Introduction to Markov Chains

George Voutsadakis[1]

[1]Mathematics and Computer Science
Lake Superior State University

LSSU Math 500

1. Continuous-time Markov chains II
   - Basic Properties
   - Class Structure
   - Hitting Times and Absorption Probabilities
   - Recurrence and Transience
   - Invariant Distributions
   - Convergence to Equilibrium
   - Time Reversal
   - Ergodic Theorem

Subsection 1

Basic Properties

Q-Matrices Revisited

- Let I be a countable set.
- Recall that a Q-**matrix** on I is a matrix

$$Q = (q_{ij} : i, j \in I),$$

satisfying the following conditions:

(i) $0 \leq -q_{ii} < \infty$, for all i;
(ii) $q_{ij} \geq 0$, for all $i \neq j$;
(iii) $\sum_{j \in I} q_{ij} = 0$, for all i.

- We set $q_i = q(i) = -q_{ii}$.

The Jump Matrix

- Associated to any Q-matrix is a **jump matrix**

$$\Pi = (\pi_{ij} : i, j \in I),$$

defined as follows:

  - For all i,

$$\pi_{ii} = \begin{cases} 0, & \text{if } q_i \neq 0, \\ 1, & \text{if } q_i = 0. \end{cases}$$

  - For all $i \neq j$,

$$\pi_{ij} = \begin{cases} \frac{q_{ij}}{q_i}, & \text{if } q_i \neq 0, \\ 0, & \text{if } q_i = 0. \end{cases}$$

- Note that Π is a stochastic matrix.

Sub-Stochastic Matrices

- A **sub-stochastic matrix** on I is a matrix
$$P = (p_{ij} : i, j \in I),$$
with nonnegative entries and such that
$$\sum_{j \in I} p_{ij} \leq 1, \quad \text{for all } i.$$
- Associated to any Q-matrix is a **semigroup** $(P(t) : t \geq 0)$ of sub-stochastic matrices
$$P(t) = (p_{ij}(t) : i, j \in I).$$
- As the name implies, we have
$$P(s)P(t) = P(s+t), \quad \text{for all } s, t \geq 0.$$

Basic Terms

- We assume familiarity with the following terms introduced in the preceding set:
  - *Minimal right-continuous random process*;
  - *Jump times*;
  - *Holding times*;
  - *Jump chain*;
  - *Explosion*.
- Briefly, a *right-continuous process*

$$(X_t)_{t \geq 0}$$

  - Runs through a sequence of states $Y_0, Y_1, Y_2, \ldots$;
  - Is held in these states for times $S_1, S_2, S_3, \ldots$, respectively;
  - Jumps to the next state at times $J_1, J_2, J_3, \ldots$.
- Thus $J_n = S_1 + \cdots + S_n$.

Basic Terms (Cont'd)

- The discrete-time process

$$(Y_n)_{n\geq 0}$$

  is the *jump chain*.
- $(S_n)_{n\geq 1}$ are the *holding times*.
- $(J_n)_{n\geq 1}$ are the *jump times*.
- The *explosion time* ζ is given by

$$\zeta = \sum_{n=1}^{\infty} S_n = \lim_{n\to\infty} J_n.$$

- For a minimal process we take a new state ∞ and insist that

$$X_t = \infty, \quad \text{for all } t \geq \zeta.$$

- An important point is that a minimal right-continuous process is determined by its jump chain and holding times.

Data for Continuous-Time Markov Chain

- The data for a continuous-time Markov chain

$$(X_t)_{t\geq 0}$$

  are:
  - A distribution λ;
  - A Q-matrix Q.
- These play the following roles.
  - The distribution λ gives the **initial distribution**, the distribution of X_0.
  - The Q-matrix is known as the **generator matrix** of $(X_t)_{t\geq 0}$.
    It determines how the process evolves from its initial state.

First Description of a Continuous-Time Markov Chain

- We established that there are two different, but equivalent, ways to describe how the process evolves.
- The first, in terms of jump chain and holding times, states that:
  (a) $(Y_n)_{n\geq 0}$ is Markov(λ, Π);
  (b) Conditional on $Y_0 = i_0, \ldots, Y_{n-1} = i_{n-1}$, the holding times $S_1, \ldots, S_n$ are independent exponential random variables of parameters $q_{i_0}, \ldots, q_{i_{n-1}}$.
- Put more simply, given that the chain starts at i:
  - It waits there for an exponential time of parameter q_i;
  - Then jumps to a new state, choosing state j with probability π_{ij}.
  - It then starts afresh, forgetting what has happened before.

Second Description of a Continuous-Time Markov Chain

- The second description, in terms of the semigroup, states that the finite dimensional distributions of the process are given by:
  (c) For all $n = 0, 1, 2, \ldots$, all times $0 \leq t_0 \leq t_1 \leq \cdots \leq t_{n+1}$ and all states $i_0, i_1, \ldots, i_{n+1}$,
  $$\mathbb{P}(X_{t_{n+1}} = i_{n+1} | X_{t_0} = i_0, \ldots, X_{t_n} = i_n) = p_{i_n i_{n+1}}(t_{n+1} - t_n).$$
- Put more simply, given that the chain starts at i:
  - By time t it is found in state j with probability $p_{ij}(t)$;
  - It then starts afresh, forgetting what has happened before.
- In the case where $\widetilde{p}_{i\infty}(t) := 1 - \sum_{j \in I} p_{ij}(t) > 0$ the chain is found at ∞ with probability $\widetilde{p}_{i\infty}(t)$.
- The semigroup $P(t)$ is the **transition matrix** of the chain.
- Its entries $p_{ij}(t)$ are the **transition probabilities**.

Remarks on the Second Description

- The second description implies that, for all $h > 0$, the discrete skeleton
$$(X_{nh})_{n \geq 0}$$
is Markov$(\lambda, P(h))$.
- Strictly, in the explosive case, that is, when $P(t)$ is strictly sub-stochastic, we should say
$$\text{Markov}(\widetilde{\lambda}, \widetilde{P}(h)),$$
where $\widetilde{\lambda}$ and $\widetilde{P}(h)$ are defined on $I \cup \{\infty\}$, extending λ and $P(h)$ by:
  - $\widetilde{\lambda}_\infty = 0$;
  - $\widetilde{p}_{\infty j}(h) = 0$.
- Usually, there is no danger of confusion in using the simpler notation.

