# Stochastic Modeling

A stochastic matrix is a square matrix whose elements are probabilities and where each row sums to one. One special example of a stochastic matrix is a matrix of transition probabilities.

***Example***
Consider the case where a student enters a college as a one major, but changes to another.

| | STEM | Humanities | Business |
|---|---|---|---|
| STEM | 0.6 | 0.3 | 0.1 |
| Hum. | 0.05 | 0.7 | 0.25 |
| Bus. | 0.2 | 0.3 | 0.5 |

In this matrix, each $a_{ij}$ entry represents the probability that a student will change to that major (or keep that major if it is the same). Suppose we have 1000 freshmen that enter the college

$$x_0 = [300, 500, 200]$$

in the order STEM, Humanities, and Business. If we want to know the number of STEM majors in the class's sophomore year, we can multiply our initial vector by the STEM column of our matrix:

$$\begin{aligned}
\text{STEM Majors} &= (0.6)(300) + (0.05)(500) + (0.2)(200) \\
&= 245 \text{ students} \\
\text{Hum. Majors} &= (0.3)(300) + (0.7)(500) + (0.3)(200) \\
&= 500 \text{ students} \\
\text{Bus. Majors} &= (0.1)(300) + (0.25)(500) + (0.5)(200) \\
&= 255 \text{ students}
\end{aligned}$$

We can continue this process for each year that the students attend the college.

This process for creating a sequence of vectors is called a Markov Chain. In general, we can start with a vector $\mathbf{x_0}$ and a matrix $P$, and we get that

$$\begin{aligned}
\mathbf{x_1} &= \mathbf{x_0} \cdot \mathbf{P} \\
\mathbf{x_2} &= \mathbf{x_1} \cdot \mathbf{P} = \mathbf{x_0} \cdot \mathbf{P} \cdot \mathbf{P} \\
\mathbf{x_3} &= \mathbf{x_2} \cdot \mathbf{P} = \mathbf{x_0} \cdot \mathbf{P} \cdot \mathbf{P} \cdot \mathbf{P},
\end{aligned}$$

which yields the recurrence relation

$$\mathbf{x_n} = \mathbf{x_0 P^n}.$$

$P^n$ is called the $n$-step transition matrix. In this matrix, the $a_{ij}$ element gives the probability of going from state $i$ to state $j$.

***Example***
Consider the movement of people from the city to the suburb.

| | Cities | Suburbs |
|---|---|---|
| Cities | 0.96 | 0.04 |
| Suburbs | 0.1 | 0.99 |

with the initial vector

$$\mathbf{x_0} = [\mathbf{60}, \mathbf{125}].$$

We can introduce some "noise" into the system by discussing the eigenvalues of this system. In the usual case, $\lambda = 1$. We can show this by solving the system.

$$\begin{aligned}(\mathbf{P} - \lambda)\mathbf{x} &\rightarrow \begin{bmatrix} -0.4 & 0.4 & 0 \\ 0.1 & 0.1 & 0 \end{bmatrix} \\ &\rightarrow \begin{bmatrix} 1 & -1 & 0 \\ 0 & 0 & 0 \end{bmatrix},\end{aligned}$$

so $x = y = 1$.

Often in Markov processes, we come across matrices raised to a power. To help us calculate these products, we can come up with a new matrix, $\mathbf{S}$. We can then calculate $\mathbf{S^{-1}DS}$, where $D$ is diagonalized matrix of the eigenvalues of $\mathbf{P}$. We know that we can diagonalize $\mathbf{P}$ by the following theorem:

An $n \times n$ matrix $\mathbf{P}$ is diagonalizable if and only if it has $n$ linearly independent eigenvectors.

***Example***
Consider three types of detergent that yield the Markov matrix

| | Tide | Bounty | Bounce |
|---|---|---|---|
| Tide | .8 | .1 | .1 |
| Bounty | .2 | .6 | .2 |
| Bounce | .3 | .3 | .4 |

where the entries indicate the probability that a user will stay with that product. Rather than finding the next state, we will consider a vector $\boldsymbol{\pi}$, called the limiting state probability vector. This vector will be representative of the probabilities of users staying with a brand. Similar to the previous examples, we have

$$\boldsymbol{\pi}\mathbf{P} = \boldsymbol{\pi}.$$

From this equation, we get

$$\begin{cases} .8\pi_1 + .2\pi_2 + .3\pi_3 = \pi_1 \\ .1\pi_1 + .6\pi_2 + .3\pi_3 = \pi_2 \\ .1\pi_1 + .2\pi_2 + .4\pi_3 = \pi_3 \end{cases}$$

The third equation in this case is a linear combination of the first two, so we cannot user it here. However, we can replace it with a linearly independent row:

$$\begin{cases} .8\pi_1 + .2\pi_2 + .3\pi_3 = \pi_1 \\ .1\pi_1 + .6\pi_2 + .3\pi_3 = \pi_2 \\ \pi_1 + \pi_2 + \pi_3 = 1 \end{cases} \rightarrow \left[\begin{array}{ccc|c} -.2 & .2 & .3 & 0 \\ .1 & -.4 & .3 & 0 \\ 1 & 1 & 1 & 0 \end{array}\right] \rightarrow \boldsymbol{\pi} = \begin{bmatrix} 6/11 \\ 3/11 \\ 2/11 \end{bmatrix}.$$

# Irreducible Matrices

We attempted to find the $\pi$ vector in the previous examples. We can use the following theorem to discuss special Markov Chains.

