STAT-F-407
Markov Chains

Thomas Verdebout

Université Libre de Bruxelles

Outline of the course

1. A short introduction.
2. Basic probability review.
3. Martingales.
4. Markov chains.
    4.1. Definitions and examples.
    4.2. Strong Markov property, number of visits.
    4.3. Classification of states.
    4.4. Computation of R and F.
    4.5. Asymptotic behavior.
5. Markov processes, Poisson processes.
6. Brownian motions.

Outline of the course

1. A short introduction.
2. Basic probability review.
3. Martingales.
4. Markov chains.
   - $\boxed{\text{4.1. Definitions and examples.}}$
   - 4.2. Strong Markov property, number of visits.
   - 4.3. Classification of states.
   - 4.4. Computation of R and F.
   - 4.5. Asymptotic behavior.
5. Markov processes, Poisson processes.
6. Brownian motions.

Markov chains

The importance of these processes comes from two facts:

- there is a large number of physical, biological, economic, and social phenomena that can be described in this way, and
- there is a well-developed theory that allows for doing the computations and obtaining explicit results...

Let S be a finite or countable set (number its elements using $i = 1, 2, \ldots$
Let $(X_n)_{n\in\mathbb{N}}$ be a SP with $X_n : (\Omega, \mathcal{A}, P) \to S$ for all n.

Definition: (X_n) is a Markov chain (MC) on $S \Leftrightarrow$

$$\mathbb{P}[X_{n+1} = j | X_0, X_1, \ldots, X_n] = \mathbb{P}[X_{n+1} = j | X_n] \quad \forall n \, \forall j.$$

Remarks:

- The equation above is the so-called Markov property. It states that the future does only depend on the present state of the process, and not on its past.
- S is the state space.
- The elements of S are the states.

Definition: The MC (X_n) is homogeneous ($\rightsquigarrow$ HMC) $\Leftrightarrow$

$$\mathbb{P}[X_{n+1} = j | X_n = i] = \mathbb{P}[X_1 = j | X_0 = i] \quad \forall n \, \forall i, j.$$

For a HMC, one can define the transition probabilities

$$p_{ij} = \mathbb{P}[X_{n+1} = j | X_n = i] \quad \forall i, j,$$

which are usually collected in the transition matrix $P = (p_{ij})$.

The transition matrix P is a "stochastic matrix", which means that

- $p_{ij} \in [0, 1]$ for all i, j.
- $\sum_j p_{ij} = 1$ for all i.

In vector notation, $P\mathbf{1} = \mathbf{1}$, where $\mathbf{1}$ stands for the vector of ones with the appropriate dimension.

Example 1: random walk

Let $X_1, X_2, \ldots$ be i.i.d., with $\mathbb{P}[X_i = 1] = p$ and $\mathbb{P}[X_i = -1] = q = 1 - p$
Let $Y_i := \sum_{i=1}^{n} X_i$ be the corresponding random walk.

$\rightsquigarrow$ (Y_n) is a HMC on $S = \mathbb{Z}$ with transition matrix

$$P = \begin{pmatrix} \ddots & \ddots & \ddots & & & & \\ & q & 0 & p & & & \\ & & q & 0 & p & & \\ & & & q & 0 & p & \\ & & & & \ddots & \ddots & \ddots \end{pmatrix}.$$

Example 2: rw with absorbing barriers

Let $X_1, X_2, \ldots$ be i.i.d., with $\mathbb{P}[X_i = 1] = p$ and $\mathbb{P}[X_i = -1] = q = 1 - p$
Let $Y_0 := k \in \{1, 2, \ldots, m-1\}$ be the initial state.
Let $Y_{n+1} := (Y_n + X_{n+1})\, \mathbb{I}_{[Y_n \notin \{0,m\}]} + Y_n\, \mathbb{I}_{[Y_n \in \{0,m\}]}$.

$\rightsquigarrow$ (Y_n) is a HMC on $S = \{0, 1, \ldots, m\}$ with transition matrix

$$P = \begin{pmatrix} 1 & 0 & \ldots & & & & \\ q & 0 & p & & & & \\ & q & 0 & p & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & & q & 0 & p \\ & & & & \ldots & 0 & 1 \end{pmatrix}.$$

Example 3: success runs

Let $X_1, X_2, \ldots$ be i.i.d., with $\mathbb{P}[X_i = 1] = p$ and $\mathbb{P}[X_i = 0] = q = 1 - p$.
Let $Y_0 := 0$ be the initial state.
Let $Y_{n+1} := (Y_n + 1)\,\mathbb{I}_{[X_{n+1}=1]} + 0 \times \mathbb{I}_{[X_{n+1}=0]}$.