Relation Between P and Q

- Note that we have not yet said how the semigroup $P(t)$ is associated to the Q-matrix Q, except via the process!
- We recall that the semigroup is characterized as the minimal non-negative solution of the backward equation

$$P'(t) = QP(t), \quad P(0) = I.$$

- In component form

$$p'_{ij}(t) = \sum_{k \in I} q_{ik} p_{kj}(t), \quad p_{ij}(0) = \delta_{ij}.$$

- The semigroup is also the minimal non-negative solution of the forward equation

$$P'(t) = P(t)Q, \quad P(0) = I.$$

- In the case where I is finite, $P(t)$ is simply the matrix exponential e^{tQ}, and is the unique solution of the backward and forward equations.

Subsection 2

Class Structure

Leading and Communicating

- Recall we deal only with minimal chains, those that die after explosion.
- The **class structure** is simply the discrete-time class structure of the jump chain $(Y_n)_{n\geq 0}$.
- We say that i **leads to** j and write $i \to j$ if

$$\mathbb{P}_i(X_t = j \text{ for some } t \geq 0) > 0.$$

- We say i **communicates with** j and write $i \leftrightarrow j$ if both $i \to j$ and $j \to i$.
- Communication is an equivalence relation between states.

Inherited Notions

- The notions of communicating class, closed class, absorbing state and irreducibility are inherited from the jump chain.
- A **communicating class** is an equivalence class of the communicating equivalence relation $\leftrightarrow$.
- A class C is **closed** if
$$i \in C \quad \text{and} \quad i \to j \quad \text{imply} \quad j \in C.$$
- Thus, a closed class is one from which there is no escape.
- A state i is **absorbing** if $\{i\}$ is a closed class.
- A chain whose state space I consists of a single communicating class is called **irreducible**.

Characterization of the Leading Relation

Theorem

For distinct states i and j the following are equivalent:

(i) $i \to j$;

(ii) $i \to j$ for the jump chain;

(iii) $q_{i_0 i_1} q_{i_1 i_2} \cdots q_{i_{n-1} i_n} > 0$ for some states $i_0, i_1, \ldots, i_n$ with $i_0 = i$, $i_n = j$;

(iv) $p_{ij}(t) > 0$, for all $t > 0$;

(v) $p_{ij}(t) > 0$, for some $t > 0$.

- Implications (iv)⇒(v)⇒(i)⇒(ii) are clear.
  Suppose Condition (ii) holds.
  Then, by a previous theorem, there are states $i_0, i_1, \ldots, i_n$ with $i_0 = i$, $i_n = j$ and
  $$\pi_{i_0 i_1} \pi_{i_1 i_2} \cdots \pi_{i_{n-1} i_n} > 0.$$
  This implies Condition (iii).

Characterization of the Leading Relation (Cont'd)

- Finally, suppose Condition (iii) holds.

  Note that, if $q_{ij} > 0$, then, for all $t > 0$,

$$\begin{aligned} p_{ij}(t) &\geq \mathbb{P}_i(J_1 \leq t, Y_1 = j, S_2 > t) \\ &= (1 - e^{-q_i t})\pi_{ij} e^{-q_j t} \\ &> 0. \end{aligned}$$

  Thus, for all $t > 0$,

$$p_{ij}(t) \geq p_{i_0 i_1}\left(\frac{t}{n}\right) \cdots p_{i_{n-1} i_n}\left(\frac{t}{n}\right) > 0.$$

  So (iv) holds.

Comparison with Discrete Time

- Condition (iv) of the Theorem shows that the situation is simpler than in discrete-time.
- In discrete time, it may be possible to reach a state, but:
  - Only after a certain length of time;
  - And then only periodically.

Subsection 3

Hitting Times and Absorption Probabilities

Hitting Time and Absorption Probability

- Let $(X_t)_{t \geq 0}$ be a Markov chain with generator matrix Q.
- The **hitting time** of a subset A of I is the random variable D^A defined by
$$D^A(\omega) = \inf \{t \geq 0 : X_t(\omega) \in A\},$$
with the usual convention that $\inf \emptyset = \infty$.
- We emphasize that $(X_t)_{t \geq 0}$ is minimal.
- So if H^A is the hitting time of A for the jump chain, then:
  - $\{H^A < \infty\} = \{D^A < \infty\}$;
  - On this set we have $D^A = J_{H^A}$.
- The probability, starting from i, that $(X_t)_{t \geq 0}$ ever hits A is then
$$h_i^A = \mathbb{P}_i(D^A < \infty) = \mathbb{P}_i(H^A < \infty).$$
- When A is a closed class, h_i^A is called the **absorption probability**.

Vector of Hitting Probabilities

Theorem

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

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

- Apply a previous theorem to the jump chain and rewrite

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

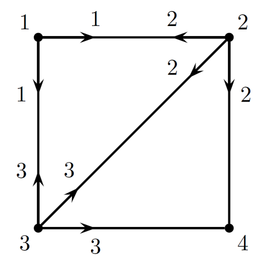
in terms of Q.

Average Hitting Times

- The average time taken, starting from i, for $(X_t)_{t \geq 0}$ to reach A is given by
$$k_i^A = \mathbb{E}_i(D^A).$$
- In calculating k_i^A we have to take account of the holding times.
- So the relationship to the discrete-time case is not quite as simple.

Example

- Consider the Markov chain $(X_t)_{t\geq 0}$ with the diagram shown. How long on average does it take to get from 1 to 4?



Set $k_i = \mathbb{E}_i(\text{time to get to } 4)$.

On starting in 1:

- We spend an average time $q_1^{-1} = \frac{1}{2}$ in 1;
- Then jump with equal probability to 2 or 3.

Thus,

$$k_1 = \frac{1}{2} + \frac{1}{2}k_2 + \frac{1}{2}k_3.$$

Similarly

$$\begin{array}{rcl} k_2 & = & \frac{1}{6} + \frac{1}{3}k_1 + \frac{1}{3}k_3, \\ k_3 & = & \frac{1}{9} + \frac{1}{3}k_1 + \frac{1}{3}k_2. \end{array}$$

On solving these linear equations we find $k_1 = \frac{17}{12}$.

