UPPSALA UNIVERSITY
Department of Mathematics

Markov Processes, 1MS012
Spring semester 2024

Lecture 10 Markov Processes, 1MS012

1 Markov processes (continuous time)

Definition: A random process $(X_t)_{t\geq 0}$ with countable state space (w.l.o.g. a set of integers) is a (discrete) **Markov process** if, for any n and time-points $t_1 < ... < t_{n+1}$ and states $j_1, ..., j_{n+1}$

$$\begin{aligned}&P(X_{t_{n+1}} = j_{n+1} \mid X_{t_n} = j_n, X_{t_{n-1}} = j_{n-1}, ..., X_{t_1} = j_1)\\ &\qquad = P(X_{t_{n+1}} = j_{n+1} \mid X_{t_n} = j_n).\end{aligned}$$

Thus the "future" is conditionally independent of the "past" given the present.

Time-homogeneous Markov process:

$$p_{ij}(t) = P(X_t = j \mid X_0 = i) = P(X_{t+s} = j \mid X_s = i),$$

for all $t, s \geq 0$.

Observe: $p_{ii}(0) = 1$ and $p_{ij}(0) = 0$ for all $i \neq j$, so if

$$\mathbf{P}(t) = \begin{pmatrix} \ddots & \vdots & . \\ \dots & p_{ij}(t) & \dots \\ . & \vdots & \ddots \end{pmatrix},$$

then $\mathbf{P}(0) = I$ (the identity matrix). The matrices $\mathbf{P}(t)$ are called the **matrices of transition probabilities**. These matrices are stochastic matrices i.e. they have non-negative entries and row-sums $= 1$.

Example: Poisson process $(N_t)_{t\geq 0}$, with intensity parameter $\lambda > 0$: Stochastic process on $S = \{0, 1, 2, \dots\}$ starting at $N_0 = 0$ with

$$P(N_{t_{n+1}} = j_{n+1} \mid N_{t_n} = j_n, N_{t_{n-1}} = j_{n-1}, ..., N_{t_1} = j_1)$$

$$\underbrace{:=}_{\text{independent increments}} \quad P(N_{t_{n+1}} - N_{t_n} = j_{n+1} - j_n)$$

$$\underbrace{:=}_{N_t - N_s \sim Po(\lambda(t-s)),\, 0 \le s < t} \quad e^{-\lambda(t_{n+1}-t_n)} \frac{(\lambda(t_{n+1}-t_n))^{j_{n+1}-j_n}}{(j_{n+1}-j_n)!},$$

if $j_k \le j_{k+1}$, for all $k = 1, 2, ..., n$.
($P(N_{t_{n+1}} = j_{n+1} \mid N_{t_n} = j_n, N_{t_{n-1}} = j_{n-1}, ..., N_{t_1} = j_1) = 0$, otherwise.)
(With $N_t - N_s \sim Po(\lambda(t-s)$ we thus mean that the increment $N_t - N_s$ has the Poisson distribution with intensity parameter $\lambda(t-s)$.)

Thus (N_t) is a Markov process with

$$p_{ij}(t) = \begin{cases} e^{-\lambda t} \frac{(\lambda t)^{j-i}}{(j-i)!}, & j \ge i \\ 0, & j < i \end{cases}$$

Chapman-Kolmogorov equations:

$$\underbrace{P(X_{t+s} = k \mid X_0 = i)}_{p_{ik}(s+t)} = \sum_{j \in S} \underbrace{P(X_s = j \mid X_0 = i)}_{p_{ij}(s)} \underbrace{P(X_{t+s} = k \mid X_s = j, X_0 = i)}_{p_{jk}(t)}$$

If we express the Chapman-Kolmogorov equations in matrix form we get:

$$\underbrace{\begin{pmatrix} \ddots & \vdots & . \\ \dots & p_{ik}(s+t) & \dots \\ . & \vdots & \ddots \end{pmatrix}}_{\mathbf{P}(s+t)} = \underbrace{\begin{pmatrix} \ddots & \vdots & . \\ p_{i1}(s) & p_{i2}(s)... & p_{i.}(s) \\ . & \vdots & \ddots \end{pmatrix}}_{\mathbf{P}(s)} \underbrace{\begin{pmatrix} \ddots & p_{1k}(t) & . \\ \dots & p_{2k}(t) & \dots \\ . & p_{.k}(t) & \ddots \end{pmatrix}}_{\mathbf{P}(t)}$$

Note:
For Markov chains (with discrete time): If we know the transition matrix $\mathbf{P}$, then we can calculate $\mathbf{P}^n$, for all n.

For Markov processes (with continuous time): If we know the matrix $\mathbf{P}(t)$, for some t, then we cannot calculate $\mathbf{P}(t)$ for all t. (only for $\mathbf{P}(t)$, $\mathbf{P}(2t)$,$\mathbf{P}(3t)$, ...)

If we let $t \to 0$ and note that

$$\frac{p_{ij}(t) - p_{ij}(0)}{t} = \begin{cases} \frac{p_{ij}(t)}{t} & i \ne j \\ \frac{p_{ij}(t)-1}{t} & i = j, \end{cases}$$

we get

$$p_{ij}(t) \approx \begin{cases} tp'_{ij}(0) & i \neq j \\ 1 + tp'_{ij}(0) & i = j, \end{cases}$$

for small t. (provided these derivatives exist)

It therefore seems plausible that the matrix of derivatives $(p'_{ij}(t))$ at time $t = 0$ for a Markov process would play a similar role as the transition matrix $\mathbf{P}$ for a Markov chain in characterizing the process.

Remark: It can be proved that the derivatives $(p'_{ij}(t))$ exist for Markov processes with right-continuous trajectories satisfying the property that only a finite number of jumps can be made in each finite time-interval.

Definition: Let (X_t) be a (discrete, time-homogeneous) Markov process. Suppose there exists $q_{ij} \geq 0$ for $j \neq i$ and $q_{ii} \leq 0$ such that

$$P(X_t = j \mid X_0 = i) = q_{ij}t + o(t),$$

and

$$P(X_t \neq i \mid X_0 = i) = q_i t + o(t),$$

where $q_i = \sum_{j \neq i} q_{ij} =: -q_{ii}$, and $o(t)$, small ordo of t, is a function satisfying $\lim_{t \to 0} o(t)/t = 0$. The numbers q_{ij} are called the **transition intensities** from state i to state j and the array

$$\mathbf{Q} = \begin{pmatrix} \ddots & \vdots & . \\ \cdots & q_{ij} & \cdots \\ . & \vdots & \ddots \end{pmatrix}_{i,j \in S}$$

is called the **intensity matrix, or the infinitesimal generator** of the Markov process.