$\rightsquigarrow (Y_n)$ is a HMC on $S = \mathbb{N}$ with transition matrix

$$P = \begin{pmatrix} q & p & 0 & \cdots & & & \\ q & 0 & p & 0 & \cdots & & \\ q & 0 & 0 & p & 0 & \cdots & \\ \vdots & & & \ddots & \ddots & & \end{pmatrix}.$$

Example 4: discrete queue models

Let Y_n be the number of clients in a queue at time n ($Y_0 = 0$).
Let X_n be the number of clients entering the shop between time $n-1$ and n (X_n i.i.d., with $\mathbb{P}[X_n = i] = p_i$; $\sum_{i=0}^{\infty} p_i = 1$).
Assume a service needs exactly one unit of time to be completed.

Then

$$Y_{n+1} = (Y_n + X_n - 1)\,\mathbb{I}_{[Y_n>0]} + X_n\,\mathbb{I}_{[Y_n=0]},$$

and (Y_n) is a HMC on $S = \mathbb{N}$ with transition matrix

$$P = \begin{pmatrix} p_0 & p_1 & p_2 & \dots & & \\ p_0 & p_1 & p_2 & \dots & & \\ 0 & p_0 & p_1 & p_2 & \dots & \\ 0 & 0 & p_0 & p_1 & p_2 & \dots \\ \vdots & & \ddots & \ddots & \ddots & \ddots \end{pmatrix}.$$

Example 5: stock management

Let X_n be the number of units on hand at the end of day n ($X_0 = M$).
Let D_n be the demand on day n (D_n i.i.d., $\mathbb{P}[D_n = i] = p_i$; $\sum_{i=0}^{\infty} p_i = 1$).
Assume that if $X_n \leq m$, it is (instantaneously) set to M again.

Then, letting $x^+ = \max(x, 0)$, we have

$$X_{n+1} = (X_n - D_{n+1})^+ \mathbb{I}_{[X_n > m]} + (M - D_{n+1})^+ \mathbb{I}_{[X_n \leq m]},$$

and (X_n) is a HMC on $S = \{0, 1, \dots, M\}$ (exercise: derive P).

Questions:

- if we make 12\$ profit on each unit sold but it costs 2\$ a day to store items, what is the long-run profit per day of this inventory policy?
- How to choose (m, M) to maximize profit?

Example 6: income classes

Assume that from one generation to the next, families change their income group "Low", "Middle", or "High" (state 1,2, and 3, respectively) according to a HMC with transition matrix

$$P = \begin{pmatrix} .6 & .3 & .1 \\ .2 & .7 & .1 \\ .1 & .3 & .6 \end{pmatrix}.$$

Questions:

- Do the fractions of the population in the three income classes stabilize as time goes on?
- If this happens, how can we compute the limiting proportions from P?

Higher-order transition probabilities

We let $P = (p_{ij})$, where $p_{ij} = \mathbb{P}[X_1 = j | X_0 = i]$.
Now, define $P^{(n)} = (p_{ij}^{(n)})$, where $p_{ij}^{(n)} = \mathbb{P}[X_n = j | X_0 = i]$.

What is the link between P and $P^{(n)}$?

$\rightsquigarrow$ **Theorem**: $P^{(n)} = P^n$.

Proof: the result holds for $n = 1$. Now, assume it holds for n. Then $(P^{(n+1)})_{ij} = \mathbb{P}[X_{n+1} = j | X_0 = i] = \sum_k \mathbb{P}[X_{n+1} = j, X_n = k | X_0 = i] = \sum_k \mathbb{P}[X_{n+1} = j | X_n = k, X_0 = i] \mathbb{P}[X_n = k | X_0 = i]$
$= \sum_k \mathbb{P}[X_{n+1} = j | X_n = k] \mathbb{P}[X_n = k | X_0 = i] =$
$\sum_k (P^{(n)})_{ik} (P^{(1)})_{kj} = (P^{(n)} P)_{ij} = (P^n P)_{ij} = (P^{n+1})_{ij}$, so that the result holds for $n + 1$ as well. $\square$

Higher-order transition probabilities

Of course, this implies that
$P^{(n+m)} = P^{n+m} = P^n P^m = P^{(n)} P^{(m)}$, that is,

$$\mathbb{P}[X_{n+m} = j | X_0 = i] = \sum_k \mathbb{P}[X_n = k | X_0 = i] \mathbb{P}[X_m = j | X_0 = k],$$

which are the so-called Chapman-Kolmogorov equations.

Higher-order transition probabilities

Clearly, the distribution of X_n is of primary interest.

Let $a^{(n)}$ be the (line) vector with jth component $(a^{(n)})_j = \mathbb{P}[X_n = j]$.
$\rightsquigarrow$ **Theorem**: $a^{(n)} = a^{(0)} P^n$.

Proof: using the total probability formula, we obtain $(a^{(n)})_j = \mathbb{P}[X_n = j] = \sum_k \mathbb{P}[X_n = j | X_0 = k] \mathbb{P}[X_0 = k] = \sum_k (a^{(0)})_k (P^{(n)})_{kj} = (a^{(0)} P^{(n)})_j = (a^{(0)} P^n)_j$, which establishes the result. $\square$

This shows that one can very easily compute the distribution of X_n in terms of

- the distribution of X_0, and
- the transition matrix P.