Example

- We consider the Markov chain with state space $\{1, 2, 3, 4\}$ and generator matrix
$$\begin{pmatrix} -1 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & -\frac{1}{2} & 0 & \frac{1}{4} \\ \frac{1}{6} & 0 & -\frac{1}{3} & \frac{1}{6} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$
We calculate the probability of hitting 3 starting from 1.

  We set
$$h_i = \text{probability of hitting 3 starting from } i.$$

Example (Cont'd)

- We must find the minimal nonnegative solution of the system

$$\left\{\begin{array}{rcl} h_3 & = & 1 \\ -h_1 + \frac{1}{2}h_2 + \frac{1}{2}h_3 & = & 0 \\ \frac{1}{4}h_1 - \frac{1}{2}h_2 + \frac{1}{4}h_4 & = & 0 \end{array}\right\} \Rightarrow \left\{\begin{array}{rcl} h_2 & = & 2h_1 - 1 \\ h_3 & = & 1 \\ h_4 & = & 3h_1 - 2 \end{array}\right\}$$

$$\Rightarrow \left\{\begin{array}{rcl} h_1 & = & \frac{2}{3} \\ h_2 & = & \frac{1}{3} \\ h_3 & = & 1 \\ h_4 & = & 0 \end{array}\right\}$$

Thus, the probability of hitting 3 starting from 1 is $\frac{2}{3}$.

Vector of Expected Hitting Times

Theorem

Assume that $q_i > 0$, for all $i \notin A$. The vector of expected hitting times $k^A = (k_i^A : i \in I)$ is the minimal non-negative solution to the system of linear equations

$$\begin{cases} k_i^A = 0, & \text{for } i \in A, \\ -\sum_{j \in I} q_{ij} k_j^A = 1, & \text{for } i \notin A. \end{cases}$$

- First we show that k^A satisfies the system of equations.

  Suppose $X_0 = i \in A$.

  Then $D^A = 0$.

  So $k_i^A = 0$.

Vector of Expected Hitting Times (Cont'd)

- Suppose $X_0 = i \notin A$.
  Then $D^A \geq J_1$.
  By the Markov Property of the jump chain,
  $$\mathbb{E}_i(D^A - J_1 | Y_1 = j) = \mathbb{E}_j(D^A).$$
  So we get
  $$\begin{array}{rcl} k_i^A & = & \mathbb{E}_i(D^A) \\ & = & \mathbb{E}_i(J_1) + \sum_{j \neq i} \mathbb{E}(D^A - J_1 | Y_1 = j) \mathbb{P}_i(Y_1 = j) \\ & = & q_i^{-1} + \sum_{j \neq i} \pi_{ij} k_j^A. \end{array}$$
  Rewriting, $q_i(k_i^A - \sum_{j \neq i} \pi_{ij} k_j^A) = 1$.
  Equivalently,
  $$-\sum_{j \in I} q_{ij} k_j^A = 1.$$

Vector of Expected Hitting Times (Cont'd)

- Suppose now that $y = (y_i : i \in I)$ is another solution of the system.
  Then $k_i^A = y_i = 0$ for $i \in A$.
  Suppose $i \notin A$. Then we have

$$\begin{array}{rcl} y_i & = & q_i^{-1} + \sum_{j \notin A} \pi_{ij} y_j \\ & = & q_i^{-1} + \sum_{j \notin A} \pi_{ij} (q_j^{-1} + \sum_{k \notin A} \pi_{jk} y_k) \\ & = & \mathbb{E}_i(S_1) + \mathbb{E}_i(S_2 1_{\{H^A \geq 2\}}) + \sum_{j \notin A} \sum_{k \notin A} \pi_{ij} \pi_{jk} y_k. \end{array}$$

  By repeated substitution for y in the final term we obtain after n steps

$$y_i = \mathbb{E}_i(S_1) + \cdots + \mathbb{E}_i(S_n 1_{\{H^A \geq n\}}) + \sum_{j_1 \notin A} \cdots \sum_{j_n \notin A} \pi_{ij_1} \cdots \pi_{j_{n-1} j_n} y_{j_n}.$$

Vector of Expected Hitting Times (Cont'd)

- We obtained

$$y_i = \mathbb{E}_i(S_1) + \cdots + \mathbb{E}_i(S_n 1_{\{H^A \geq n\}}) + \sum_{j_1 \notin A} \cdots \sum_{j_n \notin A} \pi_{ij_1} \cdots \pi_{j_{n-1}j_n} y_{j_n}.$$

So, if y is non-negative, using $H^A \wedge n = \min\{H^A, n\}$,

$$y_i \geq \sum_{m=1}^{n} \mathbb{E}_i(S_m 1_{H^A \geq m}) = \mathbb{E}_i\left(\sum_{m=1}^{H^A \wedge n} S_m\right).$$

Now $\sum_{m=1}^{H^A} S_m = D^A$.

By Monotone Convergence,

$$y_i \geq \mathbb{E}_i(D^A) = k_i^A.$$

Example

- We consider again the Markov chain with state space $\{1, 2, 3, 4\}$ and generator matrix

$$\begin{pmatrix} -1 & \frac{1}{2} & \frac{1}{2} & 0 \\ \frac{1}{4} & -\frac{1}{2} & 0 & \frac{1}{4} \\ \frac{1}{6} & 0 & -\frac{1}{3} & \frac{1}{6} \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

We calculate the expected time of hitting 4 starting from 1.

We set

$$k_i = \text{expected time of hitting 4 starting from } i.$$

Example (Cont'd)

- We must find the minimal nonnegative solution of the system

$$\left\{\begin{array}{rcl} k_4 & = & 0 \\ k_1 - \frac{1}{2}k_2 - \frac{1}{2}k_3 & = & 1 \\ -\frac{1}{4}k_1 + \frac{1}{2}k_2 - \frac{1}{4}k_4 & = & 1 \\ -\frac{1}{6}k_1 + \frac{1}{3}k_3 - \frac{1}{6}k_4 & = & 1 \end{array}\right\} \Rightarrow \left\{\begin{array}{rcl} k_1 & = & 7 \\ k_2 & = & \frac{1}{2}k_1 + 2 \\ k_3 & = & \frac{1}{2}k_1 + 3 \\ k_4 & = & 0 \end{array}\right\}$$

$$\Rightarrow \left\{\begin{array}{rcl} k_1 & = & 7 \\ k_2 & = & \frac{11}{2} \\ k_3 & = & \frac{13}{2} \\ k_4 & = & 0 \end{array}\right\}$$

Thus, the expected time of hitting 4 starting from 1 is 7.