Remark: Note that

$$\sum_j q_{ij} = \sum_{j \neq i} q_{ij} + q_{ii} = -q_{ii} + q_{ii} = 0,$$

for any state i, so row sums in $\mathbf{Q}$ are zero by definition.

Remark: Note that

$$P(X_t = i \mid X_0 = i) = 1 - q_i t + o(t),$$

for any state i, since $P(X_t = i \mid X_0 = i) + P(X_t \neq i \mid X_0 = i) = 1$.

Remark: Note that if S is finite, then $\mathbf{P} := \frac{1}{\max q_i}\mathbf{Q} + I$ is a stochastic matrix. This makes it possible to translate some properties for discrete time Markov chains to the continuous time case, see e.g. Lawler, exercise 3.4.

Theorem: Let $(X_t)_{t \geq 0}$ be a (discrete, time-homogeneous) Markov process with intensity matrix $\mathbf{Q}$. If T_1 denotes the time for the first jump of the process, then:

(i) $P(T_1 \leq t \mid X_0 = i) = 1 - e^{-q_i t}$ i.e. if the process starts in state i then T_1 is exponentially distributed with expectation $1/q_i$.

(ii) $P(X_{T_1} = j \mid X_0 = i) = \frac{q_{ij}}{q_i}$, for $j \neq i$ i.e. if the process starts in state i then the probability that the next state is j is q_{ij}/q_i.

(iii) T_1 and X_{T_1} are conditionally independent given that $X_0 = i$.

Sketch of proof: By the Markov property and homogeneity of the chain it follows that T_1 has the "lack of memory" property (see the appendix below), and therefore T_1 must be exponentially distributed. The remaining part of the proof follows from properties of exponentially distributed random variables stated in the appendix.

Construction of a Markov process, $(X_t)_{t \geq 0}$, from an intensity matrix $\mathbf{Q}$:

1. Choose a starting point Z_0.

2. If $Z_0 = j$ then the process stays in state j a random time interval $U_0 \sim \mathrm{Exp}(q_j)$. ($U_0$ is called the first holding time.) (We will typically only consider cases when $0 < q_j < \infty$. If $q_j = 0$ then the process will never leave state j, and if $q_j = \infty$ then the process leaves j immediately.)

3. Next choose a new state, Z_1, according to the matrix $R = (r_{ij})$, with $r_{ij} = q_{ij}/q_i$, $i \neq j$, and $r_{ii} = 0$.

4. Repeat inductively steps 2 and 3. If $Z_j = k$ let $U_j \sim \mathrm{Exp}(q_k)$ (be holding times independent of $U_0, \ldots, U_{j-1}$)

If $J_n = \sum_{i=0}^{n-1} U_i$ and $J_n \to \infty$ then we can define X_t for any $t \geq 0$ by $X_t = Z_m$ if $J_m \leq t < J_{m+1}$.

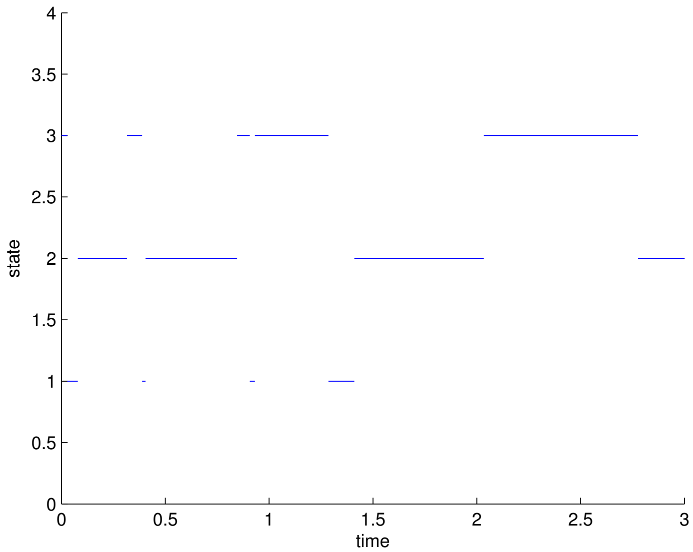

Figure 1: Trajectory of a Markov process with state space $S = (1, 2, 3)$

Theorem: The process (X_t) defined as above is a (homogeneous) Markov process.

Remark: The Markov chain (Z_n) with transition matrix R is called the jump-chain corresponding to (X_t). Note that the definition of R above presumes that $q_i > 0$. To avoid ambiguity we define $r_{ij} = 0$, for $i \neq j$ and $r_{ii} = 1$ in case $q_i = 0$.

Example: Construction of the Poisson process:
$q_j = \lambda$ for all j (for some parameter $\lambda > 0$)
$r_{i,i+1} = 1$ for any $i \geq 0$.

2 Appendix: Some properties of exponentially distributed random variables

The exponential distribution may be viewed as a continuous counterpart to the geometric distribution.

We say that a random variable X is exponentially distributed with parameter $\lambda > 0$ (notation $X \sim \mathrm{Exp}(\lambda)$) if its density function is given by $f_X(x) = \lambda e^{-\lambda x}$, $x \geq 0$, $f_X(x) = 0$, otherwise.

Distribution function: $F_X(x) = 1 - e^{-\lambda x}$, $x \geq 0$.

If $X \sim \mathrm{Exp}(\lambda)$ then $E(X) = \int_0^\infty P(X > x)dx = \int_0^\infty e^{-\lambda x} = \frac{1}{\lambda}$.

Exponentially distributed random variables have many interesting properties:

Theorem: (Lack of memory property.) Let $X \sim \mathrm{Exp}(\lambda)$. Then

$$P(X > x + y \mid X > y) = P(X > x), \text{ for all } x, y > 0.$$

The exponential distribution is the only non-negative continuous distribution having the lack of memory property.