Higher-order transition probabilities

Proposition: let (X_n) be a HMC on S. Then

$$\mathbb{P}[X_1 = i_1, X_2 = i_2, \dots, X_n = i_n | X_0 = i_0],$$
$$\mathbb{P}[X_{m+1} = i_1, X_{m+2} = i_2, \dots, X_{m+n} = i_n | X_m = i_0],$$

and

$$\mathbb{P}[X_{m+1} = i_1, X_{m+2} = i_2, \dots, X_{m+n} = i_n | X_0 = j_0, X_1 = j_1, \dots, X_m = i_0]$$

all are equal to $p_{i_0 i_1} p_{i_1 i_2} \dots p_{i_{n-1} i_n}$.

Proof: exercise...

Outline of the course

1. A short introduction.
2. Basic probability review.
3. Martingales.
4. $\boxed{\text{4. Markov chains.}}$
    4.1. Definitions and examples.
    4.2. $\boxed{\text{4.2. Strong Markov property, number of visits.}}$
    4.3. Classification of states.
    4.4. Computation of R and F.
    4.5. Asymptotic behavior.
5. Markov processes, Poisson processes.
6. Brownian motions.

Strong Markov property

Quite similarly as for the optional stopping theorem for martingales, $\mathbb{P}[X_{n+1} = j | X_n = i] = p_{ij}$ does also hold at stopping times T. This is the so-called “strong Markov property" (SMP).

An illustration: for $0 < p, q < 1$ ($p + q = 1$), consider the HMC with graph
Let T be the time of first visit in $2 \rightsquigarrow \mathbb{P}[X_{T+1} = 1 | X_T = 2] = q(= p_{21})$.

Let T be the time of last visit in $2 \rightsquigarrow \mathbb{P}[X_{T+1} = 1 | X_T = 2] = q^\infty = 0$ $(\neq p_{21})$, which shows that the SMP may be violated if T is not a ST.

Numbers of visits

Of particular interest is also the total number of visits in j, that is $N_j = \sum_{n=0}^{\infty} I_{[X_n=j]}$.

To determine the distribution of N_j, let

- $T_j = \inf\{n \in \mathbb{N}_0 | X_n = j\}$, which is the time of first visit in j (if $X_0 \neq j$) or of first return to j (if $X_0 = j$), and
- $f_{ij} = \mathbb{P}[T_j < \infty | X_0 = i]$.

Let θ_k be the time of kth visit of the chain in j (if there are only k visits in j, we let $\theta_\ell = \infty$ for all $\ell \geq k+1$ and $\theta_{\ell+1} - \theta_\ell = \infty$ for all $\ell \geq k$).

Then, for $k = 1, 2, \ldots$,

$$\begin{aligned}
& \mathbb{P}[N_j = k | X_0 = i] \\
= \; & \mathbb{P}[\theta_1 < \infty, \ldots, \theta_k < \infty, \theta_{k+1} = \infty | X_0 = i] \\
= \; & \mathbb{P}[\theta_1 < \infty | X_0 = i] \ldots \mathbb{P}[\theta_{k+1} = \infty | \theta_1 < \infty, ..., \theta_k < \infty, X_0 = i] \\
= \; & \mathbb{P}[\theta_1 < \infty | X_0 = i] (\mathbb{P}[\theta_1 < \infty | X_0 = j])^{k-1} \mathbb{P}[\theta_1 = \infty | X_0 = j] \\
= \; & f_{ij} f_{jj}^{k-1} (1 - f_{jj}).
\end{aligned}$$

Numbers of visits

Working similarly, one shows that

$$\mathbb{P}[N_j = k | X_0 = i] = \begin{cases} f_{ij} f_{jj}^{k-1}(1 - f_{jj}) & \text{if} \quad k > 0 \\ 1 - f_{ij} & \text{if} \quad k = 0 \end{cases}$$

for $i \neq j$, and $\mathbb{P}[N_j = k | X_0 = j] = f_{jj}^{k-1}(1 - f_{jj}), \quad k > 0.$

Hence, letting $r_{ij} = E[N_j | X_0 = i]$ be the expected number of visits in j when starting from i, we have, for $i \neq j$,

$$r_{ij} = \sum_{k=0}^{\infty} k \, \mathbb{P}[N_j = k | X_0 = i] = f_{ij}(1 - f_{jj}) \sum_{k=1}^{\infty} k \, f_{jj}^{k-1} = \frac{f_{ij}}{1 - f_{jj}}$$

and

$$r_{jj} = \sum_{k=0}^{\infty} k \, \mathbb{P}[N_j = k | X_0 = j] = (1 - f_{jj}) \sum_{k=1}^{\infty} k \, f_{jj}^{k-1} = \frac{1}{1 - f_{jj}}.$$

Numbers of visits

Similarly as for the transition probabilities p_{ij}, the $r_{ij} = \mathbb{E}[N_j | X_0 = i]$ will be collected in some matrix $R = (r_{ij})$.