Subsection 4

Recurrence and Transience

Recurrent and Transient States

- Let $(X_t)_{t\geq 0}$ be a Markov chain with generator matrix Q.
- We insist $(X_t)_{t\geq 0}$ be minimal.
- We say a state i is **recurrent** if

$$\mathbb{P}_i(\{t \geq 0 : X_t = i\} \text{ is unbounded}) = 1.$$

- We say that i is **transient** if

$$\mathbb{P}_i(\{t \geq 0 : X_t = i\} \text{ is unbounded}) = 0.$$

- Note that, if $(X_t)_{t\geq 0}$ can explode starting from i, then i is certainly not recurrent.

Recurrence and Transience and the Jump Chain

Theorem

(i) If i is recurrent for the jump chain $(Y_n)_{n\geq 0}$, then i is recurrent for $(X_t)_{t\geq 0}$;

(ii) If i is transient for the jump chain, then i is transient for $(X_t)_{t\geq 0}$;

(iii) Every state is either recurrent or transient;

(iv) Recurrence and transience are class properties.

(i) Suppose i is recurrent for $(Y_n)_{n\geq 0}$.

If $X_0 = i$, then $(X_t)_{t\geq 0}$ does not explode.

Moreover, $J_n \to \infty$, by a previous theorem.

Also $X(J_n) = Y_n = i$ infinitely often.

So $\{t \geq 0 : X_t = i\}$ is unbounded, with probability 1.

Recurrence and Transience and the Jump Chain (Cont'd)

(ii) Suppose i is transient for $(Y_n)_{n\geq 0}$.

If $X_0 = i$, then

$$N = \sup\{n \geq 0 : Y_n = i\} < \infty.$$

So $\{t \geq 0 : X_t = i\}$ is bounded by $J(N+1)$.

Now $(Y_n : n \geq N)$ cannot include an absorbing state.

So $J(N+1)$ is finite, with probability 1.

For (iii) and (iv), we apply previous theorems to the jump chain.

Conditions for Recurrence and Transience

- We denote by T_i the **first passage time** of $(X_t)_{t \geq 0}$ to state i, defined by
$$T_i(\omega) = \inf \{t \geq J_1(\omega) : X_t(\omega) = i\}.$$

Theorem

The following dichotomy holds:

(i) If $q_i = 0$ or $\mathbb{P}_i(T_i < \infty) = 1$, then i is recurrent and
$$\int_0^\infty p_{ii}(t)dt = \infty;$$

(ii) If $q_i > 0$ and $\mathbb{P}_i(T_i < \infty) < 1$, then i is transient and
$$\int_0^\infty p_{ii}(t)dt < \infty.$$

Conditions for Recurrence and Transience (Cont'd)

- If $q_i = 0$, then $(X_t)_{t \geq 0}$ cannot leave i. So:
  - i is recurrent;
  - $p_{ii}(t) = 1$, for all t;
  - $\int_0^\infty p_{ii}(t)dt = \infty$.

  Suppose then that $q_i > 0$.

  Let N_i be the first passage time of the jump chain $(Y_n)_{n \geq 0}$ to state i.

  Then $\mathbb{P}_i(N_i < \infty) = \mathbb{P}_i(T_i < \infty)$.

  So, by the preceding theorem and the corresponding result for the jump chain, i is recurrent if and only if $\mathbb{P}_i(T_i < \infty) = 1$.

Conditions for Recurrence and Transience (Cont'd)

- Write $\pi_{ij}^{(n)}$ for the (i,j) entry in Π^n.

  We show that $\int_0^\infty p_{ii}(t)dt = \frac{1}{q_i}\sum_{n=0}^\infty \pi_{ii}^{(n)}$.

  Then i is recurrent if and only if $\int_0^\infty p_{ii}(t)dt = \infty$, by the preceding theorem and the corresponding result for the jump chain.

  We use Fubini's Theorem:

$$\begin{aligned}\int_0^\infty p_{ii}(t)dt &= \int_0^\infty \mathbb{E}_i(1_{\{X_t=i\}})dt \\ &= \mathbb{E}_i \int_0^\infty 1_{\{X_t=i\}}dt \\ &= \mathbb{E}_i \sum_{n=0}^\infty S_{n+1}1_{\{Y_n=i\}} \\ &= \sum_{n=0}^\infty \mathbb{E}_i(S_{n+1}|Y_n=i)\mathbb{P}_i(Y_n=i) \\ &= \frac{1}{q_i}\sum_{n=0}^\infty \pi_{ii}^{(n)}.\end{aligned}$$

Recurrence, Transience and Samplings

Theorem

Let $h > 0$ be given and set $Z_n = X_{nh}$.

(i) If i is recurrent for $(X_t)_{t \geq 0}$, then i is recurrent for $(Z_n)_{n \geq 0}$.

(ii) If i is transient for $(X_t)_{t \geq 0}$, then i is transient for $(Z_n)_{n \geq 0}$.

- Claim (ii) is obvious.
  We now prove Claim (i).
  By the Markov Property, for $nh \leq t < (n+1)h$, we have
  $$p_{ii}((n+1)h) \geq e^{-q_i h} p_{ii}(t).$$
  By Monotone Convergence,
  $$\int_0^\infty p_{ii}(t) dt \leq h e^{q_i h} \sum_{n=1}^{\infty} p_{ii}(nh).$$
  So the result follows by previous theorems.

Subsection 5

Invariant Distributions

Invariant Measures

- We say that λ is **invariant** if $\lambda Q = 0$.

Theorem

Let Q be a Q-matrix with jump matrix Π and let λ be a measure. The following are equivalent:

(i) λ is invariant;

(ii) $\mu\Pi = \mu$ where $\mu_i = \lambda_i q_i$.