If $X_1 \sim \mathrm{Exp}(\lambda_1)$ and $X_2 \sim \mathrm{Exp}(\lambda_2)$ are independent then

$$\begin{aligned} P(X_1 < X_2) &= \int_{\mathbb{R}} P(X_1 < X_2 \mid X_1 = s) f_{X_1}(s)dt \\ &= \int_0^\infty \underbrace{P(X_2 > s)}_{e^{-\lambda_2 s}} \underbrace{f_{X_1}(s)}_{\lambda_1 e^{-\lambda_1 s}} ds \\ &= \lambda_1 \int_0^\infty e^{-(\lambda_1+\lambda_2)s} ds = \frac{\lambda_1}{\lambda_1 + \lambda_2}, \end{aligned} \tag{1}$$

and if $T = \min(X_1, X_2)$ then $T \sim \mathrm{Exp}(\lambda_1 + \lambda_2)$ since

$$\begin{aligned} P(T > t) &= P(X_1 > t, X_2 > t) \\ &\underbrace{=}_{by\ indep.} P(X_1 > t)P(X_2 > t) = e^{-(\lambda_1+\lambda_2)t}. \end{aligned}$$

The events $\{\min(X_1, X_2) > t\}$ and $\{X_1 < X_2\}$ are also independent, since a slight generalization of (1) gives $P(t < X_1 < X_2) = \lambda_1 e^{-(\lambda_1+\lambda_2)t}/(\lambda_1 + \lambda_2)$.

More generally:

Theorem: Let $X_1 \sim \mathrm{Exp}(\lambda_1), X_2 \sim \mathrm{Exp}(\lambda_2), \ldots$ be independent exponentially distributed random variables with $0 < \sum_i \lambda_i < \infty$. Let $X = \inf_k X_k$.

Then this infimum is attained at a unique random value K of k (with probability 1). Moreover X and K are independent and

$$X \sim \mathrm{Exp}(\sum_i \lambda_i),$$

and

$$P(K = k) = \frac{\lambda_k}{\sum_i \lambda_i}.$$

3 Matlab

3.1 Program for simulating a Markov processes with finite state space and given generator Q (Figure 1)

```
Q=[-12 10 2; 1 -2 1; 2 2 -4];                    % Choose intensity matrix
Tmax=3;                       % The process will be simulated up to time Tmax
z(1) = 3;          % Choose initial state in 1,2,...,m (in case Q has size m × m)
                                  % (We store Z_n in z(n+1) and J_n in J(n+1).)


for i=1:length(Q)                                   % corresponding jump-chain
  for j=1:length(Q)
    if i==j
      R(i,j)=0;
    else
      R(i,j)=-Q(i,j)/Q(i,i);
    end
  end
end
J(1) = 0;                                                   % start time at t=0
i = 1;
while J(i) < Tmax,
  J(i+1) = J(i) + log(rand)/Q(z(i),z(i));                   % Next jump-time
  u = rand;                                 % Next state is chosen according to R
  j = 1;
  s = R(z(i),1);
  while ((u > s) & (j < length(Q) )),
    j=j+1;
    s=s+R(z(i),j);
  end
  z(i+1) = j;
  i=i+1;
end
cla;
hold on                                 % A trajectory plot consists of many...
for i=1:(length(J)-1),
  plot([J(i) J(i+1)], [z(i) z(i)]);                         % ... line-segments
end
axis([0 Tmax 0 (length(Q)+1)]);
xlabel('time');
ylabel('state');
```

4 Suggested exercises

Basic exercises:
18, 20

Exercises Lawler:
3.1, 3.3

Lecture 11 Markov Processes, 1MS012

5 Forward & Backward equations

Recall:
A (well behaved) Markov process gives rise to an intensity matrix or generator $\mathbf{Q}$. Conversely given a matrix $\mathbf{Q} = (q_{jk})$ where q_{jk}, for all $j, k \in S$ satisfy
(a) $q_{jk} \geq 0,\, j \neq k$.
(b) $-\infty < q_{jj} \leq 0$.
(c) $\sum_{k \in S} q_{jk} = 0,\ j \in S$,
we can construct a Markov process (at least if S is finite), with holding times, exponentially distributed with parameters $-q_{jj} = q_j < \infty$, and jump chain also specified by the entries in $\mathbf{Q}$.

The matrix $\mathbf{Q}$ gives all information about the Markov process in the sense that for small time steps h

$$\begin{aligned} p_{ij}(h) &= P(X_{t+h} = j \mid X_t = i) = q_{ij}h + o(h),\ i \neq j \\ p_{jj}(h) &= P(X_{t+h} = j \mid X_t = j) = 1 + q_{jj}h + o(h) = 1 - q_j h + o(h) \end{aligned}$$

Markovprocesses in discrete time:
$P(X_n = k)$ is "easy" to find if we know the distribution of X_0 and the transition matrix $\mathbf{P}$.

Markovprocesses continuous time:
Not obvious how to use $\mathbf{Q}$ to find $P(X_t = k)$. We can derive equations for $p_{ij}(t) = P(X_t = j | X_0 = i)$ that hopefully can be solved:

Forward equations: (S finite)
We study the derivatives of $p_{ij}(t)$.

$$\begin{aligned} \lim_{h \to 0} \frac{p_{ij}(t+h) - p_{ij}(t)}{h} \quad & \underbrace{=}_{CK-equations} \quad \lim_{h \to 0} \frac{\sum_{k \in S} p_{ik}(t) p_{kj}(h) - p_{ij}(t)}{h} \\ & \underbrace{=}_{S\ finite} \quad \sum_{k \in S \setminus \{j\}} p_{ik}(t) \underbrace{\lim_{h \to 0} \frac{p_{kj}(h)}{h}}_{q_{kj}} + p_{ij}(t) \underbrace{\lim_{h \to 0} \frac{p_{jj}(h) - 1}{h}}_{q_{jj}} \\ & = \quad \sum_{k \in S} p_{ik}(t) q_{kj}. \end{aligned}$$

The equations

$$p'_{ij}(t) = \sum_{k \in S} p_{ik}(t) q_{kj}$$

are called **Kolmogorov's forward equations**.

