Markov Chains, Roulette and Google Search[1]

Henk Tijms, 15 December 2021[2]

Abstract The goal of this tutorial is to give a first introduction to Markov chain methods. This powerful probability tool in operations research will be illustrated with the game of Egg Russian Roulette and Google's Page-Rank search algorithm among others.

1. Introduction

A gem for teaching Markov chains to beginning students is the game of Egg Russian Roulette. This game was played for several years in The Tonight Show with Jimmy Fallon. In this show Jimmy plays the Egg Russian Roulette game with a guest of the show. The guest was always a celebrity from sports or film: Tom Cruise, Anna Kendrick, Jodie Foster, David Beckham, to name a few. The guest and Jimmy take turns picking an egg from a carton and smashing it on their heads. The carton contains a dozen eggs, four of which are raw and the rest are boiled. Neither Jimmy nor the guest knows which eggs are raw and which are boiled. The first person who has cracked two raw eggs on their head loses the game.

The entertainment value of seeing famous people with raw yolk and albumin running down their hair and faces made the game very popular. Incidentally, the origin of the game has a rich history, dating back to the Middle Ages. In the rural English hamlet of Swaton (184 inhabitants, currently), the throwing of eggs started around 1322 when the new abbot of the town, who owned all of the poultry, handed out eggs to loyal churchgoers as alms. Whenever the church was cut-off from the rest of the hamlet by the sometimes overflowing local river, the eggs were chucked to the churchgoers waiting on the other side of this watercourse. Recently, this tradition has been slightly adapted and restored: every year since 2006, this little village hosts a world championship of Russian Egg Roulette, which attracts contestants from all over the world.

Let's go back to Jimmy Fallon's Tonight Show. In the show, the guest is the first to choose an egg. Do you think each player has the same probability of losing the game? Does the guest of the show has an advantage because there are more hard boiled eggs to select from at the start? To answer these

[1]This tutorial is largely based on the book Henk Tijms, *Basic Probability, What Every Math Student Should Know*, second edition, World Scientific Press, 2021.

[2]The author is emeritus professor of operations research at the Vrije University in Amsterdam, email: h.c.tijms@xs4all.nl.

questions, the method of absorbing Markov chains will be used. This method essentially boils down to the use of conditional probabilities and matrix calculations. In some intermediate situations arising during the course of the game the probability of the guest of losing the game can be calculated by a simple argument. Suppose that Jimmy and the guest have each smashed one raw egg on their heads. Then, for $i = 2, \ldots, 10$, let g_i be the probability of the guest losing the game when i eggs are left in the carbon. In other words, g_i is the probability that the guest will pick as first a second raw egg when $i-2$ boiled eggs and 2 raw eggs are left in the carbon. If i is even, the next egg will be picked by the guest; otherwise the host Jimmy picks the next egg. Therefore

$$g_2 = 1.$$

The other g_i can be recursively computed from

$$g_i = \frac{i-2}{i} \times g_{i-1} \quad \text{for } i = 3, 5, 7, \text{ and } 9$$
$$g_i = \frac{2}{i} + \frac{i-2}{i} \times g_{i-1} \quad \text{for } i = 4, 6, 8, \text{ and } 10.$$

These equations are easily explained. For i odd, the guest loses the game only if Jimmy picks a boiled egg from the i eggs left in the carbon and the guest picks as first a raw egg from the remaining $i-1$ eggs. The joint probability of these two events is $\frac{i-2}{i}$ multiplied by g_{i-1}. For i even, the probability of the guest losing the game is the sum of the probability of the event that the guest picks directly a raw egg from the carton with i eggs left and the probability of the event that the guest picks a boiled egg from the carbon with i eggs left and loses in the remainder of the game with $i-1$ eggs left in the carbon. The first event has probability $\frac{2}{i}$ and the second event has probability $\frac{i-2}{i} \times g_{i-1}$. Starting the recursive calculation with $g_2 = 1$, you obtain from the recursion equations

$$g_2 = 1,\ g_3 = \frac{1}{3},\ g_4 = \frac{2}{3},\ g_5 = \frac{2}{5},\ g_6 = \frac{3}{5},\ g_7 = \frac{3}{7},\ g_8 = \frac{4}{7},\ g_9 = \frac{4}{9},\ g_{10} = \frac{5}{9}.$$

In order to solve the general problem of Egg Russian Roulette, it is helpful to use Markov chains. This will be the subject of the next section.