- We have $q_i(\pi_{ij} - \delta_{ij}) = q_{ij}$, for all i, j. So

$$(\mu(\Pi - I))_j = \sum_{i\in I} \mu_i(\pi_{ij} - \delta_{ij}) = \sum_{i\in I} \lambda_i q_{ij} = (\lambda Q)_j.$$

- This connection allows using the existence and uniqueness results related to the discrete-time processes.

Irreducible and Recurrent Q-Matrices

Theorem

Suppose that Q is irreducible and recurrent. Then Q has an invariant measure λ which is unique up to scalar multiples.

- Let us exclude the trivial case $I = \{i\}$.

  Then irreducibility forces $q_i > 0$ for all i.

  By previous theorems, Π is irreducible and recurrent.

  Then, by theorems addressing the discrete time case, Π has an invariant measure μ, which is unique up to scalar multiples.

  Setting $\lambda_i = \frac{\mu_i}{q_i}$, we obtain, by the preceding theorem, an invariant measure unique up to scalar multiples.

Positive Recurrence

- Recall that a state i is **recurrent** if $q_i = 0$ or $\mathbb{P}_i(T_i < \infty) = 1$.
- State i is **positive recurrent** if $q_i = 0$ or the **expected return time** $m_i = \mathbb{E}_i(T_i)$ is finite.
- Otherwise a recurrent state i is called **null recurrent**.

Theorem

Let Q be an irreducible Q-matrix. Then the following are equivalent:

(i) Every state is positive recurrent;

(ii) Some state i is positive recurrent;

(iii) Q is non-explosive and has an invariant distribution λ.

Moreover, when (iii) holds we have $m_i = \frac{1}{\lambda_i q_i}$, for all i.

- We again exclude the trivial case $I = \{i\}$.

  Irreducibility forces $q_i > 0$, for all i. Obviously, (i) implies (ii).

Positive Recurrence (Cont'd)

- Define $\mu^i = (\mu^i_j : j \in I)$ by
$$\mu^i_j = \mathbb{E}_i \int_0^{T_i \wedge \zeta} 1_{\{X_s = j\}} ds,$$
where $T_i \wedge \zeta$ denotes the minimum of T_i and ζ.
By Monotone Convergence, $\sum_{j \in I} \mu^i_j = \mathbb{E}_i(T_i \wedge \zeta)$.
Let N_i be the first passage time of the jump chain to state i.
By Fubini's Theorem,
$$\begin{aligned} \mu^i_j &= \mathbb{E}_i \sum_{n=0}^{\infty} S_{n+1} 1_{\{Y_n = j, n < N_i\}} \\ &= \sum_{n=0}^{\infty} \mathbb{E}_i(S_{n+1} | Y_n = j) \mathbb{E}_i(1_{\{Y_n = j, n < N_i\}}) \\ &= q_j^{-1} \mathbb{E}_i \sum_{n=0}^{\infty} 1_{\{Y_n = j, n < N_i\}} \\ &= q_j^{-1} \mathbb{E}_i \sum_{n=0}^{N_i - 1} 1_{\{Y_n = j\}} = \frac{\gamma^i_j}{q_j}, \end{aligned}$$
where γ^i_j is the expected time in j between visits to i for the jump chain.

Positive Recurrence ((ii)⇒(iii))

- Suppose Condition (ii) holds.

  Let state i be positive recurrent.

  Then i is certainly recurrent.

  By a previous theorem, the jump chain is recurrent and Q is non-explosive.

  Also, by a previous theorem, $\gamma^i \Pi = \gamma^i$.

  So $\mu^i Q = 0$, by one of the preceding theorems.

  But μ^i has finite total mass

  $$\sum_{j \in I} \mu_j^i = \mathbb{E}_i(T_i) = m_i.$$

  So we obtain an invariant distribution λ by setting $\lambda_j = \frac{\mu_j^i}{m_i}$.

Positive Recurrence ((iii)⇒(i))

- Suppose Condition (iii) holds.

  Fix $i \in I$ and set $\nu_j = \frac{\lambda_j q_j}{\lambda_i q_i}$.

  Then $\nu_i = 1$ and $\nu \Pi = \nu$ by a previous theorem.

  Also by a previous theorem, $\nu_j \geq \gamma_j^i$, for all j.

  So we get

$$\begin{aligned} m_i &= \sum_{j \in I} \mu_j^i = \sum_{j \in I} \frac{\gamma_j^i}{q_j} \leq \sum_{j \in I} \frac{\nu_j}{q_j} \\ &= \sum_{j \in I} \frac{\lambda_j}{\lambda_i q_i} = \frac{1}{\lambda_i q_i} < \infty. \end{aligned}$$

  This shows that i is positive recurrent.

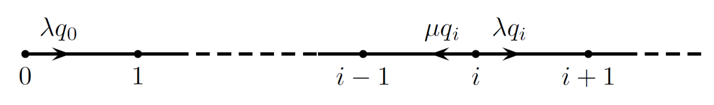
  To complete the proof we return to the preceding calculation armed with the knowledge that Q is recurrent.

  It follows that Π is recurrent, $\nu_j = \gamma_j^i$ and $m_i = \frac{1}{\lambda_i q_i}$, for all i.

Example

- The existence of an invariant distribution for a continuous-time Markov chain is not enough to guarantee positive recurrence, or even recurrence.

  Consider the Markov chain $(X_t)_{t\geq 0}$ on $\mathbb{Z}^+$ with the following diagram.



  - $q_i > 0$, for all i;
  - $0 < \lambda = 1 - \mu < 1$.

  The jump chain behaves as a simple random walk away from 0.

  So $(X_t)_{t\geq 0}$ is:

  - Recurrent, if $\lambda \leq \mu$;
  - Transient, if $\lambda > \mu$.

Example (Cont'd)

- To compute an invariant measure ν it is convenient to use the detailed balance equations $\nu_i q_{ij} = \nu_j q_{ji}$, for all i, j.

  In this case the non-zero equations read

  $$\nu_i \lambda q_i = \nu_{i+1} \mu q_{i+1}, \quad \text{for all } i.$$

  So a solution is given by $\nu_i = q_i^{-1} (\frac{\lambda}{\mu})^i$.

  If the jump rates q_i are constant, then ν can be normalized to produce an invariant distribution precisely when $\lambda < \mu$.