Forward equations expressed in matrix form: $\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q}$

$$\underbrace{\begin{pmatrix} \ddots & \vdots & . \\ \dots & p'_{ij}(t) & \dots \\ . & \vdots & \ddots \end{pmatrix}}_{\mathbf{P}'(t)} = \underbrace{\begin{pmatrix} \ddots & \vdots & . \\ p_{i1}(t) & p_{i2}(t)\dots & p_{i|S|}(t) \\ . & \vdots & \ddots \end{pmatrix}}_{\mathbf{P}(t)} \underbrace{\begin{pmatrix} \ddots & q_{1j} & . \\ \dots & q_{2j} & \dots \\ . & q_{|S|j} & \ddots \end{pmatrix}}_{\mathbf{Q}}$$

(Actually we just proved that the right-derivative exists but a similar arguments shows that also the left derivative exists.)

Example: A machine works an exponential time with mean $\frac{1}{\mu}$. It takes an exponential time with mean $\frac{1}{\lambda}$ to repair it. If the machine works at time 0, what is the probability that it will not work at time t?
We model the state of the machine with a Markov process. Let (X_t) be a Markov process with state space $S = (0, 1)$ where X_t represents the state of the process at time t: Let
$X_t = 1$ mean that the machine works at time t, and
$X_t = 0$ mean that it doesn't work at time t.
When X_t enters state 1 it stays there an $\mathrm{Exp}(\mu)$-distributed time and then goes to 0 and stays there an $\mathrm{Exp}(\lambda)$-distributed time and then goes to state 1 and so on.
We have $q_0 = \lambda = -q_{00}$, and $q_1 = \mu = -q_{11}$. Thus we get the intensity matrix

$$\mathbf{Q} = \begin{pmatrix} -\lambda & \lambda \\ \mu & -\mu \end{pmatrix},$$

(since row sums of Q are zero).
We want to find $p_{10}(t) = P(X_t = 0 | X_0 = 1)$.
The forward equations gives:

$$p'_{10}(t) = -\lambda p_{10}(t) + \mu p_{11}(t) = -\lambda p_{10}(t) + \mu(1 - p_{10}(t))$$
$$= \mu - (\lambda + \mu) p_{10}(t),$$

with the boundary condition $p_{10}(0) = 0$.
The equation

$$p'_{10}(t) + (\lambda + \mu) p_{10}(t) = \mu,$$

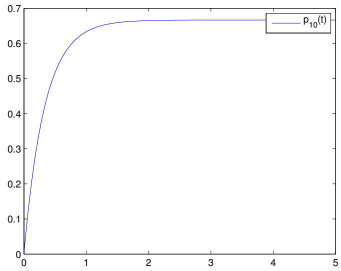

Figure 2: (X_t) with parameters $\lambda = 1$, and $\mu = 2$.
$\lim_{t\to\infty} p_{10}(t) = \mu/(\lambda+\mu) = 2/3$.

can be solved by the standard trick of equivalently writing it as

$$\underbrace{p_{10}'(t)e^{(\lambda+\mu)t} + (\lambda+\mu)e^{(\lambda+\mu)t}p_{10}(t)}_{\frac{d}{dt}(p_{10}(t)e^{(\lambda+\mu)t})} = \mu e^{(\lambda+\mu)t}.$$

By integrating both sides we get

$$p_{10}(t)e^{(\lambda+\mu)t} = \frac{\mu}{\lambda+\mu}e^{(\lambda+\mu)t} + c,$$

where c is a constant. The boundary condition $p_{10}(0) = 0 \Rightarrow c = \frac{-\mu}{\lambda+\mu}$.
Thus

$$p_{10}(t) = \frac{\mu}{\lambda+\mu}(1 - e^{-(\lambda+\mu)t}).$$

5.1 Kolmogorov's backward equations

If S is infinite then the Kolmogorov forward equations may fail to hold.
To avoid such problems we derive another set of equations.
By conditioning on the state of the process at time h we get

$$\begin{aligned}
\lim_{h\to 0}\frac{p_{ij}(t+h)-p_{ij}(t)}{h} &\underbrace{=}_{CK-equations} \lim_{h\to 0}\frac{\sum_{k\in S}p_{ik}(h)p_{kj}(t)-p_{ij}(t)}{h} \\
&\underbrace{=}_{Can\ be\ justified} \sum_{k\in S\setminus\{i\}}\underbrace{\lim_{h\to 0}\frac{p_{ik}(h)}{h}}_{q_{ik}}p_{kj}(t) + p_{ij}(t)\underbrace{\lim_{h\to 0}\frac{p_{ii}(h)-1}{h}}_{q_{ii}} \\
&= \sum_{k\in S} q_{ik}p_{kj}(t).
\end{aligned}$$

We have thus derived **Kolmogorov's backward equations**

$$p_{jk}'(t) = \sum_{r\in S} q_{jr}p_{rk}(t)$$

expressed in matrix form as:

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

Thus if the forward equation also holds then

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

Theorem: If $(X_t)_{t\geq 0}$ is a Markov process with finite state space,then

$$\mathbf{P}(t) = e^{t\mathbf{Q}} := \sum_{k=0}^{\infty} \mathbf{Q}^k t^k/k! = \mathbf{I} + t\mathbf{Q} + \frac{t^2\mathbf{Q}^2}{2!} + \frac{t^3\mathbf{Q}^3}{3!} + ...,$$

is the unique solution of the forward and backward equations, with initial condition $\mathbf{P}(0) = \mathbf{I}$ where $\mathbf{I}$ is the identity matrix.

This explicitly explains why $\mathbf{Q}$ is called the "generator" of the process.

If S is finite and $\mathbf{Q}$ can be diagonalized, with $\mathcal{P}^{-1}\mathbf{Q}\mathcal{P} = \mathbf{D}$, for some invertible matrix $\mathcal{P}$, where $\mathbf{D}$ is a diagonal matrix with diagonal entries d_i, then

$$\mathbf{P}(t) = \mathcal{P}e^{t\mathbf{D}}\mathcal{P}^{-1},$$

where $e^{t\mathbf{D}}$ is the diagonal matrix with diagonal entries $e^{d_i t}$.

If S infinite then it is possible to prove that if there exists a solution, $\mathbf{P}(t)$ to the forward and backward equations with $\sum_{j\in S} p_{ij}(t) = 1$, for all t and $i \in S$, then this is the unique solution.
In general it is hard to explicitly solve the forward and backward equations if S has many states and $\mathbf{Q}$ has no particular structure. Typically the forward and backward equations have to be solved by numerical methods.

