1 Session 16. Markov chains

Exercise 1

Here is a simple game:

- a player bets on coin tosses, a dollar each time,
- the game ends either when the player has no money left or is up to five dollars.

If the player starts with three dollars, what is the chance that the game takes at least five flips? Twenty-five flips?

At any point, this player has either \$0, or \$1, ..., or \$5. We say that the player is in the state s_0, s_1, ..., or s_5. A game consists of moving from state to state. For instance, a player now in state s_3 has on the next flip a .5 chance of moving to state s_2 and a .5 chance of moving to s_4. The boundary states are a bit different; once in state s_0 or state s_5, the player never leaves.

Let $p_i(n)$ be the probability that the player is in state s_i after n flips. Then, for instance, we have that the probability of being in state s_0 after flip $n+1$ is $p_0(n+1) = p_0(n) + 0.5 \cdot p_1(n)$. This matrix equation sumarizes.

$$\begin{pmatrix} 1 & .5 & 0 & 0 & 0 & 0 \\ 0 & 0 & .5 & 0 & 0 & 0 \\ 0 & .5 & 0 & .5 & 0 & 0 \\ 0 & 0 & .5 & 0 & .5 & 0 \\ 0 & 0 & 0 & .5 & 0 & 0 \\ 0 & 0 & 0 & 0 & .5 & 1 \end{pmatrix} \begin{pmatrix} p_0(n) \\ p_1(n) \\ p_2(n) \\ p_3(n) \\ p_4(n) \\ p_5(n) \end{pmatrix} = \begin{pmatrix} p_0(n+1) \\ p_1(n+1) \\ p_2(n+1) \\ p_3(n+1) \\ p_4(n+1) \\ p_5(n+1) \end{pmatrix}$$

With the initial condition that the player starts with three dollars, calculation gives this.

$n=0$	$n=1$	$n=2$	$n=3$	$n=4$	$\cdots$	$n=24$
0	0	0	.125	.125		.39600
0	0	.25	0	.1875		.00276
0	.5	0	.375	0	$\cdots$	0
1	0	.5	0	.3125		.00447
0	.5	0	.25	0		0
0	0	.25	.25	.375		.59676

- The player quickly ends in either state s_0 or state s_5. For instance, after the fourth flip there is a probability of 0.50 that the game is already over: The boundary states are said to be **absorbtive**.
- *The Markov chain model is historyless*: with a fixed transition matrix, the next state depends only on the current state, not on any prior states. More formally:

$$P(x_{t+1} = s_{t+1} | x_t = s_t, x_{t-1} = s_{t-1}, ..., x_1 = s_1, x_0 = s_0) = P(x_{t+1} = s_{t+1} | x_t = s_t)$$

- In addition, it is clear that the probability of a transition from state i to state j is independent of time (**stationarity property**):

$$P(x_{t+1} = j | x_t = i) = p_{ij}$$

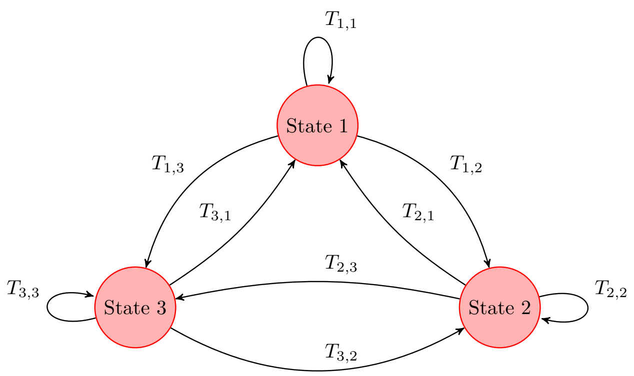
- Each column vector is a probability vector (and it sums up 1) and the matrix is a transition matrix.

Exercise 2

A study ([1], p. 202) divided occupations in the United Kingdom into upper level (executives and professionals), middle level (supervisors and skilled manual workers), and lower level (unskilled). To determine the mobility across these levels in a generation, about two thousand men were asked, "At which level are you, and at which level was your father when you were fourteen years old?"

This equation and transition diagram summarizes the results.

$$\begin{pmatrix} T_{1,1} & T_{1,2} & T_{1,3} \\ T_{2,1} & T_{2,2} & T_{2,3} \\ T_{3,1} & T_{3,2} & T_{3,3} \end{pmatrix}^t \rightarrow \begin{pmatrix} .60 & .29 & .16 \\ .26 & .37 & .27 \\ .14 & .34 & .57 \end{pmatrix} \begin{pmatrix} p_U(n) \\ p_M(n) \\ p_L(n) \end{pmatrix} = \begin{pmatrix} p_U(n+1) \\ p_M(n+1) \\ p_L(n+1) \end{pmatrix}$$



For instance, a child of a lower class worker has a .27 probability of growing up to be middle class. Notice that the Markov model assumption about history seems reasonable– we expect that while a parent's occupation has a direct influence on the occupation of the child, the grandparent's occupation has no such direct influence. With the initial distribution of the respondents's fathers given below, this table lists the distributions for the next five generations.