  Consider the case where $q_i = 2^i$, for all i, and $1 < \frac{\lambda}{\mu} < 2$.

  Then ν has finite total mass.

  So $(X_t)_{t \geq 0}$ has an invariant distribution.

  But $(X_t)_{t \geq 0}$ is also transient.

  Given the theorem, the only possibility is that $(X_t)_{t \geq 0}$ is explosive.

Characterization of Invariant Measures

Theorem

Let Q be irreducible and recurrent, and let λ be a measure. Let $s > 0$ be given. The following are equivalent:

(i) $\lambda Q = 0$;

(ii) $\lambda P(s) = \lambda$.

- There is a very simple proof in the case of finite state space.
  By the backward equation

$$\frac{d}{ds}\lambda P(s) = \lambda P'(s) = \lambda Q P(s).$$

  So $\lambda Q = 0$ implies $\lambda P(s) = \lambda P(0) = \lambda$, for all s.
  $P(s)$ is also recurrent.
  So $\mu P(s) = \mu$ implies that μ is proportional to λ.
  So $\mu Q = 0$.

Characterization of Invariant Measures

- For infinite state space, the interchange of differentiation with the summation involved in multiplication by λ is not justified.

  So an entirely different proof is needed.

  Since Q is recurrent, it is non-explosive, by a previous theorem.

  Moreover, $P(s)$ is recurrent, by a previous theorem.

  Hence, any λ satisfying (i) or (ii) is unique up to scalar multiples.

  Fix i and set

$$\mu_j = \mathbb{E}_i \int_0^{T_i} 1_{\{X_t=j\}} dt.$$

  By the proof of a previous theorem, $\mu Q = 0$.

  Thus, it suffices to show $\mu P(s) = \mu$.

Characterization of Invariant Measures (Infinite States)

- By the Strong Markov Property at T_i (which is a simple consequence of the Strong Markov Property of the jump chain),

$$\mathbb{E}_i \int_0^s 1_{\{X_t=j\}} dt = \mathbb{E}_i \int_{T_i}^{T_i+s} 1_{\{X_t=j\}} dt.$$

Hence, using Fubini's Theorem,

$$\begin{array}{rcl} \mu_j & = & \mathbb{E}_i \int_s^{s+T_i} 1_{\{X_t=j\}} dt \\ & = & \int_0^\infty \mathbb{P}_i(X_{s+t} = j, t < T_i) dt \\ & = & \int_0^\infty \sum_{k \in I} \mathbb{P}_i(X_t = k, t < T_i) p_{kj}(s) dt \\ & = & \sum_{k \in I} (\mathbb{E}_i \int_0^{T_i} 1_{\{X_t=k\}} dt) p_{kj}(s) \\ & = & \sum_{k \in I} \mu_k p_{kj}(s). \end{array}$$

Irreducibility, Non-Explosivity and Invariant Distributions

Theorem

Let Q be an irreducible non-explosive Q-matrix having an invariant distribution λ. If $(X_t)_{t\geq 0}$ is Markov(λ, Q), then so is $(X_{s+t})_{t\geq 0}$, for any $s \geq 0$.

- By the preceding theorem, for all i,

$$\mathbb{P}(X_s = i) = (\lambda P(s))_i = \lambda_i.$$

  So, by the Markov Property, conditional on $X_s = i$, $(X_{s+t})_{t\geq 0}$ is Markov(δ_i, Q).

Subsection 6

Convergence to Equilibrium

Estimate of Uniform Continuity for Transition Probabilities

Lemma

Let Q be a Q-matrix with semigroup $P(t)$. Then, for all $t, h \geq 0$,

$$|p_{ij}(t+h) - p_{ij}(t)| \leq 1 - e^{-q_i h}.$$

- We have

$$\begin{aligned} |p_{ij}(t+h) - p_{ij}(t)| &= \left|\sum_{k \in I} p_{ik}(h) p_{kj}(t) - p_{ij}(t)\right| \\ &= \left|\sum_{k \neq i} p_{ik}(h) p_{kj}(t) - (1 - p_{ii}(h)) p_{ij}(t)\right| \\ &\leq 1 - p_{ii}(h) \\ &\leq \mathbb{P}_i(J_1 \leq h) \\ &= 1 - e^{-q_i h}. \end{aligned}$$

Convergence to Equilibrium

Theorem (Convergence to Equilibrium)

Let Q be an irreducible non-explosive Q-matrix with semigroup $P(t)$, and having an invariant distribution λ. Then for all states i, j we have

$$p_{ij}(t) \to \lambda_j \text{ as } t \to \infty.$$

- Let $(X_t)_{t \geq 0}$ be Markov(δ_i, Q).

  Fix $h > 0$ and consider the *h*-**skeleton** $Z_n = X_{nh}$.

  By a previous theorem,

  $$\mathbb{P}(Z_{n+1} = i_{n+1} | Z_0 = i_0, \ldots, Z_n = i_n) = p_{i_n i_{n+1}}(h).$$

  So $(Z_n)_{n \geq 0}$ is discrete-time Markov$(\delta_i, P(h))$.

Convergence to Equilibrium (Cont'd)

- By a previous theorem, irreducibility implies $p_{ij}(h) > 0$ for all i, j.
  So $P(h)$ is irreducible and aperiodic
  By a previous theorem, λ is invariant for $P(h)$.
  So, by discrete-time convergence to equilibrium, for all i, j,
  $$p_{ij}(nh) \to \lambda_j \text{ as } n \to \infty.$$
  Thus we have a lattice of points along which the desired limit holds.

Convergence to Equilibrium (Cont'd)

- We fill in the gaps using uniform continuity. Fix a state i.

  Given $\varepsilon > 0$, we can find $h > 0$, such that

  $$1 - e^{-q_i s} \leq \frac{\varepsilon}{2}, \quad \text{for } 0 \leq s \leq h.$$

  Then find N, such that

  $$|p_{ij}(nh) - \lambda_j| \leq \frac{\varepsilon}{2}, \quad \text{for } n \geq N.$$

  For $t \geq Nh$, we have $nh \leq t < (n+1)h$, for some $n \geq N$.

  Moreover, by the preceding lemma,

  $$|p_{ij}(t) - \lambda_j| \leq |p_{ij}(t) - p_{ij}(nh)| + |p_{ij}(nh) - \lambda_j| \leq \varepsilon.$$

  Hence, $p_{ij}(t) \to \lambda_j$ as $n \to \infty$.