6 Matlab

6.1 Program for Figure 2

```
lambda=1;      % parameter values
mu=2;          %
x0=0;          % starting point
[t,y]=ode45(@hogerled,[0,5],x0,[],lambda,mu);
               % Numerically solves ODEs of the form y'(t)=f(t,y)
               % with given starting values, on a specified interval
               % with specified parameters.
               % The function f(t,y) is specified in the m-file
               % hogerled.m below.
plot(t,y)
legend('p_{10}(t)')


% In the file hogerled.m
function fty=hogerled(t,y,lambda,mu);
fty=mu-(lambda+mu)*y;
```

7 Suggested exercises

Basic exercises:
19

Extra problems:
c2

Exercises Lawler:
3.5

Lecture 12 Markov Processes, 1MS012

8 Class structure

Definition: State k is **accessible** from state j ($j \to k$) if $p_{jk}(t) > 0$ for some $t \geq 0$.
k and j are intercommunicating ($j \leftrightarrow k$) if $j \to k$ and $k \to j$

Theorem: Let (X_t) be a discrete Markov process with intensity matrix $\mathbf{Q}$. Let j and k be two (distinct) states. Then

$$j \to k$$

$$\Leftrightarrow$$

there exists an admissible path from j to k with respect to the matrix $\mathbf{Q}$ i.e. if either $q_{jk} > 0$ or there exists states $j_1, ..., j_n$ such that $q_{jj_1} > 0$, $q_{j_1 j_2} > 0$,..., $q_{j_n k} > 0$.

$$\Leftrightarrow$$

$$p_{jk}(t) > 0, \text{ for } \textit{all } t > 0,$$

The last equivalence can be seen since holding times of arbitrary small length have positive probability.

Conclusion:
Don't need to bother about aperiodicity for Markov processes in continuous time.

Closed sets and irreducible sets/processes are defined in analogy with the discrete time case.

The definition of recurrent and transient states has to be presented with care since all states are re-visited an infinite amount of time during the first holding time;

Definition: We say that j is recurrent if j is recurrent for the jump-chain, otherwise j is transient.

A state j is recurrent if, starting at j the process visits j at arbitrary large time-points (with probability 1). A state j is transient otherwise.

Let T_{jj} be the first time the process returns to j after the first jump. Thus if $X_0 = j$ then $T_{jj} = \inf(t > J_1 : X_t = j)$, where J_1 is the time for the first jump. If j is a recurrent state, and $ET_{jj} = \infty$, then j is said to be null-recurrent. Otherwise we call j non-null, (or positive) recurrent.

If $q_j = 0$ then we may define j to be non-null (positive) recurrent.

Theorem: j is recurrent iff $\int_0^\infty p_{jj}(t)dt = \infty$.

Similarly as for Markov chains in discrete time it can also be proved that
• if i and j are two states with $i \leftrightarrow j$ then i is positive recurrent iff j is positive recurrent.
• All states are positive recurrent for a finite irreducible Markov process.

8.1 Stationary distributions and the long run

Suppose $S = (1, 2, ..., n)$ is finite and

$$p_{jk}(t) \to \pi_k \qquad \text{as } t \to \infty, \tag{2}$$

for all j and k, i.e.

$$\mathbf{P}(t) = \begin{pmatrix} p_{11}(t) & \cdot & \cdot & p_{1n}(t) \\ p_{21}(t) & \cdot & \cdot & p_{2n}(t) \\ \cdot & \cdot & \cdot & \cdot \\ p_{n1}(t) & \cdot & \cdot & p_{nn}(t) \end{pmatrix} \to \begin{pmatrix} \pi_1 & \pi_2 & \cdot & \pi_n \\ \pi_1 & \pi_2 & \cdot & \pi_n \\ \pi_1 & \pi_2 & \cdot & \pi_n \\ \pi_1 & \pi_2 & \cdot & \pi_n \end{pmatrix}$$

Then, $\pi = (\pi_1, \pi_2, \dots, \pi_n)$, has the following properties:

1. If $\boldsymbol{\mu}(t) = (\mu_1(t), \dots, \mu_n(t)) := (P(X_t = 1), \dots, P(X_t = n))$ is the $1 \times n$ vector representing the probability distribution of X_t then

   $$\boldsymbol{\mu}(t) = \boldsymbol{\mu}(0)\mathbf{P}(t) \to \pi,$$

   as $t \to \infty$. This can be seen from (2) since

   $$\mu_k(t) = P(X_t = k) = \sum_{j \in S} P(X_t = k \mid X_0 = j)P(X_0 = j) = \sum_{j \in S} p_{jk}(t)\mu_j(0)$$

   $$\to \sum_{j \in S} \pi_k \mu_j(0) = \pi_k, \qquad \text{as } t \to \infty.$$

2.
$$\pi = \pi\mathbf{P}(s),$$
for all s, i.e. π is a stationary (or equilibrium) distribution.

This can be seen by letting $t \to \infty$ in the Chapman-Kolmogorov identity since
$$\underbrace{\boldsymbol{\mu}(0)\mathbf{P}(t+s)}_{\to\pi} = \underbrace{\boldsymbol{\mu}(0)\mathbf{P}(t)}_{\to\pi}\mathbf{P}(s)$$
(Thus if $X_0 \overset{d}{\sim} \pi$ then (X_t) forms a stationary process.)

3.
$$\pi\mathbf{Q} = \mathbf{0}$$
This can be seen since $\lim_{t\to\infty} p_{jk}(t) = \pi_k \Rightarrow \lim_{t\to\infty} p'_{jk}(t) = 0$ and since by the forward equations, $\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q}$,
$$\underbrace{\pi\mathbf{P}'(t)}_{\to\mathbf{0}} = \underbrace{\pi\mathbf{P}(t)}_{\pi}\mathbf{Q}.$$

4. The distribution
$$\nu = \frac{1}{\sum \pi_k q_k}(\pi_1 q_1, ..., \pi_n q_n),$$
is a stationary distribution for the jump chain, i.e. $\nu\mathbf{R} = \nu$, where
$$\mathbf{R} = \begin{pmatrix} r_{11} & \cdot & \cdot & r_{1n} \\ r_{21} & \cdot & \cdot & r_{2n} \\ \cdot & \cdot & \cdot & \cdot \\ r_{n1} & \cdot & \cdot & r_{nn} \end{pmatrix}$$
is defined by $r_{ij} = q_{ij}/q_i$ if $i \neq j$, and $r_{ii} = 0$ for any i. (Recall the implicit assumption that $q_i > 0$ for all i, so r_{ij} is well defined.)