2. A Primer on Markov chains

Markov chains represent the most important stochastic process in probability. This section gives a first impression of the fascinating world of Markov chains. This branch of probability was founded by the Russian mathematician A.A. Markov (1856–1922) at the beginning of the 20th century.[3] Before Markov the theory of probability involved observing a series of events that were independent of each other. The classic examples are coin-flipping and dice-rolling. Markov added the idea of interdependence to probability, the notion that what happens next is linked to what is happening now. In his famous 1913 lecture for the Imperial Academy of Sciences in St. Petersburg, Markov used this notion to analyze the frequencies at which vowels and consonants occur in Pushkin's novel "Eugene Onegin." Markov's model is a very versatile probability model that is used today in countless applications in many different areas, such as voice recognition, DNA analysis, stock control, telecommunications and a host of others. Markov chains are everywhere in science today.

A Markov chain can be seen as a dynamic stochastic process that randomly moves from state to state with the property that only the current state is relevant for the next state. In other words, the memory of the process goes back only to the most recent state. A picturesque illustration of this would show the image of a frog jumping from lily pad to lily pad with appropriate transition probabilities that depend only on the position of the last lily pad visited. In order to plug a specific problem into a Markov chain model, the state variable(s) should be appropriately chosen in order to ensure the characteristic memoryless property of the process. The basic steps of the modeling approach are:

- Choosing the state variable(s) such that the current state summarizes everything about the past that is relevant to the future states.
- The specification of the one-step transition probabilities of moving from state to state in a single step.

Using the concept of state and choosing the state in an appropriate way, surprisingly many probability problems can be solved within the framework

[3]Markov lived through a period of great political activity in Russia and, having firm opinions, he became heavily involved. Maksim Gorky, the Russian short-story writer, novelist and left wing activist, was elected a member of the Russian Academy of Sciences in 1902, but his election was soon withdrawn for political reasons on the Tsar's orders. Markov protested strongly and refused to accept honours awarded him on the following year. In June 1907 Tsar Nicholas dissolved the Second Duma which had been elected with majority on the left. Markov repudiated his membership and might have expected to suffer severe consequences but the authorities chose not to make an example of an elderly and distinguished mathematician. In 1913 the Romanov dynasty, which had been in power in Russia since 1613, celebrated their 300 years of power. This was not likely to improve their already weak position. Markov showed his disapproval of the celebration but holding celebrations of his own - he celebrated 200 years of the Law of Large Numbers!

of a Markov chain. The set of states is denoted by I and is assumed to be *finite*. The one-step transition probabilities are denoted by:

$$p_{ij} = \text{the probability of going from state } i \text{ to state } j \text{ in one step}$$

for $i, j \in I$. The one-step probabilities must satisfy

$$p_{ij} \geq 0 \text{ for all } i, j \in I \quad \text{and} \quad \sum_{j \in I} p_{ij} = 1 \text{ for all } i \in I.$$

It is convenient to summarize the one-step transition probabilities in a matrix $\mathbf{P}$ having p_{ij} as its (i, j)th element.