Let $P$ be the transition matrix for an irreducible and aperiodic Markov Chain. Then there exists a unique probability vector $\pi$ such that

$$\pi \cdot P = \pi.$$

It follows that if $P$ is a regular matrix, then we can also find the $\pi$ vector. A more interesting phenomenon occurs when we do not have a regular $P$ matrix. A typical model of this can be seen in the Drunkard's walk without reflection.

## The Drunkard's Walk

The Drunkard's walk represents a random walk with several states. At some states, the drunk man may either end up in his bed or on the street. In the reflective version of this walk, it is possible that after going into his bed, the drunkard goes back out. In this example, we want to consider absorbing states, which are states that are not able to be escaped.

In this walk, we want to model the drunkard by saying that if he falls asleep in the street or he goes to his bed, then the walk ends. Those conditions are representative of absorbing states.

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

More generally, an absorbing matrix takes the form of an $r \times r$ matrix with $k$ absorbing states. It is a fact that ultimately everything will end up in the absorbed state. Markov Chains are only absorbing when all absorbing state can be reached. If a state is not absorbing, it is called a transient state.

For the drunkard's walk, we can partition the matrix into four parts:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 1/2 & 0 & 1/2 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 1/2 & 0 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

which allows us to calculate a new matrix $N = (I - Q)^{-1}$. Every $n_{ij}$ entry in the $N$ matrix tells us the average number of times that the process is in the $j$ state if it starts in the $i$ state.

# Poisson Probabilities

Suppose that $f(x,t)$ is the probability of getting $x$ successes during a time interval of length $t$. Assume the prob of success during a very small time internal from $t$ to $t + \delta t$ is $\mu$, the probability of more than one success during a time interval is negligible, and the probability of a success during such a time interval does not depend on what happened prior to time $t$. The Poisson distribution is given by

$$f(x, t + \Delta t) = f(x,t) \cdot [1 - \mu \Delta t] + f(x-1,t)\mu\Delta t = \frac{e^{-\mu}\mu^x}{x!}$$

The variable $x$ is indicative of the number of successes we want in the next time interval $t + \delta t$. One example of the use of Poisson distributions is to calculate the probability that an earthquake will occur again in a certain area.

***Example***
Suppose that in southern California, 5 earthquakes occur every 100 years on average. What is the probability that 2 earthquakes will occur in the next century?

The answer is very straightforward:

$$P = \frac{e^{-5}5^2}{2!}$$

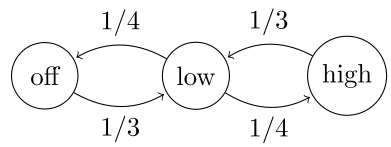
If we wanted to know the probability in the next 10 years, we can calculate

$$P = \frac{e^{-0.5}0.5^2}{2!}.$$

# Markov Chains and Continuity

We can form a relationship between Markov Chains and differential equations.

***Example***
In the summer, an AC unit is in one of 3 states: (0) off, (1) low, or (2) high. While off, transitions to low occur after an exponential time with an expected time of 3 minutes.




We assign a $\lambda$ value to each state where $\lambda \in (0, \infty)$. The density function is given by

$$f(t) = \lambda e^{-\lambda t}$$

for $t > 0$.

# Various Distributions

| | | | |
|---|---|---|---|
| Binomial Distribution | $B(x; n, p)$ | $\binom{n}{x} p^x q^{n-x}$ | Discreet |
| Poisson Distribution | $\rho(x, \mu)$ | $\frac{e^{-\mu}\mu^x}{x!}$ | Discreet |
| Geometric Distribution | $G(x; p)$ | $p(1-p)^{x-1}$ | Discreet, Memoryless |
| Bell Curve | $B(\sigma, x, \mu)$ | $\int_a^b \frac{1}{\sigma\sqrt{2\pi}} e^{-0.5\left(\frac{x-\mu}{\sigma}\right)^2}\, dx$ | Continuous |
| Exponential Distribution | $E(t, \theta)$ | $\int_a^b \frac{1}{\theta} e^{-t/\theta}\, dt$ | Continuous, Memoryless |