Limiting Behavior for Irreducible Chains

- The complete description of limiting behavior for irreducible chains in continuous time is provided by the following result.

Theorem

Let Q be an irreducible Q-matrix and let ν be any distribution. Suppose that $(X_t)_{t\geq 0}$ is Markov(ν, Q). Then

$$\mathbb{P}(X_t = j) \to \frac{1}{q_j m_j}, \quad \text{as } t \to \infty, \text{ for all } j \in I,$$

where m_j is the expected return time to state j.

- This follows from a previous theorem by the same argument we used in the preceding result.

Example

- Consider the Markov chain with state space $\{1, 2, 3\}$ and Q-matrix

$$\begin{pmatrix} -2 & 1 & 1 \\ 4 & -4 & 0 \\ 2 & 1 & -3 \end{pmatrix}.$$

We find an invariant distribution λ.

We have

$$(\lambda_1 \ \lambda_2 \ \lambda_3) \begin{pmatrix} -2 & 1 & 1 \\ 4 & -4 & 0 \\ 2 & 1 & -3 \end{pmatrix} = 0 \Rightarrow \left\{ \begin{array}{rcl} -2\lambda_1 + 4\lambda_2 + 2\lambda_3 & = & 0 \\ \lambda_1 - 4\lambda_2 + \lambda_3 & = & 0 \\ \lambda_1 - 3\lambda_3 & = & 0 \end{array} \right\}$$

$$\Rightarrow \left\{ \begin{array}{rcl} \lambda_1 & = & 3\lambda_3 \\ \lambda_2 & = & \lambda_3 \\ \lambda_1 + \lambda_2 + \lambda_3 & = & 1 \end{array} \right\} \Rightarrow \left\{ \begin{array}{rcl} \lambda_1 & = & \frac{3}{5} \\ \lambda_2 & = & \frac{1}{5} \\ \lambda_3 & = & \frac{1}{5} \end{array} \right\}.$$

Example (Cont'd)

- We discover $p_{11}(t)$.

  The matrix has characteristic equation

$$x(x+4)(x+5)=0.$$

  Hence, its eigenvalues are $x=0$, $x=-4$ and $x=-5$.

  It follows that

$$p_{11}(t)=a+be^{-4t}+ce^{-5t}.$$

  Moreover, we have

$$\left\{\begin{array}{rcl} p_{11}(0) & = & 1 \\ p'_{11}(0) & = & q_{11} \\ p''_{11}(0) & = & q^{(2)}_{11} \end{array}\right\} \Rightarrow \left\{\begin{array}{rcl} a+b+c & = & 1 \\ -4b-5c & = & -2 \\ 16b+25c & = & 10 \end{array}\right\} \Rightarrow \left\{\begin{array}{rcl} a & = & \frac{3}{5} \\ b & = & 0 \\ c & = & \frac{2}{5} \end{array}\right.$$

  So $p_{11}(t)=\frac{3}{5}+\frac{2}{5}e^{-5t} \stackrel{t\to\infty}{\longrightarrow} \frac{3}{5}=\lambda_1$.

Subsection 7

Time Reversal

Introducing Time Reversal

- In time reversal right-continuous processes become left-continuous.
- We can redefine the time-reversed process to equal its right limit at the jump times, thus obtaining again a right-continuous process.
- We suppose implicitly that this is done, and ignore this problem.

Time Reversal

Theorem

Let Q be irreducible and non-explosive and suppose that Q has an invariant distribution λ. Let $T \in (0, \infty)$ be given and let $(X_t)_{0 \leq t \leq T}$ be Markov(λ, Q). Set $\widehat{X}_t = X_{T-t}$. Then the process $(\widehat{X}_t)_{0 \leq t \leq T}$ is Markov$(\lambda, \widehat{Q})$, where $\widehat{Q} = (\widehat{q}_{ij} : i, j \in I)$ is given by $\lambda_j \widehat{q}_{ji} = \lambda_i q_{ij}$. Moreover, $\widehat{Q}$ is also irreducible and non-explosive with invariant distribution λ.

- By a previous theorem, the semigroup $(P(t) : t \geq 0)$ of Q is the minimal non-negative solution of the forward equation

  $$P'(t) = P(t)Q, \quad P(0) = I.$$

  Also, for all $t > 0$, $P(t)$ is an irreducible stochastic matrix with invariant distribution λ. Define $\widehat{P}(t)$ by

  $$\lambda_j \widehat{p}_{ji}(t) = \lambda_i p_{ij}(t).$$

Time Reversal (Cont'd)

- We defined $\widehat{P}(t)$ by $\lambda_j \widehat{p}_{ji}(t) = \lambda_i p_{ij}(t)$.

  Then $\widehat{P}(t)$ is an irreducible stochastic matrix with invariant distribution λ.

  We can rewrite the forward equation transposed as

  $$\widehat{P}'(t) = \widehat{Q}\widehat{P}(t).$$

  This is the backward equation for $\widehat{Q}$, which is itself a Q-matrix.

  Furthermore, $\widehat{P}(t)$ is its minimal non-negative solution.

  Hence $\widehat{Q}$ is irreducible and non-explosive and has invariant distribution λ.

Time Reversal (Cont'd)

- Finally, consider:
  - $0 = t_0 < \cdots < t_n = T$;
  - $s_k = t_k - t_{k-1}$.

  We have, by a previous theorem,

$$\begin{aligned} \mathbb{P}(\widehat{X}_{t_0} = i_0, \ldots, \widehat{X}_{t_n} = i_n) &= \mathbb{P}(X_{T-t_0} = i_0, \ldots, X_{T-t_n} = i_n) \\ &= \lambda_{i_n} p_{i_n i_{n-1}}(s_n) \cdots p_{i_1 i_0}(s_1) \\ &= \lambda_{i_0} \widehat{p}_{i_0 i_1}(s_1) \cdots \widehat{p}_{i_{n-1} i_n}(s_n). \end{aligned}$$

  So, again by a previous theorem, $(\widehat{X}_t)_{0 \leq t \leq T}$ is Markov$(\lambda, \widehat{Q})$.
- The chain $(\widehat{X}_t)_{0 \leq t \leq T}$ is called the **time-reversal** of $(X_t)_{0 \leq t \leq T}$.