In Markov chains a key role is played by the n-step transition probabilities. For any $n = 1, 2, \ldots$, these probabilities are defined as

$$p_{ij}^{(n)} = \text{the probability of going from state } i \text{ to state } j \text{ in } n \text{ steps}$$

for all $i, j \in I$. Note that $p_{ij}^{(1)} = p_{ij}$. How to calculate the n-step transition probabilities? It will be seen that they can be calculated by matrix products. This key fact is based on the so-called Chapman–Kolmogorov equations

$$\boxed{p_{ij}^{(n)} = \sum_{k \in I} p_{ik}^{(n-1)} p_{kj} \quad \text{for all } i, j \in I \text{ and } n = 2, 3, \ldots .}$$

This recurrence relation can be seen by noting that the probability of going from state i to state j in n steps is obtained by summing the probabilities of the mutually exclusive events of going from state i to some state k in the first $n-1$ steps and then going from state k to state j in the nth step.

An extremely useful observation is that the n-step transition probabilities $p_{ij}^{(n)}$ can be calculated by multiplying the matrix $\mathbf{P}$ of one-step transition probabilities by itself n times. Let's verify this for $n = 2$. Then,

$$p_{ij}^{(2)} = \sum_{k \in I} p_{ik} p_{kj} \quad \text{for all } i, j \in I.$$

This is the definition for the elements of the matrix product $\mathbf{P} \times \mathbf{P} = \mathbf{P}^2$. The argument can be extended to conclude that for any $n \geq 1$ and $i, j \in I$:

$$\boxed{p_{ij}^{(n)} = \text{the } (i, j)\text{th element of the } n\text{-fold matrix product } \mathbf{P}^n.}$$

This is a very important conclusion: many computations for finite-state Markov chains can be boiled down to matrix calculations!

Absorbing Markov chains

Many applied probability problems can be analyzed by using an absorbing Markov chain with an appropriate choice of the state variable(s). A Markov chain is said to be *absorbing* if there are one or more states i with $p_{ii} = 1$ and thus $p_{ij} = 0$ for $j \neq i$. That is, once the process is in an absorbing state it always stays there. We give two examples of absorbing Markov chains before we apply this concept to solve the Egg Russian Roulette problem.

Desperate Joe goes for roulette

An instructive example of an absorbing Markov chain is the following random-walk type of problem. Joe Dalton desperately wants to raise his bankroll of \$600 to \$1,000 in order to pay his debts before midnight; otherwise he will get into big trouble with a loan shark. He enters a casino to play European roulette. He decides to bet on red each time using bold play, that is, Joe bets either his entire bankroll or the amount needed to reach the target bankroll, whichever is smaller. Thus the stake is \$200 if his bankroll is \$200 or \$800 and the stake is \$400 if his bankroll is \$400 or \$600. In European roulette a bet on red is won with probability $\frac{18}{37}$ and is lost with probability $\frac{19}{37}$. What is the probability that Joe will reach his goal?

To solve Joe's problem, take a Markov chain with six states $i = 0, 1, \ldots, 5$, where state i means that Joe's bankroll is $i \times 200$ dollars. The states 0 and 5 are absorbing and the game is over as soon one of these states is reached. Thus $p_{00} = p_{55} = 1$. The other p_{ij} are easily found. For example, the only possible one-step transitions from state $i = 2$ are to the states 0 and 4, because Joe bets \$400 in state 2. Thus $p_{20} = \frac{19}{37}$ and $p_{24} = \frac{18}{37}$. The other p_{ij} are given in the following matrix $\mathbf{P}$ of one-step transition probabilities:

$$
\begin{array}{ll}
\text{from}\backslash\text{to} & \begin{array}{cccccc} 0 & 1 & 2 & 3 & 4 & 5 \end{array} \\
\begin{array}{l} 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array} &
\begin{pmatrix}
1 & 0 & 0 & 0 & 0 & 0 \\
\frac{19}{37} & 0 & \frac{18}{37} & 0 & 0 & 0 \\
\frac{19}{37} & 0 & 0 & 0 & \frac{18}{37} & 0 \\
0 & \frac{19}{37} & 0 & 0 & 0 & \frac{18}{37} \\
0 & 0 & 0 & \frac{19}{37} & 0 & \frac{18}{37} \\
0 & 0 & 0 & 0 & 0 & 1
\end{pmatrix}
\end{array}.
$$