Note that

$$r_{ij} = \mathbb{E}\Big[\sum_{n=0}^{\infty} I_{[X_n = j]} | X_0 = i\Big] = \sum_{n=0}^{\infty} \mathbb{E}[I_{[X_n=j]} | X_0 = i]$$

$$= \sum_{n=0}^{\infty} \mathbb{P}[X_n = j | X_0 = i] = \sum_{n=0}^{\infty} p_{ij}^{(n)} = \sum_{n=0}^{\infty} (P^n)_{ij},$$

which shows that

$$R = \sum_{n=0}^{\infty} P^n.$$

Outline of the course

1. A short introduction.
2. Basic probability review.
3. Martingales.
4. Markov chains.
   - 4.1. Definitions and examples.
   - 4.2. Strong Markov property, number of visits.
   - 4.3. Classification of states.
   - 4.4. Computation of R and F.
   - 4.5. Asymptotic behavior.
5. Markov processes, Poisson processes.
6. Brownian motions.

Classification of states

Definition:

- the state j is transient $\Leftrightarrow f_{jj} < 1$.
- the state j is recurrent $\Leftrightarrow f_{jj} = 1$.

Remarks:

- j transient $\Leftrightarrow r_{jj} < \infty$; j recurrent $\Leftrightarrow r_{jj} = \infty$.
- j transient $\Rightarrow \mathbb{P}[T_j = \infty | X_0 = j] > 0 \Rightarrow \mathbb{E}[T_j | X_0 = j] = \infty$.
- j recurrent $\Rightarrow \mathbb{P}[T_j = \infty | X_0 = j] = 0$, but $\mathbb{E}[T_j | X_0 = j]$ can be finite or infinite...

$\rightsquigarrow$ Definition:

- j is positive-recurrent $\Leftrightarrow$
  j is recurrent and $\mathbb{E}[T_j | X_0 = j] < \infty$.
- j is null-recurrent $\Leftrightarrow$ j is recurrent and $\mathbb{E}[T_j | X_0 = j] = \infty$.

Classification of states

Definition:
j is accessible from i (not. $i \to j$) $\Leftrightarrow \exists n \in \mathbb{N}$ such that $p_{ij}^{(n)} > 0$ (that is, there is some path, from i to j, in the graph of the HMC).

Letting $\alpha_{ij} = \mathbb{P}[\text{go to } j \text{ before coming back to } i | X_0 = i]$, the following are equivalent

- $i \to j$.
- $\exists n \in \mathbb{N}$ such that $(P^n)_{ij} > 0$.
- $f_{ij} > 0$.
- $\alpha_{ij} > 0$.

Classification of states

Definition: i and j communicate (not.: $i \leftrightarrow j$) $\Leftrightarrow$ $i \to j$ and $j \to i$.

This allows for a partition of the state space S into classes (=subsets of S in which states communicate with each other).

Classification of states

$\rightsquigarrow$ two types of classes:

- $\mathcal{C}$ is open $\Leftrightarrow \forall i \in \mathcal{C}$, there is some $j \notin \mathcal{C}$ such that $i \to j$.
- $\mathcal{C}$ is closed $\Leftrightarrow \forall i \in \mathcal{C}$, there is no $j \notin \mathcal{C}$ such that $i \to j$.

There are strong links between the types of classes and the types of states...

Proposition: *all states in an open class $\mathcal{C}$ are transient.*

Proof: let $i \in \mathcal{C}$. Then there is some $j \notin \mathcal{C}$ such that $i \to j$ (and hence $j \not\to i$). We then have

$$\begin{aligned} 1 - f_{ii} &= \mathbb{P}[T_i = \infty | X_0 = i] \\ &\geq \mathbb{P}[\text{go to } j \text{ before coming back to } i | X_0 = i] \\ &= \alpha_{ij} > 0, \end{aligned}$$

so that i is transient. $\square$

Classification of states

What about states in a closed class?