Check:
Let $\nu_j = \frac{\pi_j q_j}{\sum \pi_k q_k}$. The distribution $\nu = (\nu_1, ..., \nu_n)$ is stationary, since
$$\nu_j = \frac{\pi_j q_j}{\sum \pi_k q_k} = \frac{-\pi_j q_{jj}}{\sum \pi_k q_k} = \frac{\sum_{i\neq j} \pi_i q_{ij}}{\sum \pi_k q_k} = \sum_{i\neq j} \frac{\pi_i q_i}{(\sum \pi_k q_k)} \frac{q_{ij}}{q_i} = \sum_i \nu_i r_{ij},$$
for each j, where in the third equality we used that $\pi\mathbf{Q} = \mathbf{0}$.

5. The stationary distribution π satisfies
$$\pi_k = \frac{1}{q_k E(T_{kk})},$$

i.e. π_k is the proportion of time spent in state k.
The proof of this uses the renewal theorem.
(Interpretation: On average it takes a time of length $E(T_{kk})$ between arrivals to the state k. During this time we spend on average a time of length $1/q_k$ in the state k and the rest of the time in other states.)

Theorem: If (X_t) is an irreducible (discrete, time-homogeneous) Markov process such that all states are positive recurrent then

$$p_{jk}(t) \to \pi_k = \frac{1}{q_k E(T_{kk})},$$

as $t \to \infty$ for all j and k, and $\pi = (\pi_j)_{j \in S}$, is uniquely stationary for the process, and the unique solution to the equation $\pi \mathbf{Q} = \mathbf{0}$, with $\sum_{i \in S} \pi_i = 1$, and $\pi_i \geq 0$, for any i.

How can we check if an irreducible Markov process is positive recurrent?

Theorem: Let (X_t) be an irreducible Markov process with intensity matrix $\mathbf{Q}$ such that there is a unique solution to the forward and backward equations. Then

$$(X_t) \text{ is positive recurrent}$$

$$\Leftrightarrow$$

there exists a solution to $\pi \mathbf{Q} = \mathbf{0}$, with $\sum_{i \in S} \pi_i = 1$, and $\pi_i \geq 0$, for any i.

Remark: If S is finite then $P(t) = e^{\mathbf{Q}t}$ is the unique solution to the forward and backward equations.

Example: Let $(X_t)_{t \geq 0}$ be a Markov process on $S = (0, 1, 2)$ with intensity matrix

$$\mathbf{Q} = \begin{pmatrix} q_{00} & q_{01} & q_{02} \\ q_{10} & q_{11} & q_{12} \\ q_{20} & q_{21} & q_{22} \end{pmatrix} = \begin{pmatrix} -1 & 1 & 0 \\ 2 & -3 & 1 \\ 0 & 2 & -2 \end{pmatrix}.$$

Suppose we wish to find the limit

$$\lim_{t \to \infty} P(X_t = 2 \mid X_0 = 0).$$

Since the Markovprocess (X_t) is irreducible with a finite state space it follows from the convergence theorem that

$$\lim_{t \to \infty} P(X_t = 2 \mid X_0 = 0) = \pi_2,$$

where $\pi = (\pi_0, \pi_1, \pi_2)$ is the unique probability vector solving the equation

$$\pi \mathbf{Q} = \mathbf{0}.$$

This system has solution $\pi = (4/7, 2/7, 1/7)$.
Thus

$$\lim_{t\to\infty} P(X_t = 2 \mid X_0 = 0) = 1/7.$$

Example: (continued)
Consider the same Markov process as in the previous example.
Suppose, starting at state 0, we wish to find the expected time it takes to reach state 2.
Let $T_{i2} = \inf(t > J_1 : X_t = 2 \mid X_0 = i)$ be the first passage time from state i to state 2, $i = 0, 1, 2$.
In order to find $E(T_{02})$ we will use the idea of "first step analysis" but since time is continuous the first "interesting time-step" is naturally the time when the first jump occurs. More precisely

$$E(T_{02}) = \sum_{k=0}^{2} E(T_{02} \mid X_{J_1} = k, X_0 = 0)\underbrace{P(X_{J_1} = k \mid X_0 = 0)}_{r_{0k}},$$

where $r_{ik} = -q_{ik}/q_{ii}$, $i \neq j$ are the transition probabilities of the jump chain, i.e.

$$\mathbf{R} = \begin{pmatrix} r_{00} & r_{01} & r_{02} \\ r_{10} & r_{11} & r_{12} \\ r_{20} & r_{21} & r_{22} \end{pmatrix} = \begin{pmatrix} 0 & 1 & 0 \\ 2/3 & 0 & 1/3 \\ 0 & 1 & 0 \end{pmatrix}.$$

Since T_{02} is the time it takes to leave state 0 plus the time it takes to reach state 2 from the new state reached after the first jump from 0, it follows from the Markov property that

$$E(T_{02} \mid X_{J_1} = k, X_0 = 0) = E(J_1 \mid X_0 = 0) + E(T_{k2}),\ k = 0, 1,$$

$E(T_{02}|X_{J_1} = 2, X_0 = 0) = E(J_1|X_0 = 0)$.
From the intensity matrix we see that $J_1|(X_0 = 0) \sim \text{Exp}(q_0)$ and $J_1|(X_0 = 1) \sim \text{Exp}(q_1)$, where $q_0 = -q_{00} = 1$, and $q_1 = -q_{11} = 3$. Thus $E(J_1|X_0 = 0) = 1$ and $E(J_1|X_0 = 1) = 1/3$.
Therefore

$$E(T_{02}) = \underbrace{E(J_1 \mid X_0 = 0)}_{=1} + E(T_{12}) r_{01} = 1 + E(T_{12}).$$

Similarly

$$E(T_{12}) = (1/q_1) + r_{10} E(T_{02}) = 1/3 + (2/3) \cdot E(T_{02}).$$

Thus $E(T_{02}) = 1 + E(T_{12}) = 1 + 1/3 + (2/3)E(T_{02})$, i.e. $E(T_{02}) = 4$.