Q-Matrix and Measure in Detailed Balance

- A Q-matrix Q and a measure λ are said to be in **detailed balance** if

$$\lambda_i q_{ij} = \lambda_j q_{ji}, \quad \text{for all } i, j.$$

Lemma

If Q and λ are in detailed balance then λ is invariant for Q.

- We have

$$(\lambda Q)_i = \sum_{j \in I} \lambda_j q_{ji} = \sum_{j \in I} \lambda_i q_{ij} = 0.$$

Reversibility

- Let $(X_t)_{t \geq 0}$ be Markov(λ, Q), with Q irreducible and non-explosive.

  We say that $(X_t)_{t \geq 0}$ is **reversible** if, for all $T > 0$, $(X_{T-t})_{0 \leq t \leq T}$ is also Markov(λ, Q).

Theorem

Let Q be an irreducible and non-explosive Q-matrix and let λ be a distribution. Suppose that $(X_t)_{t \geq 0}$ is Markov(λ, Q). Then the following are equivalent:

(a) $(X_t)_{t \geq 0}$ is reversible;

(b) Q and λ are in detailed balance.

- Both Conditions (a) and (b) imply that λ is invariant for Q.

  Then both Conditions (a) and (b) are equivalent to the statement that $\widehat{Q} = Q$ in the preceding theorem.

Subsection 8

Ergodic Theorem

Ergodic Theorem

Theorem (Ergodic Theorem)

Let Q be irreducible and let ν be any distribution. If $(X_t)_{t\geq 0}$ is Markov(ν, Q), then

$$\mathbb{P}\left(\frac{1}{t}\int_0^t 1_{\{X_s=i\}}ds \to \frac{1}{m_i q_i} \text{ as } t \to \infty\right) = 1,$$

where $m_i = \mathbb{E}_i(T_i)$ is the expected return time to state i. Moreover, in the positive recurrent case, for any bounded function $f : I \to \mathbb{R}$ we have

$$\mathbb{P}\left(\frac{1}{t}\int_0^t f(X_s)ds \to \overline{f} \text{ as } t \to \infty\right) = 1,$$

where $\overline{f} = \sum_{i\in I} \lambda_i f_i$ and where $(\lambda_i : i \in I)$ is the unique invariant distribution.

Ergodic Theorem (Cont'd)

- Suppose Q is transient.

  Then the total time spent in any state i is finite.

  So
  $$\frac{1}{t}\int_0^t 1_{\{X_s=i\}}ds \leq \frac{1}{t}\int_0^\infty 1_{\{X_s=i\}}ds \to 0 = \frac{1}{m_i}.$$

  Suppose then that Q is recurrent and fix a state i.

  Then $(X_t)_{t\geq 0}$ hits i with probability 1.

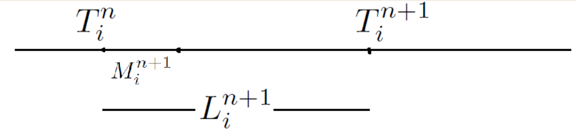
  The long run proportion of time in i equals the long run proportion of time in i after first hitting i.

  By the Strong Markov Property (of the jump chain), it suffices to consider the case $\nu = \delta_i$.

Ergodic Theorem (Cont'd)

- Denote by:
  - M_i^n the length of the n-th visit to i;
  - T_i^n the time of the n-th return to i;
  - L_i^n the length of the n-th excursion to i.

  Thus for $n = 0, 1, 2, \ldots$, setting $T_i^0 = 0$, we have:



$$\begin{aligned} M_i^{n+1} &= \inf\{t > T_i^n : X_t \neq i\} - T_i^n; \\ T_i^{n+1} &= \inf\{t > T_i^n + M_i^{n+1} : X_t = i\}; \\ L_i^{n+1} &= T_i^{n+1} - T_i^n. \end{aligned}$$

Ergodic Theorem (Cont'd)

- By the Strong Markov Property (of the jump chain) at the stopping times T_i^n, for $n \geq 0$, we find that:
  - $L_i^1, L_i^2, \ldots$ are independent and identically distributed with mean m_i;
  - $M_i^1, M_i^2, \ldots$ are independent and identically distributed with mean $\frac{1}{q_i}$.

  Hence, by the Strong Law of Large Numbers, as $n \to \infty$,

$$\frac{L_i^1 + \cdots + L_i^n}{n} \to m_i \quad \text{and} \quad \frac{M_i^1 + \cdots + M_i^n}{n} \to \frac{1}{q_i}.$$

  Therefore,

$$\frac{M_i^1 + \cdots + M_i^n}{L_i^1 + \cdots + L_i^n} \to \frac{1}{m_i q_i} \text{ as } n \to \infty, \text{ with probability } 1.$$

  In particular, we note that $\frac{T_i^n}{T_i^{n+1}} \to 1$ as $n \to \infty$, with probability 1.

Ergodic Theorem (Cont'd)

- Now, for $T_i^n \leq t < T_i^{n+1}$ we have

$$\frac{T_i^n}{T_i^{n+1}} \frac{M_i^1 + \cdots + M_i^n}{L_i^1 + \cdots + L_i^n} \leq \frac{1}{t} \int_0^t 1_{\{X_s = i\}} ds \leq \frac{T_i^{n+1}}{T_i^n} \frac{M_i^1 + \cdots + M_i^{n+1}}{L_i^1 + \cdots + L_i^{n+1}}.$$

Letting $t \to \infty$, we obtain that, with probability 1,

$$\frac{1}{t} \int_0^t 1_{\{X_s = i\}} ds \to \frac{1}{m_i q_i}.$$

In the positive recurrent case, for $\lambda_i = \frac{1}{m_i q_i}$, we can write

$$\frac{1}{t} \int_0^t f(X_s) ds - \overline{f} = \sum_{i \in I} f_i \left(\frac{1}{t} \int_0^t 1_{\{X_s = i\}} ds - \lambda_i \right).$$

By the same argument used in the discrete case, $\frac{1}{t} \int_0^t f(X_s) ds \to \overline{f}$ as $t \to \infty$, with probability 1.