Proposition: *let $\mathcal{C}$ be a closed class. Then if there is some recurrent state $i \in \mathcal{C}$, all states in $\mathcal{C}$ are recurrent.*

Proof: let $j \in \mathcal{C}$. Choose $r, s \in \mathbb{N}$ such that $(P^r)_{ij} > 0$ and $(P^s)_{ji} > 0$ (existence since $i \leftrightarrow j$). Then j is recurrent since

$$\begin{aligned} r_{jj} &= \sum_{n=0}^{\infty} (P^n)_{jj} \geq \sum_{n=r+s}^{\infty} (P^n)_{jj} = \sum_{m=0}^{\infty} (P^s P^m P^r)_{jj} \\ &= \sum_{m=0}^{\infty} \sum_{k,\ell} (P^s)_{jk} (P^m)_{k\ell} (P^r)_{\ell j} \\ &\geq \sum_{m=0}^{\infty} (P^s)_{ji} (P^m)_{ii} (P^r)_{ij} \\ &= (P^s)_{ji} r_{ii} (P^r)_{ij} = \infty. \end{aligned}$$

$\square$

Classification of states

Proposition: *let $\mathcal{C}$ be a closed class. Then if there is some recurrent state $i \in \mathcal{C}$, all states in $\mathcal{C}$ are recurrent.*

This result shows that recurrent and transient states do not mix in a closed class. Actually, it can be shown that:

Consequently, a closed class contains either

- transient states only, or
- positive-recurrent states only, or
- null-recurrent states only.

The following result is very useful:

Proposition: *let $\mathcal{C}$ be a closed class, with $\#\mathcal{C} < \infty$. Then all states in $\mathcal{C}$ are positive-recurrent.*

How would look a closed class with transient states?

An example: with $p + q = 1$, consider the chain

If $p > \frac{1}{2}$, one can show all states are transient...

Classification of states

A last result in this series:

Proposition: *let $\mathcal{C}$ be a closed class, with recurrent states. Then $f_{ij} = 1$ for all $i, j \in \mathcal{C}$.*

Proof: let $i, j \in \mathcal{C}$. Since j is recurrent, $f_{jj} = 1$, so that

$$\begin{aligned} 0 = 1 - f_{jj} = \mathbb{P}[T_j = \infty | X_0 = j] \\ &\geq \mathbb{P}[\text{go to } i \text{ before coming back to } j, \\ &\qquad \text{and then never come back to } j | X_0 = j] \\ &= \alpha_{ji}(1 - f_{ij}). \end{aligned}$$

Hence, $\alpha_{ji}(1 - f_{ij}) = 0$. Since $\alpha_{ji} > 0$ ($j \to i$), we must have $f_{ij} = 1$. $\square$

Outline of the course

1. A short introduction.
2. Basic probability review.
3. Martingales.
4. Markov chains.
    - 4.1. Definitions and examples.
    - 4.2. Strong Markov property, number of visits.
    - 4.3. Classification of states.
    - 4.4. Computation of R and F.
    - 4.5. Asymptotic behavior.
5. Markov processes, Poisson processes.
6. Brownian motions.

Computation of R and F

In this section, we describe a systematic method that allows for computing the matrices

$$R = (r_{ij})$$

where

$$r_{ij} = \mathbb{E}[N_j | X_0 = i]$$

is the expected number of visits in j when starting from i, and

$$F = (f_{ij})$$

where

$$f_{ij} = \mathbb{P}[T_j < \infty | X_0 = i]$$

is the probability that, being in i, the HMC will visit j in the future.

Computation of R and F

The first step consists in renumerating the states in such a way the indices of recurrent states are smaller than those of transient ones. (remark: we assume $\#S < \infty$ in this section)

Consequently, the transition matrix can be partitioned into

$$P = \begin{pmatrix} P_{rr} & P_{rt} \\ P_{tr} & P_{tt} \end{pmatrix},$$

where P_{tr} is the transition matrix from transient states to recurrent ones, P_{rr} that between recurrent states, and so on...

Of course, we will partition accordingly

$$R = \begin{pmatrix} R_{rr} & R_{rt} \\ R_{tr} & R_{tt} \end{pmatrix} \quad \text{and} \quad F = \begin{pmatrix} F_{rr} & F_{rt} \\ F_{tr} & F_{tt} \end{pmatrix}.$$

Computation of R and F

Actually, $P_{rt} = 0$.

Indeed, if i is recurrent and j is transient, i belongs to some closed class $\mathcal{C}_1$, while j belongs to another class $\mathcal{C}_2$ (otherwise, j would be recurrent as well). Hence, $i \nrightarrow j$, so that $p_{ij} = 0$.

Clearly, this also implies that $R_{rt} = 0$ and $F_{rt} = 0$.

(a) Computation of R

We start with the computation of

$$R = \begin{pmatrix} R_{rr} & R_{rt} \\ R_{tr} & R_{tt} \end{pmatrix} = \begin{pmatrix} R_{rr} & 0 \\ R_{tr} & R_{tt} \end{pmatrix}.$$

Previously, we showed that $R = \sum_{n=0}^{\infty} P^n$, so that

$$\begin{pmatrix} R_{rr} & 0 \\ R_{tr} & R_{tt} \end{pmatrix} = R = \sum_{n=0}^{\infty} \begin{pmatrix} ? & 0 \\ ? & P_{tt}^n \end{pmatrix} = \begin{pmatrix} ? & 0 \\ ? & \sum_{n=0}^{\infty} P_{tt}^n \end{pmatrix},$$

which yields that

$$R_{tt} = \sum_{n=0}^{\infty} P_{tt}^n = I + \sum_{n=1}^{\infty} P_{tt}^n = I + P_{tt} \sum_{n=1}^{\infty} P_{tt}^{n-1} = I + P_{tt} R_{tt},$$

so that $R_{tt} = (I - P_{tt})^{-1}$.