For any starting state, the process will ultimately absorbed in either state 0 or state 5. The absorption probabilities can be obtained by calculating $\mathbf{P}^n$ for n sufficiently large. Trying several values of n, it was found that $n = 20$

is large enough to have convergence of all $p_{ij}^{(n)}$ in four or more decimals:

$$\mathbf{P}^{20} = \mathbf{P}^{21} = \ldots = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 0.8141 & 0 & 0 & 0. & 0 & 0.1859 \\ 0.6180 & 0 & 0 & 0 & 0 & 0.3820 \\ 0.4181 & 0 & 0 & 0 & 0 & 0.5819 \\ 0.2147 & 0 & 0 & 0 & 0 & 0.7853 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

You read off from row 4 that the probability of Joe reaching his goal when starting with \$600 is equal to

$$p_{35}^{(20)} = p_{35}^{(21)} = \ldots = 0.5819.$$

This probability is the maximum probability for Joe of reaching the goal of getting \$1000. The intuitive explanation that bold play is optimal in Joe's situation is that the shorter Joe exposes his bankroll to the casino's house advantage, the better it is (e.g., if Joe bets \$50 each time, he reaches his goal with probability 0.4687, and with probability 0.2917 if each bet is for \$20).

Alternatively, the win probability 0.5819 can be calculated by solving four linear equations. To do so, define f_i as the probability of ever getting absorbed in state 5 when the starting state is i. By definition, $f_0 = 0$ and $f_5 = 1$. By conditioning on the next state after state i and using the law of conditional probability, you get the four linear equations

$$f_1 = \frac{19}{37} \times 0 + \frac{18}{37} f_2, \ f_2 = \frac{19}{37} \times 0 + \frac{18}{37} f_4,$$
$$f_3 = \frac{19}{37} f_1 + \frac{18}{37} \times 1, \ f_4 = \frac{19}{37} f_3 + \frac{18}{37} \times 1.$$

The solution of these equations is $f_1 = 0.1859$, $f_2 = 0.3820$, $f_3 = 0.5819$, and $f_4 = 0.7853$. The same solution as found by matrix multiplication.

A similar set of linear equations can be used to calculate $E(N)$, where the random variable N is the number of bets made by Joe. To do so, parametrize again and define e_i as the expected value of the number of bets when the starting state is i. By definition, $e_0 = e_5 = 0$. By conditioning on the next state after state i and using the law of conditional expectation, you get

$$e_1 = 1 + \frac{19}{37} \times e_0 + \frac{18}{37} \times e_2, \ e_2 = 1 + \frac{19}{37} \times e_0 + \frac{18}{37} \times e_4,$$
$$e_3 = 1 + \frac{19}{37} \times e_1 + \frac{18}{37} \times e_5, \ e_4 = 1 + \frac{19}{37} \times e_3 + \frac{18}{37} \times e_5.$$

The solution of these linear equations is $e_1 = 1.9675$, $e_2 = 1.9887$, $e_3 = 2.0323$, and $e_4 = 2.0103$. Thus $E(N) = 2.0323$.

Another nice example of an absorbing Markov chain is provided by the episode *Glass Stepping Stones* in the 2021 Netflix series of *Squid Game*.

The deadly glass bridge game

The squid game is one of the most watched series in Netflix. In an insanely sadistic rat race − based on old-fashioned children's games − a few hundred downtrodden people are given the chance to still make something of their lives. The blood-curdling episode 'glass stepping stones' of the originally Korean series is about 16 players crossing a floating bridge of 18 steps, see https://www.youtube.com/watch?v=19oFNytu0z0. For each step, the player has a choice to pick the left pane or the right pane. One of these two panes is of tempered glass, capable of supporting the weight of a person, and the other of normal glass, which breaks when stepped on. It is impossible to see the difference between the panes of tempered glass and normal glass, which are randomly assigned to the steps. Sixteen players will try to cross the bridge one after the other, choosing one of the two panes at each step. The bad news is that if a player jumps onto a panel with normal glass, the glass breaks and the player tumbles down, resulting in death. The good news is that the sacrifice was not in vain, because the broken panel gives all remaining players valuable information about what the right path to safety is. Furthermore, it is assumed that each player also knows the safe panels chosen by previous players. In sequence, each player attempts to cross the bridge and keeps moving until the player has either successfully crossed all 18 steps on the bridge or has tumbled down in between. What is the expected number of survivors, what is the probability of survival for each player, and what is the probability distribution of the number of survivors?