9 Matlab

9.1 Code for solving $\pi\mathbf{Q} = \mathbf{0} \Leftrightarrow \mathbf{Q}'\pi' = \mathbf{0}'$ for probability vectors π.

```
Q=[-1 1 0; 2 -3 1; 0 2 -2]

pi= null(transpose(Q))/ sum(null(transpose(Q)))
```

10 Suggested exercises

Basic exercises:
21–22, 27

Extra problems:
b3, B3, B4

Exercises Lawler:
3.8

Lecture 13 Markov Processes, 1MS012

11 Birth-processes

Definition: A birth-process on $S = (0, 1, 2, \ldots)$ is a Markov process with generator

$$\mathbf{Q} = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdot & \cdot \\ 0 & -\lambda_1 & \lambda_1 & 0 & \cdot & \cdot \\ 0 & 0 & -\lambda_2 & \lambda_2 & 0 & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdot & \cdot \end{pmatrix}.$$

The λ_j:s are called birthrates.
A birth-process is thus a Markovprocess with (trivial) jump chain $r_{j,j+1} = 1$ and holding times λ_j for all states j.

Birth-processes can be used as models for the size of a growing population, where new individuals are born with intensity λ_i if the size of the population is i. The Poisson process corresponds to the special case when $\lambda_i = \lambda$ i.e. the case when the birthrates do not depend on the population size.

11.1 Transition probabilities for a birth-process

Forward equations $\mathbf{P}'(t) = \mathbf{P}(t)\mathbf{Q}$:

$$p'_{ij}(t) = \lambda_{j-1}p_{i,j-1}(t) - \lambda_j p_{i,j}(t), \quad j \geq 1.$$

Since $p_{ij}(t) = 0$, if $j < i$, we get
(I):

$$p'_{ii}(t) = -\lambda_i p_{ii}(t)$$

(II):

$$p'_{ij}(t) = \lambda_{j-1}p_{i,j-1}(t) - \lambda_j p_{ij}(t), \; j > i$$

Since $p_{ii}(0) = 1$, (I) has solution $p_{ii}(t) = e^{-\lambda_i t}$, for all $i \geq 0$.
We now solve (II):

$$\underbrace{e^{\lambda_j t}(p'_{ij}(t) + \lambda_j p_{ij}(t))}_{\frac{d}{dt}(e^{\lambda_j t}p_{ij}(t))} = (\lambda_{j-1}p_{i,j-1}(t))e^{\lambda_j t}$$

By integrating both sides and using $p_{ij}(0) = 0$, for $j > i$, we finally get the recursive equations

$$p_{ij}(t) = e^{-\lambda_j t} \lambda_{j-1} \int_0^t p_{i,j-1}(s) e^{\lambda_j s} ds \tag{3}$$

for $j > i$.

Remark: It is an exercise to explicitly solve the recursive equations (3) in the special case $\lambda_i = \lambda$ (corresponding to the Poisson process), to obtain the solution

$$p_{ij}(t) = e^{-\lambda t} (\lambda t)^{j-i} / (j-i)!, \quad j \geq i,$$

in that case.

11.2 Honest and dishonest processes

Let J_n be the time for the n:th jump. Let $J_\infty = \lim_{n \to \infty} J_n$.

If $J_\infty < \infty$ then the process makes an infinite number of jumps just before time J_∞. The process is then explosive and $\mathbf{Q}$ does not define the process for all times in this case. This occurs if λ_i grows sufficiently fast.

Since $U_i := (J_{i+1} - J_i) \sim \mathrm{Exp}(\lambda_i)$ and thus

$$EJ_\infty = E(\sum_{i=0}^{\infty} U_i) = \sum_{i=0}^{\infty} E(U_i) = \sum_{i=0}^{\infty} 1/\lambda_i,$$

we have:

Theorem: (Feller-Lundberg)

$$\sum_{i=0}^{\infty} 1/\lambda_i < \infty \Rightarrow P(J_\infty < \infty) = 1 \qquad \text{(explosion always)}$$

It is also possible to prove

$$\sum_{i=0}^{\infty} 1/\lambda_i = \infty \Rightarrow P(J_\infty = \infty) = 1 \qquad \text{(no explosion)}$$

Example: Poisson process $\lambda_i = \lambda$, does not explode since $\sum_{i=0}^{\infty} 1/\lambda = \infty$.

Definition: A Markov process $(X_t)_{t \geq 0}$ is said to be

honest if $P(J_\infty = \infty) = 1$
dishonest if $P(J_\infty = \infty) < 1$

We thus have a simple criteria for telling if a birth-process is honest or not. For more general Markov processes it is much harder:

Theorem: The Markov process $(X_t)_{t\geq 0}$ with generator $\mathbf{Q}$ is honest if one of the following conditions hold.

- $\sup_j q_j < \infty$
- The state space S is finite
- The process starts in a recurrent state.

Note that these are only sufficient conditions for a Markov process to be honest.

Theorem: If $(X_t)_{t\geq 0}$ is irreducible with generator $\mathbf{Q}$ then:

$$\text{all states are positive recurrent} \quad \Leftrightarrow \quad \begin{array}{l} (X_t)_{t\geq 0} \text{ is honest and } \pi\mathbf{Q} = \mathbf{0} \\ \text{for some probability distribution } \pi. \end{array}$$

Obs:
Birth-processes are not irreducible, and stationary distributions do not exist for such processes.

12 Birth-death processes

Definition: A birth death process is a Markov process on $S = (0, 1, 2, ...)$ with generator

$$\mathbf{Q} = \begin{pmatrix} -\lambda_0 & \lambda_0 & 0 & 0 & \cdot & \cdot \\ \mu_1 & -(\lambda_1+\mu_1) & \lambda_1 & 0 & \cdot & \cdot \\ 0 & \mu_2 & -(\lambda_2+\mu_2) & \lambda_2 & 0 & \cdot \\ . & . & . & . & . & . \\ . & . & . & . & . & . \\ . & . & . & . & . & . \end{pmatrix}$$

Thus $i \mapsto i+1$ (birth) and $i \mapsto i-1$ (death) are the only possible transitions.

$\lambda_0, \lambda_1, \ldots$ birthrates/arrival rates
$\mu_1, \mu_2, \ldots$ deathrates/departure rates

(the latter notation if regarded as the number of people in a system)

The process will be irreducible if $\lambda_n > 0$, and $\mu_n > 0$ for all n.