(a) Computation of R

It remains to compute the entries r_{ij}, where j is recurrent.

$\rightsquigarrow$ **Proposition**: *for such entries, (i)* $r_{ij} = \infty$ *if* $i \to j$ *and (ii)* $r_{ij} = 0$ *if* $i \not\to j$.

Proof:

(i) in the previous lecture, we have shown that $r_{ij} = f_{ij}/(1 - f_{jj})$ and $r_{jj} = 1/(1 - f_{jj})$, so that $r_{ij} = f_{ij} r_{jj}$. Now, if $i \to j$, we have $f_{ij} > 0$, so that $r_{ij} = f_{ij} r_{jj} = f_{ij} \times \infty = \infty$ (since j is recurrent).

(ii) is trivial, since $i \not\to j$ implies that $N_j | [X_0 = i] = 0$ a.s., which yields $r_{ij} = \mathbb{E}[N_j | X_0 = i] = 0$. $\square$

(b) Computation of F

We now go to the computation of

$$F = \begin{pmatrix} F_{rr} & F_{rt} \\ F_{tr} & F_{tt} \end{pmatrix} = \begin{pmatrix} F_{rr} & 0 \\ F_{tr} & F_{tt} \end{pmatrix}.$$

(i) $F_{rr} = ?$

If $i \nrightarrow j$, $f_{ij} = \mathbb{P}[T_j < \infty | X_0 = i] = 0$.

If $i \to j$, then we must also have $j \to i$ (indeed, $j \nrightarrow i$ would imply that i belongs to an open class, and hence that i is transient). Therefore, i and j are recurrent states belonging to the same class, so that $f_{ij} = 1$ (cf. previously).

(b) Computation of F

(ii) $F_{tt} = ?$

By inverting

$$\begin{cases} r_{jj} = \dfrac{1}{1 - f_{jj}} \\ r_{ij} = \dfrac{f_{ij}}{1 - f_{jj}}, \end{cases}$$

we obtain

$$\begin{cases} f_{jj} = 1 - \dfrac{1}{r_{jj}} \\ f_{ij} = \dfrac{r_{ij}}{r_{jj}}, \end{cases}$$

which does the job since $R = (r_{ij})$ has already been obtained...

(b) Computation of F

(iii) $F_{tr} = ?$

Complicated...(discussion).

We start with a lemma:

$\rightsquigarrow$ **Lemma**: *let i be transient. Let j, k be recurrent states in the same class $\mathcal{C}$. Then $f_{ij} = f_{ik}$.*

Proof: since j, k are recurrent states in the same class, $f_{jk} = 1$. Hence,

$f_{ik} = \mathbb{P}[T_k < \infty | X_0 = i] \geq \mathbb{P}[\text{go to } j, \text{then go to } k | X_0 = i] = f_{ij} f_{jk} = f_{ij}.$

Similarly, we obtain $f_{ij} \geq f_{ik}$, so that $f_{ik} = f_{ij}$ $\square$

Therefore, it is sufficient to compute $\mathbb{P}[T_{\mathcal{C}} < \infty | X_0 = i]$ for each transient state i and for each class of recurrent states $\mathcal{C}$.

(b) Computation of F

To achieve this, consider the new HMC $(\tilde{X}_n)$ on $\tilde{S}$, for which

- the transient states of S remain transient states in $\tilde{S}$, and
- each class $\mathcal{C}_k$ $(k = 1, \ldots, K)$ of recurrent states gives birth to a single recurrent state k in $\tilde{S}$.

The transition matrix $\tilde{P}$ of $(\tilde{X}_n)$ is

$$\tilde{P} = \begin{pmatrix} \tilde{P}_{rr} & \tilde{P}_{rt} \\ \tilde{P}_{tr} & \tilde{P}_{tt} \end{pmatrix} = \begin{pmatrix} I_K & 0 \\ B & P_{tt} \end{pmatrix},$$

where $B_{ik} = \mathbb{P}[\tilde{X}_1 = k | \tilde{X}_0 = i] = \sum_{j \in \mathcal{C}_k} \mathbb{P}[X_1 = j | X_0 = i]$.

Now, letting $T_{\mathcal{C}_k} := \inf\{n \in \mathbb{N} | X_n \in \mathcal{C}_k\} = \inf\{n \in \mathbb{N} | \tilde{X}_n = k\}$, the previous lemma states that $g_{ik} = \mathbb{P}[T_{\mathcal{C}_k} < \infty | X_0 = i]$ is the common value of the f_{ij}'s, $j \in \mathcal{C}_k$.