To answer these questions, consider an absorbing Markov with 20 states $i = 0, 1, \ldots, 18, 19$. State i with $1 \leq i \leq 18$ means that the game has moved forward to step i where a player jumped on normal glass and was eliminated, state 19 means that a player has safely reached the final step 18, and state 0 is an auxiliary state corresponding to the beginning of the game. State 19 is the absorbing state of the Markov chain, that is, $p_{19,19} = 1$. For $i = 0, 1, \ldots, 18$, the one-step transition probabilities are

$$p_{ij} = \left(\frac{1}{2}\right)^{j-i} \text{ for } j = i+1, .., 18 \quad \text{and} \quad p_{i,19} = \left(\frac{1}{2}\right)^{18-i}.$$

The other p_{ij} are 0. You calculate the matrix products $\mathbf{P}^k$ for $k = 1, \ldots, 16$. Let a_k be the probability that player k survives and d_k be the probability that exactly k players survive for $k = 0, 1, \ldots, 16$. Then

$$a_k = p_{0,19}^{(k)} \quad \text{for } k = 1, 2, \ldots, 16,$$

and so a_k is given by the $(0, 19)$th element of $\mathbf{P}^k$. The d_j's can next be calculated from

$$d_{16-k+1} = a_k - a_{k-1} \quad \text{for } k = 1, 2, \ldots, 16,$$

where $a_0 = 0$. The explanation is that a_k also gives the probability that $16-k+1$ or more players survive, because each player knows what panes were chosen by the previous players and so, if player k crosses safely the bridge, any player behind player k crosses safely the bridge. Thus the probability that the total number of survivors is exactly equal to $16 - k + 1$ is obtained by subtracting the probability a_{k-1} from the probability a_k. Of course, the probability d_0 is $1 - \sum_{k=1}^{16} d_k$. The expected value of the number of survivors can be calculated as $\sum_{k=1}^{16} k\, d_k$. The matrix calculations lead to

$$\begin{aligned}
&a_1 = 0.000, a_2 = 0.000, a_3 = 0.001, a_4 = 0.004, a_5 = 0.015, a_6 = 0.048,\\
&a_7 = 0.119, a_8 = 0.240, a_9 = 0.407, a_{10} = 0.593, a_{11} = 0.760,\\
&a_{12} = 0.881, a_{13} = 0.952, a_{14} = 0.985, a_{15} = 0.996, a_{16} = 0.999.
\end{aligned}$$

The expected value of the number of survivors is 7.000076 , a value very close to 7 (in the Netflix episode the actual number of survivors was 3, a remarkably small number of survivors in the light of $d_0 + d_1 + d_2 + d_3 = 0.047$). The value 7 is obtained by the following heuristic argument: if the number of steps is large enough, the number of panes 'unlocked' by a player is approximately geometrically distributed with parameter $\frac{1}{2}$ and expected value 2, and this makes it plausible that on average about 9 players have to be sacrificed in order for the remaining $16 - 9 = 7$ players to cross the bridge safely.

