{\mathbf{T}} = \left[t_{i,j}\right]_{1 \leq i,j \leq N}$ and the column vector $\mathbf{T} = (T(i), i = 1, \ldots, N)^T$ of dimension N.

Below is the main result:

Proposition 11 (Expected absorption time for an absorbing CTMC)**.**

$$\tilde{\mathbf{T}} = -\tilde{\mathbf{Q}}^{-1} \ , \tag{57}$$
$$\mathbf{T} = -\tilde{\mathbf{Q}}^{-1}.\mathbf{1} \ , \tag{58}$$

where $\tilde{\mathbf{Q}}$ *is defined in* (54) *and* $\mathbf{1}$ *is the column vector of dimension* N *which has all its entries equal to* 1. ■

Proof. The time in state j follows an exponential distribution with mean $-1/q_{j,j}$. Since the expected number of visits to transient state j starting from transient state i is $n_{i,j}$, we deduce that the expected time spent in transient state j before absorption starting from transient state i is

$$t_{i,j} = n_{i,j} \times \frac{-1}{q_{j,j}}, \quad i = 1, 2, \ldots, N \ . \tag{59}$$

We introduce the notation $\mathbf{diag}(x_1, \ldots, x_n)$ to refer to an n-by-n diagonal matrix whose (i,i)-entry is x_i. Defining $\mathbf{D} = \mathbf{diag}(1/q_{1,1}, \ldots, 1/q_{N,N})$, (59) can be rewritten in matrix form

$$\tilde{\mathbf{T}} = -\mathbf{N}\,\mathbf{D} = -(\mathbf{I} - \mathbf{A})^{-1}\,\mathbf{D} = -\left(\mathbf{D}^{-1}\,(\mathbf{I} - \mathbf{A})\right)^{-1}.$$

Since $\mathbf{D}^{-1} = \mathbf{diag}(q_{1,1}, \ldots, q_{N,N})$, it is easily seen from (55) and the definition of the matrix $\tilde{\mathbf{Q}}$ that

$$\mathbf{D}^{-1}(\mathbf{I} - \mathbf{A}) = \tilde{\mathbf{Q}}$$

from which we deduce that $\tilde{\mathbf{T}}$ is the inverse of the matrix $-\tilde{\mathbf{Q}}$, concluding the proof of (57). From the definitions of $t_{i,j}$ and $T(i)$, it is clear that $T(i) = \sum_{j=1}^{N} t_{i,j}$. We can readily write in matrix notation

$$\mathbf{T} = \tilde{\mathbf{T}}.\mathbf{1} \tag{60}$$

where $\mathbf{1}$ is an N-sized column vector of 1's. Combining (60) with (57) gives (58). ★

Observe that in Proposition 11, the matrix $-\tilde{\mathbf{Q}}^{-1}$ plays the same role as the fundamental matrix in an absorbing DTMC (see Equation (52)).

Corollary 2 (Expected absorption time for an absorbing CTMC)**.** *The column vector* $\mathbf{T}$ *is the solution of the system*

$$\tilde{\mathbf{Q}}\mathbf{T} = -\mathbf{1}\ . \tag{61}$$

Example 4. Consider an absorbing homogeneous CTMC with generator

$$\mathbf{Q} = \begin{bmatrix} -3 & 2 & 1 \\ 2 & -3 & 1 \\ 0 & 0 & 0 \end{bmatrix} .$$

From $\mathbf{Q}$ and the transition diagram, it is clear that State 3 is absorbing so we must have $T(3) = 0$. The times until absorption when initially in states 1 or 2 are the solution of the system $\tilde{\mathbf{Q}}\mathbf{T} = -\mathbf{1}$. Namely

$$\left.\begin{array}{r} -3\,T(1) + 2\,T(2) = -1 \\ 2\,T(1) - 3\,T(2) = -1 \end{array}\right\} \Rightarrow T(1) = T(2) = 1.$$

The expected time spent in a transient state is given by

$$\tilde{\mathbf{T}} = -\tilde{\mathbf{Q}}^{-1} = \frac{1}{5}\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} .$$

The embedded Markov chain at jump times has transition matrix

$$\mathbf{P} = \begin{bmatrix} 0 & \frac{2}{3} & \frac{1}{3} \\ \frac{2}{3} & 0 & \frac{1}{3} \\ 0 & 0 & 1 \end{bmatrix} \Rightarrow \mathbf{A} = \begin{bmatrix} 0 & \frac{2}{3} \\ \frac{2}{3} & 0 \end{bmatrix} , \mathbf{R} = \begin{bmatrix} \frac{1}{3} \\ \frac{1}{3} \end{bmatrix} .$$

The fundamental matrix giving the expected number of visits in transient states is

$$\mathbf{N} = \begin{bmatrix} \frac{9}{5} & \frac{6}{5} \\ \frac{6}{5} & \frac{9}{5} \end{bmatrix} .$$

The absorption probabilities are

$$\mathbf{B} = \mathbf{N}\mathbf{R} = \frac{3}{5}\begin{bmatrix} 3 & 2 \\ 2 & 3 \end{bmatrix} \frac{1}{3} \begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix} .$$

♦