(b) Computation of F

$\rightsquigarrow$ **Proposition**: *let $G = (g_{ik})$, where $g_{ik} = \mathbb{P}[T_{\mathcal{C}_k} < \infty | X_0 = i]$. Then $G = R_{tt}B$.*

Proof:

$$\begin{aligned} g_{ik} &= \mathbb{P}[T_{\mathcal{C}_k} < \infty | X_0 = i] = \lim_{n\to\infty} \mathbb{P}[X_n \in \mathcal{C}_k | X_0 = i] \\ &= \lim_{n\to\infty} \mathbb{P}[\tilde{X}_n = k | \tilde{X}_0 = i] = \lim_{n\to\infty} (\tilde{P}^n)_{ik}. \end{aligned}$$

Now, it is easy to check that

$$\tilde{P}^n = \begin{pmatrix} I_K & 0 \\ B^{(n)} & P_{tt}^n \end{pmatrix},$$

where $B^{(n)} = B + P_{tt}B + P_{tt}^2 B + \ldots + P_{tt}^{n-1}B$. Hence,

$$\begin{aligned} G &= \lim_{n\to\infty} B^{(n)} = \lim_{n\to\infty} (B + P_{tt}B + P_{tt}^2 B + \ldots + P_{tt}^{n-1}B) \\ &= \Big(\sum_{n=0}^{\infty} P_{tt}^n \Big) B = R_{tt}B. \end{aligned}$$

$\square$

An example

A and B own together 6\$. They sequentially bet 1\$ when flipping a (fair) coin. Let X_n be the fortune of A after game n.
The game ends as soon as some player is ruined.

$\rightsquigarrow$ (X_n) is a HMC with transition matrix

$$P = \left(\begin{array}{c|ccccc|c} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ \hline 1/2 & 0 & 1/2 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ \hline 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{array}\right).$$

An example

We first have to renumerate the states in such a way recurrent states come before transient ones:

$\rightsquigarrow$ (X_n) is a HMC with transition matrix

$$P = \left(\begin{array}{cc|ccccc} 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ \hline 1/2 & 0 & 0 & 1/2 & 0 & 0 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 0 & 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 1/2 & 0 & 0 & 0 & 1/2 & 0 \end{array}\right).$$

An example

The computation of R is immediate, but for the block R_{tt}, which is given by $R_{tt} = (I - P_{tt})^{-1}$

$$
= \begin{pmatrix} 1 & -\frac{1}{2} & 0 & 0 & 0 \\ -\frac{1}{2} & 1 & -\frac{1}{2} & 0 & 0 \\ 0 & -\frac{1}{2} & 1 & -\frac{1}{2} & 0 \\ 0 & 0 & -\frac{1}{2} & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & -\frac{1}{2} & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{5}{3} & \frac{4}{3} & 1 & \frac{2}{3} & \frac{1}{3} \\ \frac{4}{3} & \frac{8}{3} & 2 & \frac{4}{3} & \frac{2}{3} \\ 1 & 2 & 3 & 2 & 1 \\ \frac{2}{3} & \frac{4}{3} & 2 & \frac{8}{3} & \frac{4}{3} \\ \frac{1}{3} & \frac{2}{3} & 1 & \frac{4}{3} & \frac{5}{3} \end{pmatrix},
$$

from which we learn, e.g., that $\mathbb{E}[N_6 | X_0 = 3] = r_{36} = \frac{2}{3}$, or that the expected number of flips required to end the game, when starting from state 3, is

$$
\sum_{j=2}^{6} r_{3j} = 8.
$$

An example

The computation of F is immediate, but for the blocks F_{tt} and F_{tr}. The latter, in this simple case, is given by $F_{tr} = G = R_{tt}B = R_{tt}P_{tr}$

$$= \begin{pmatrix} \frac{5}{3} & \frac{4}{3} & 1 & \frac{2}{3} & \frac{1}{3} \\ \frac{4}{3} & \frac{8}{3} & 2 & \frac{4}{3} & \frac{2}{3} \\ 1 & 2 & 3 & 2 & 1 \\ \frac{2}{3} & \frac{4}{3} & 2 & \frac{8}{3} & \frac{4}{3} \\ \frac{1}{3} & \frac{2}{3} & 1 & \frac{4}{3} & \frac{5}{3} \end{pmatrix} \begin{pmatrix} \frac{1}{2} & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & 0 \\ 0 & \frac{1}{2} \end{pmatrix} = \begin{pmatrix} \frac{5}{6} & \frac{1}{6} \\ \frac{2}{3} & \frac{1}{3} \\ \frac{1}{2} & \frac{1}{2} \\ \frac{1}{3} & \frac{2}{3} \\ \frac{1}{6} & \frac{5}{6} \end{pmatrix},$$

from which we learn, e.g., that the probability A loses the game, when he starts with 2$ (=state 3), is

$$f_{30} = \frac{2}{3}.$$

An example

Remarks:

- These results were previously obtained, in the chapter about martingales, by using the optional stopping theorem.