A true slaughter of the players would have been the case if the players could only have seen the broken panels and had no further information. Simulation seems the only practical approach for the probabilistic analysis of this variant. One hundred thousand simulation runs give the estimate 0.24 for the expected number of survivors and the estimate 0.11 for the probability of the last player surviving.[4]

3. Markov chain analysis for Egg Russian Roulette

An absorbing Markov chain is used to analyze the game of Egg Russian Roulette. The state of the Markov chain is described by the triple (i, r_1, r_2), where i denotes the number of smashed eggs, r_1 is the number of raw eggs

[4]In support of this result, Ad Ridder pointed out to me: the probability that the first 15 players do not survive is $(1 - (\frac{1}{2})^{18}) \times (1 - (\frac{1}{2})^{17}) \times \cdots \times (1 - (\frac{1}{2})^4) = 0.8801$, and so a lower bound for $P(\text{last player survives})$ is $(\frac{1}{2})^3 \times 0.8801 = 0.110$. In the same way, a lower bound for $P(\text{second last player survives})$ is $(\frac{1}{2})^4 \times 0.9388 = 0.059$, etc.

picked by the guest and r_2 is the number of raw eggs picked by the host of the game. The states satisfy $0 \leq i \leq 11$ and $r_1 + r_2 \leq 3$. The process starts in state $(0,0,0)$ and ends when one of the absorbing states $(i,2,0)$, $(i,2,1)$, $(i,0,2)$, or $(i,1,2)$ is reached. The guest loses the game if the game ends in a state $(i,2,0)$ or $(i,2,1)$ with i odd. In a non-absorbing state (i,r_1,r_2) with i even, the guest picks an egg and the process goes either to state $(i+1,r_1+1,r_2)$ with probability $\frac{4-r_1-r_2}{12-i}$ or to state $(i+1,r_1,r_2)$ with probability $1-\frac{4-r_1-r_2}{12-i}$. In a non-absorbing state (i,r_1,r_2) with i odd, the host picks an egg and the process goes either to state $(i+1,r_1,r_2+1)$ with probability $\frac{4-r_1-r_2}{12-i}$ or to state $(i+1,r_1,r_2)$ with probability $1-\frac{4-r_1-r_2}{12-i}$. This sets the matrix $\mathbf{P}$ of one-step transition probabilities. The probability that the guest will lose can be computed by calculating $\mathbf{P}^{11}$. This requires that the states are ordered in a one-dimensional array. It is easier to use a recursion to calculate the probability of the guest losing the game. For that, you reason in the same way as in the above gambling problem. For any state (i,r_1,r_2), let $p(i,r_1,r_2)$ be the probability that the guest will lose if the process starts in state (i,r_1,r_2). The sought probability $p(0,0,0)$. The sought probability can be calculated by a recursion with the boundary conditions can be calculated by a recursion with the boundary conditions $p(i,2,0) = p(i,2,1) = 1$ and $p(i+1,0,2) = p(i+1,1,2) = 0$ for $i = 3$, 5, 7, 9 and 11. The recursive calculations are

$$p(i,r_1,r_2) = \frac{4-r_1-r_2}{12-i} p(i+1,r_1+1,r_2) + \Big(1 - \frac{4-r_1-r_2}{12-i}\Big) p(i+1,r_1,r_2)$$

for $i = 0$, 2, 4, 6, 8 and 10, and

$$p(i,r_1,r_2) = \frac{4-r_1-r_2}{12-i} p(i+1,r_1,r_2+1) + \Big(1 - \frac{4-r_1-r_2}{12-i}\Big) p(i+1,r_1,r_2)$$

for $i = 1$, 3, 5, 7, 9 and 11. The recursive computations lead to the value $\frac{5}{9}$ for the probability that the guest of the show will lose the game. Interestingly enough, the game turns out to be fair for the case of three raw eggs and nine boiled eggs. For the case of five raw eggs and seven boiled eggs, the guest will lose with probability 0.563. Similar recursive computations give that the expected number of trials has the values 8.41, 6.86 and 5.73 for the respective cases of three, four and five raw eggs. All of these results might also be verified by computer simulation – a Python program is easily written. In fact, simulations of the problem are provided by the videos online of episodes of Egg Russian Roulette in The Tonight Show by Jimmy Fallon, with Higgins as unsurpassed sidekick with his characteristically shrill voice, reminiscent of the character Igor from the parody movie Young Frankenstein. In the 18 episodes I found on Internet the guest lost 9 times the game. Remarkably, the experimental probability of 50% resulting from this very small sample size is not far away from the theoretical probability of 55.6%.