$n=0$	$n=1$	$n=2$	$n=3$	$n=4$	$n=5$
$\begin{pmatrix} .12 \\ .32 \\ .56 \end{pmatrix}$	$\begin{pmatrix} .23 \\ .34 \\ .42 \end{pmatrix}$	$\begin{pmatrix} .29 \\ .34 \\ .37 \end{pmatrix}$	$\begin{pmatrix} .31 \\ .34 \\ .35 \end{pmatrix}$	$\begin{pmatrix} .32 \\ .33 \\ .34 \end{pmatrix}$	$\begin{pmatrix} .33 \\ .33 \\ .34 \end{pmatrix}$

Exercise 3

The World Series of American baseball is played between the team winning the American League and the team winning the National League. The series is won by the first team to win four games. That means that a series is in one of twenty-four states: 0-0 (no games won yet by either team), 1-0 (one game won for the American League team and no games for the National League team), etc. If we assume that there is a probability p that the American League team wins each game then we have the following transition matrix.

$$\begin{pmatrix} 0 & 0 & 0 & 0 & \ldots \\ p & 0 & 0 & 0 & \ldots \\ 1-p & 0 & 0 & 0 & \ldots \\ 0 & p & 0 & 0 & \ldots \\ 0 & 1-p & p & 0 & \ldots \\ 0 & 0 & 1-p & 0 & \ldots \\ \vdots & \vdots & \vdots & \vdots & \end{pmatrix} \begin{pmatrix} p_{\text{0-0}}(n) \\ p_{\text{1-0}}(n) \\ p_{\text{0-1}}(n) \\ p_{\text{2-0}}(n) \\ p_{\text{1-1}}(n) \\ p_{\text{0-2}}(n) \\ \vdots \end{pmatrix} = \begin{pmatrix} p_{\text{0-0}}(n+1) \\ p_{\text{1-0}}(n+1) \\ p_{\text{0-1}}(n+1) \\ p_{\text{2-0}}(n+1) \\ p_{\text{1-1}}(n+1) \\ p_{\text{0-2}}(n+1) \\ \vdots \end{pmatrix}$$

An especially interesting special case is $p = 0.50$; this table lists the resulting components of the $n = 0$ through $n = 7$ vectors. Check what happens with other values of p.

	$n = 0$	$n = 1$	$n = 2$	$n = 3$	$n = 4$	$n = 5$	$n = 6$	$n = 7$
$0-0$	1	0	0	0	0	0	0	0
$1-0$	0	0.5	0	0	0	0	0	0
$0-1$	0	0.5	0	0	0	0	0	0
$2-0$	0	0	0.25	0	0	0	0	0
$1-1$	0	0	0.5	0	0	0	0	0
$0-2$	0	0	0.25	0	0	0	0	0
$3-0$	0	0	0	0.125	0	0	0	0
$2-1$	0	0	0	0.375	0	0	0	0
$1-2$	0	0	0	0.375	0	0	0	0
$0-3$	0	0	0	0.125	0	0	0	0
$4-0$	0	0	0	0	0.0625	0.0625	0.0625	0.062
$3-1$	0	0	0	0	0.25	0	0	0
$2-2$	0	0	0	0	0.375	0	0	0
$1-3$	0	0	0	0	0.25	0	0	0
$0-4$	0	0	0	0	0.0625	0.0625	0.0625	0.062
$4-1$	0	0	0	0	0	0.125	0.125	0.125
$3-2$	0	0	0	0	0	0.3125	0	0
$2-3$	0	0	0	0	0	0.3125	0	0
$1-4$	0	0	0	0	0	0.125	0.125	0.125
$4-2$	0	0	0	0	0	0	0.15625	0.156
$3-3$	0	0	0	0	0	0	0.3125	0
$2-4$	0	0	0	0	0	0	0.15625	0.156
$4-3$	0	0	0	0	0	0	0	0.156
$3-4$	0	0	0	0	0	0	0	0.156

Evenly-matched teams are likely to have a long series– there is a probability of 0.625 that the series goes at least six games.

In summary, a system can be modeled as a Markov process if it has the following four properties:

- Property 1: A finite number of states can be used to describe the dynamic behavior of the system.

- Property 2: Initial probabilities are specified for the system.

- Property 3: Markov property—We assume that a transition to a new state depends only on the current state and not on past conditions.
- Property 4: Stationarity property—The probability of a transition between any two states does not vary in time.

If Markov analysis is appropriate for the system being studiedone can try answering questions like:

- How many transitions (steps) will it likely take for the system to move from some specified state to another specified state?
- What is the probability that it will take some given number of steps to go from one specified state to another?
- In the long run, which state is occupied by the system most frequently?
- Over a long period of time, what fraction of the time does the system occupy each of the possible states?
- Will the system continue indefinitely to move among all N states, or will it eventually settle into a certain few states?

[1]

References

[1] Kenneth Macdonald and John Ridge. Social Mobility. In A. H. Halsey, editor, *British Social Trends since 1900: A Guide to the Changing Social Structure of Britain*, pages 202–225. Palgrave Macmillan UK, London, 1988.

[1]dynamic programming: https://www.youtube.com/watch?v=jqyfhFziDjw
https://www.techiedelight.com/word-break-problem/