- It should be noted however that the methodology developed in this chapter applies to arbitrary graph structures...

Outline of the course

1. A short introduction.
2. Basic probability review.
3. Martingales.
4. Markov chains.
    4.1. Definitions and examples.
    4.2. Strong Markov property, number of visits.
    4.3. Classification of states.
    4.4. Computation of R and F.
    4.5. Asymptotic behavior.
5. Markov processes, Poisson processes.
6. Brownian motions.

Asymptotic behavior: an example

Let $0 \leq p, q \leq 1$ (with $0 < p + q < 2$) and consider the chain
We are interested in $a^{(n)} = (\mathbb{P}[X_n = 0], \mathbb{P}[X_n = 1])$ for large n.
We have $a^{(n)} = a^{(0)} P^n$ and

$$a^{(0)} P^n = (\xi, 1 - \xi) \left[\frac{1}{p+q} \begin{pmatrix} q & p \\ q & p \end{pmatrix} + \frac{(1-p-q)^n}{p+q} \begin{pmatrix} p & -p \\ -q & q \end{pmatrix} \right],$$

so that

$$\lim_{n \to \infty} a^{(n)} = (\xi, 1 - \xi) \frac{1}{p+q} \begin{pmatrix} q & p \\ q & p \end{pmatrix} = \left(\frac{q}{p+q}, \frac{p}{p+q} \right),$$

which does not depend on $a^{(0)}$ (not so amazing! Why?)

Asymptotic behavior

Let (X_n) be a HMC with transition matrix P.

Definition: (X_n) admits a limiting distribution $\Leftrightarrow$

- $\exists \pi$ such that $\lim_{n\to\infty} a^{(n)} = \pi$,
- $\pi_j \geq 0$ for all j and $\pi 1 = \sum_j \pi_j = 1$,
- π does not depend on $a^{(0)}$.

Remarks:

- π is called the limiting distribution.
- The existence of π does only depend on P.
- Not every HMC does admit some limiting distribution:

Asymptotic behavior

Consider the chain
We have

$$a^{(n)} = a^{(0)}P^n = (\xi, 1-\xi)\left[\frac{1}{2}\begin{pmatrix} 1 & 1 \\ 1 & 1 \end{pmatrix} + \frac{(-1)^n}{2}\begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\right]$$
$$= \ldots = \Big(\frac{1}{2} + (-1)^n(\xi - \frac{1}{2}), \frac{1}{2} + (-1)^{n+1}(\xi - \frac{1}{2})\Big),$$

which does only converge for $\xi = \frac{1}{2}$. Hence, this HMC does not admit a limiting distribution...

Asymptotic behavior

How to determine the limiting distribution (if it exists)?

$\rightsquigarrow$ **Theorem 1**: *assume the HMC is (i) irreducible (that is, contains only one class) and (ii) non-periodic. Then all states are positive-recurrent* $\Leftrightarrow$ *The system of equations*

$$\begin{cases} xP = x \\ x1 = 1 \end{cases}$$

has a nonnegative solution (and, in that case, $x = \pi$ *is the limiting distribution).*

Remark: π is also called the stationary (or invariant distribution). This terminology is explained by the fact that if one takes $a^{(0)} = \pi$, then
$a^{(n)} = a^{(0)}P^n = a^{(0)}P^{n-1} = a^{(0)}P^{n-2} = \ldots = a^{(0)}P = a^{(0)}$ for all n.

Asymptotic behavior

How to determine the limiting distribution (if it exists)?

$\rightsquigarrow$ **Theorem 2**: *assume the HMC has a finite state space and that P is regular (that is, $\exists n$ such that $(P^n)_{ij} > 0$ for all i, j). Then it admits a limiting distribution, which is given by the solution of*

$$\begin{cases} xP = x \\ x1 = 1. \end{cases}$$

$\rightsquigarrow$ **Theorem 3**: *assume the eigenvalue 1 of P has multiplicity 1 and that all other eigenvalues $\lambda_j (\in \mathbb{C})$ satisfy $|\lambda_j| < 1$. Then the conclusion of Theorem 2 holds.*

Asymptotic behavior

A simple (artificial) example...

Consider the chain with transition matrix

$$P = \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ 0 & 1 \end{pmatrix}.$$

Clearly, Theorem 2 does not apply, but Theorem 3 does.
The limiting distribution is given by

$$(\pi_0, \pi_1) \begin{pmatrix} \frac{3}{4} & \frac{1}{4} \\ 0 & 1 \end{pmatrix} = (\pi_0, \pi_1), \quad \pi_0 + \pi_1 = 1, \quad \pi_0 \geq 0, \quad \pi_1 \geq 0,$$

which yields $\pi = (\pi_0, \pi_1) = (0, 1)$... which is not very surprising.