4. Long-run behavior of Markov chains

Let's consider a Markov chain with no absorbing states. What about the probability distribution of the state after many, many transitions? Does the effect of the starting state ultimately fade away? To answer these questions, the following assumption is made:

(**a**) There is a state s that can be reached from any other state, that is, for any state j there is some $n \geq 1$ such that $p_{js}^{(n)} > 0$.
(**b**) The set of states cannot be split into multiple disjoint sets $S_1, \ldots, S_d$ with $d \geq 2$ such that a one-step transition from a state in S_k is always to a state in S_{k+1}, where $S_{d+1} = S_1$.

The condition (**a**) is satisfied in nearly any application, but this is not the case for condition (**b**). The condition (**b**) rules out periodicity in the state transitions, as in the example of the three-state Markov chain with $p_{12} = p_{13} = 0.5$, $p_{21} = p_{31} = 1$, and $p_{ij} = 0$ otherwise. In this example $d = 2$ with $S_1 = \{1\}$ and $S_2 = \{2, 3\}$, and so $\lim_{n\to\infty} p_{ij}^{(n)}$ does not exist for any i, j, e.g. $p_{11}^{(n)}$ is 1 for n even and 0 for n odd.

Under the conditions (**a**) and (**b**), it can be shown that the limiting probability (or equilibrium probability)

$$\pi_j = \lim_{n\to\infty} p_{ij}^{(n)}$$

exists for all $i, j \in I$ and is independent of the starting state i. The π_j can be calculated as the unique solution to the balance equations

$$\boxed{\pi_j = \sum_{k\in I} p_{kj}\pi_k \qquad \text{for } j \in I}$$

together with the normalization equation $\sum_{j\in I} \pi_j = 1$ (you can delete one arbitrarily chosen balance equation as a redundant equation in order to get a square system of linear equations). The balance equations can be easily explained from the Chapman-Kolmogorov equations $p_{ij}^{(n)} = \sum_{k\in I} p_{ik}^{(n-1)} p_{kj}$. Letting $n \to \infty$ in both sides of these equations and interchanging limit and summation (justified by the finiteness of I), you get the balance equations.

The probability π_j can be interpreted as the probability of finding the Markov chain in state j many, many transitions later, whatever the current state is.

This interpretation is obvious. Another interpretation that is more useful for practical purposes is the following:

$$\begin{aligned}&\text{the long-run proportion of times the Markov chain will visit state } j\\ &= \pi_j \quad \text{with probability 1,}\end{aligned}$$

independently of the starting state.[5] This is a kind of large numbers for Markov chains.

As an illustration, consider the following example. On the Island of Hope the weather each day is classified as sunny, cloudy, or rainy. The next day's weather depends only on today's weather and not on the weather of the previous days. If the present day is sunny, the next day will be sunny, cloudy, or rainy with probabilities 0.70, 0.10, and 0.20. The transition probabilities for the weather are 0.50, 0.28, and 0.22 when the present day is cloudy and they are 0.40, 0.30, and 0.30 when the present day is rainy. What are the proportions of sunny days, cloudy days and rainy days over the long run?