In order to find stationary distributions we solve $\pi\mathbf{Q} = \mathbf{0}$.

$$-\lambda_0\pi_0 + \mu_1\pi_1 = 0 \tag{I}$$
$$\lambda_{n-1}\pi_{n-1} - (\lambda_n + \mu_n)\pi_n + \mu_{n+1}\pi_{n+1} = 0, \qquad n \geq 1 \tag{II}$$

Thus

$$\mu_{n+1}\pi_{n+1} - \lambda_n\pi_n \underbrace{=}_{\text{(II)}} \mu_n\pi_n - \lambda_{n-1}\pi_{n-1} \underbrace{=}_{\text{(II)}} \ldots = \mu_1\pi_1 - \lambda_0\pi_0 \underbrace{=}_{\text{(I)}} 0,\ n \geq 0,$$

i.e.

$$\mu_{n+1}\pi_{n+1} = \lambda_n\pi_n, \qquad n \geq 0, \qquad\qquad \text{(detailed balance equations)}$$

so

$$\pi_n = \frac{\lambda_{n-1}}{\mu_n}\pi_{n-1} = \ldots = \frac{\lambda_{n-1}\lambda_{n-2}\cdots\lambda_0}{\mu_n\mu_{n-1}\cdots\mu_1}\pi_0, \qquad n \geq 1.$$

If π is a probability distribution then

$$\sum_{n=0}^{\infty}\pi_n = 1 \Leftrightarrow (1 + \sum_{n=1}^{\infty}\frac{\lambda_{n-1}\lambda_{n-2}\cdots\lambda_0}{\mu_n\mu_{n-1}\cdots\mu_1})\pi_0 = 1$$

This partly proves

Theorem:
Let $(X_t)_{t\geq 0}$ be a birth-death process with birth intensities $\lambda_i > 0$, $i \geq 0$, and death intensities $\mu_i > 0$, $i \geq 1$.
If

$$\sum_{k=1}^{\infty}\frac{\lambda_0\lambda_1\cdots\lambda_{k-1}}{\mu_1\mu_2\cdots\mu_k} < \infty$$

and

$$\sum_{k=1}^{\infty}\frac{\mu_1\mu_2\cdots\mu_k}{\lambda_1\lambda_2\cdots\lambda_k} = \infty \tag{4}$$

then a unique stationary distribution π, exists and

$$p_{ij}(t) = P(X_t = j|X_0 = i) \to \pi_j,$$

as $t \to \infty$, where

$$\pi_k = \frac{\lambda_0\lambda_1\cdots\lambda_{k-1}}{\mu_1\mu_2\cdots\mu_k}\pi_0, \qquad \text{and} \qquad \pi_0 = (1+\sum_{k=1}^{\infty}\frac{\lambda_0\lambda_1\cdots\lambda_{k-1}}{\mu_1\mu_2\cdots\mu_k})^{-1}.$$

Remark: It can be proved that (4) is a condition for recurrence of the state 0.

Remark: If $\lambda_{n_0} = 0$, for some n_0, then we regard $S = (0, \ldots, n_0)$ as the state space and if the process is irreducible (regarded as a process on S), then

$$p_{ij}(t) = P(X_t = j|X_0 = i) \to \pi_j,$$

as $t \to \infty$, where

$$\pi_k = \frac{\lambda_0\lambda_1\cdots\lambda_{k-1}}{\mu_1\mu_2\cdots\mu_k}\pi_0, \qquad \text{and} \qquad \pi_0 = (1+\sum_{k=1}^{n_0}\frac{\lambda_0\lambda_1\cdots\lambda_{k-1}}{\mu_1\mu_2\cdots\mu_k})^{-1}.$$

12.1 Population models

Let $X_t =$ number of individuals in a population at time t.
Suppose $X_t = i$, at some fixed time t. Suppose $B \sim \mathrm{Exp}(\lambda_i)$ is the time until next birth, and $D \sim \mathrm{Exp}(\mu_i)$ is the time until the next death, and B and D are independent. If $D < B$ then the next jump will be $i \mapsto i-1$. By properties of the exponential distribution if follows that $P(D < B) = \frac{\mu_i}{\lambda_i+\mu_i}$.
If J_1 is the time for the next jump, then $J_1 = \min(B, D) \sim \mathrm{Exp}(\lambda_i + \mu_i)$.
It follows from the memoryless property of the exponential distribution that conditional on $X_{J_1} = j$, the process starts afresh from j at time J_1. Thus $(X_t)_{t\geq 0}$ is a birth-death process with birthrates λ_i, and deathrates μ_i.

Example: Each individual produces new individuals with rate λ and dies with rate μ.
This is a birth-death process with $\lambda_n = n\lambda$, $\mu_n = n\mu$.

12.2 Some common queueing models

$X_t =$ number of people on line for service.
Independent interarrival times $\sim \mathrm{Exp}(\lambda)$, $\lambda > 0$, and service times $\sim \mathrm{Exp}(\mu)$, $\mu > 0$.

Example: $M/M/1$ queue
One server.
This is a birth-death process with $\lambda_n = \lambda$, $\mu_n = \mu$.

(The two M:s refer to the fact that both inter-arrival and service times are exponential and thus memoryless, and hence arrival and service times are Markovian.)
If $\lambda < \mu$ then $\pi_k = (1 - \frac{\lambda}{\mu})(\frac{\lambda}{\mu})^k$, is the unique stationary distribution.
(Shifted geometric distribution with parameter $1 - \frac{\lambda}{\mu}$.)

Example: $M/M/k$ queue
k servers.
This is a birth-death process with $\lambda_n = \lambda$, $\mu_n = \begin{cases} n\mu & \text{if } n \leq k \\ k\mu & \text{if } n > k \end{cases}$.
If $\lambda < k\mu$ then there is a unique stationary distribution.

Example: $M/M/\infty$ queue
infinitely many servers.
This is a birth-death process with $\lambda_n = \lambda$, $\mu_n = n\mu$.
The distribution $\pi_k = e^{-\lambda/\mu}\frac{(\lambda/\mu)^k}{k!}$, is the unique stationary distribution.
(Poisson distribution with parameter λ/μ.)

13 Suggested exercises

Basic exercises:
23–26, 28

Extra problems:
c1, B2

Exercises Lawler:
3.12