This question can be answered by using a three-state Markov chain. Let's say that the weather is in state S if it is sunny, in state C if it is cloudy and in state R if it is rainy. The evolution of the weather is described by a Markov chain with state space $I = \{S, C, R\}$. The matrix $\mathbf{P}$ of the one-step transition probabilities of this Markov chain is given by

$$\begin{array}{ll} \text{from}\backslash\text{to} & \begin{array}{ccc} \;S\; & \;C\; & \;R\; \end{array} \\ \begin{array}{l} S \\ C \\ R \end{array} & \begin{pmatrix} 0.70 & 0.10 & 0.20 \\ 0.50 & 0.28 & 0.22 \\ 0.40 & 0.30 & 0.30 \end{pmatrix}. \end{array}$$

In the context of this weather problem, it will be intuitively clear that the effect of the starting state ultimately fades away and that the limiting probabilities also give the proportions of sunny days, cloudy days and rainy days over the long run. Solving the two balance equations

$$\begin{aligned}\pi_S &= 0.70\pi_S + 0.50\pi_C + 0.40\pi_R\\ \pi_C &= 0.10\pi_S + 0.28\pi_C + 0.30\pi_R\end{aligned}$$

together with the normalization equation $\pi_S + \pi_C + \pi_R = 1$ gives that the long-run proportions of sunny days, cloudy days and rainy days are equal to $\pi_S = 0.5967$, $\pi_C = 0.1771$ and $\pi_R = 0.2262$. Alternatively, these probabilities

[5] The balance equations also apply when the aperiodicity condition (**b**) is not satisfied, in which case π_j can only be interpreted as the long-run average number of visits per unit time to state j.

can be estimated by calculating $\mathbf{P}^n$ for sufficiently large n. Trying several values of n, it was found that $n = 7$ matrix multiplications suffice to have that the elements of the matrix $\mathbf{P}^7$ agree row-to-row to four decimal places:

$$\mathbf{P}^n = \begin{pmatrix} 0.5967 & 0.1771 & 0.2262 \\ 0.5967 & 0.1771 & 0.2262 \\ 0.5967 & 0.1771 & 0.2262 \end{pmatrix} \quad \text{for all } n \geq 7.$$

The Page-Rank algorithm in Google search

The Page-Rank algorithm devised by Larry Page and Sergey Brin, the founders of Google, is based on a Markov chain whose states are the pages of the World Wide Web. This algorithm is one of the methods Google uses to determine a page's relevance or importance. A rudiment of the algorithm will be discussed. Suppose that you have n interlinked web pages. Let n_j be the number of outgoing links on page j. It is assumed that $n_j > 0$ for all j. Let α be a given number with $0 < \alpha < 1$. Imagine that a random surfer jumps from his current page by choosing with probability α a random page amongst those that are linked from the current page, and by choosing with probability $1-\alpha$ a completely random page. Hence the random surfer jumps around the web from page to page according to a Markov chain with the one-step transition probabilities

$$p_{jk} = \alpha\, r_{jk} + (1-\alpha)\,\frac{1}{n} \quad \text{for } j, k = 1, \ldots, n,$$

where $r_{jk} = \frac{1}{n_j}$ if page k is linked from page j and $r_{jk} = 0$ otherwise. The parameter α was originally set to 0.85. The inclusion of the term $(1-\alpha)/n$ can be justified by assuming that the random surfer occasionally gets bored and then randomly jumps to any page on the web. Since the probability of such a jump is rather small, it is reasonable that it does not influence the ranking very much. By the term $(1-\alpha)/n$ in the p_{jk}, the Markov chain satisfies the conditions (**a**) and (**b**). Thus the Markov chain has a unique equilibrium distribution $\{\pi_j\}$. These probabilities can be estimated by multiplying the matrix $\mathbf{P}$ of one-step transition probabilities by itself repeatedly (the computational effort can be considerably reduced by using the trick $\mathbf{P}^m = \mathbf{P}^{\frac{1}{2}m} \times \mathbf{P}^{\frac{1}{2}m}$ for $m = 2^n$). Because of the constant $(1-\alpha)/n$ in the matrix $\mathbf{P}$, things mix better up so that the n-fold matrix product $\mathbf{P}^n$ converges very quickly to its limit. The equilibrium probability π_j gives us the long-run proportion of time that the random surfer will spend on page j. If $\pi_j > \pi_k$, then page j is more important than page k and should be ranked higher.