Markov Chains

Review of Markov Chains

Let's see an example called city-suburb problem:

Suppose the population of a city and its suburbs were measured each years. Because the total population might be changed, we can use percentage to represent the population. For example, 0.7 of the population lives in the city and 0.3 in the suburbs in 2017. If one want to know the annual migration between these two parts of the metropolitan region, there are four moving cases need to be considered:

- From city to city
- From city to suburbs
- From suburbs to city
- From suburbs to suburbs

For example, each year 10% of the city population moves to the suburbs, and 2% of the suburban population moves to the city. Based on the information, we know 90% the city population stays in city and 98% of the suburban population stay in suburbs. If we assume no significant change of the moving percentages in the future, we can use the following way to predict the population in the next year: Since the city population is 70% and the suburban population is 30% in 2017, so in 2018 we obtain

- From city to city: $70\% \times 90\% = 63\%$
- From city to suburbs: $70\% \times 10\% = 7\%$
- From suburbs to city: $30\% \times 2\% = 0.6\%$
- From suburbs to suburbs: $30\% \times 98\% = 29.4\%$

which means $63\% + 0.6\% = 63.6\%$ of the population lives in the city and $29.4\% + 7\% = 36.4\%$ in the suburbs in 2018.

If we want know the population in 2019, or 2021, we need to do these again and again.

Here we can use a matrix to represent the annual migration percentage

$$\begin{bmatrix} 90\% & 2\% \\ 10\% & 98\% \end{bmatrix}$$

and a vector to represent the population in 2017

$$\begin{bmatrix} 70\% \\ 30\% \end{bmatrix}.$$

Then the population in 2018 will be the multiplcation of the migartion matrix and population vector:

$$\begin{bmatrix} 90\% & 2\% \\ 10\% & 98\% \end{bmatrix} \begin{bmatrix} 70\% \\ 30\% \end{bmatrix} = \begin{bmatrix} 63.6\% \\ 36.4\% \end{bmatrix}$$

We can use the multiplcation to predict the population in 2019, 2020, or 2021:

- In 2019,

$$\begin{bmatrix} 90\% & 2\% \\ 10\% & 98\% \end{bmatrix} \begin{bmatrix} 90\% & 2\% \\ 10\% & 98\% \end{bmatrix} \begin{bmatrix} 70\% \\ 30\% \end{bmatrix} = \begin{bmatrix} 90\% & 2\% \\ 10\% & 98\% \end{bmatrix} \begin{bmatrix} 63.6\% \\ 36.4\% \end{bmatrix} = \begin{bmatrix} 57.97\% \\ 42.03\% \end{bmatrix}.$$

- In 2020,

$$\begin{bmatrix} 90\% & 2\% \\ 10\% & 98\% \end{bmatrix} \begin{bmatrix} 90\% & 2\% \\ 10\% & 98\% \end{bmatrix} \begin{bmatrix} 90\% & 2\% \\ 10\% & 98\% \end{bmatrix} \begin{bmatrix} 70\% \\ 30\% \end{bmatrix} = \begin{bmatrix} 53.01\% \\ 46.99\% \end{bmatrix}.$$

- In 2020,

$$\begin{bmatrix} 90\% & 2\% \\ 10\% & 98\% \end{bmatrix} \begin{bmatrix} 90\% & 2\% \\ 10\% & 98\% \end{bmatrix} \begin{bmatrix} 90\% & 2\% \\ 10\% & 98\% \end{bmatrix} \begin{bmatrix} 90\% & 2\% \\ 10\% & 98\% \end{bmatrix} \begin{bmatrix} 70\% \\ 30\% \end{bmatrix} = \begin{bmatrix} 48.65\% \\ 51.35\% \end{bmatrix}.$$

We can put these five-year vectors of population together:

$$\begin{array}{ccccc} 2017 & 2018 & 2019 & 2020 & 2021 \\ \begin{bmatrix} 70\% \\ 30\% \end{bmatrix} & \begin{bmatrix} 63.6\% \\ 36.4\% \end{bmatrix} & \begin{bmatrix} 57.97\% \\ 42.03\% \end{bmatrix} & \begin{bmatrix} 53.01\% \\ 46.99\% \end{bmatrix} & \begin{bmatrix} 48.65\% \\ 51.35\% \end{bmatrix} \end{array}$$

You may find the entries of vectors of population are all nonnegative and the sum of the entries in each vector is 1. We call this kind of vector **probability vector**. The migartion matrix

$$\begin{bmatrix} 90\% & 2\% \\ 10\% & 98\% \end{bmatrix}$$

is called **stochastic matrix** or **transition matrix** whose columns are probability vectors and Those population vectors are called **Markov chain** which is a squence of probability vectors together with the stochastic matrix.

We can check all the results by using MATLAB:

```
> A=[0.9  0.02;  0.1  0.98];  x=[0.7;0.3];
>> x1=A*x
x1 =
     0.6360
     0.3640

>> x2=A*x1
x2 =
```

```
      0.5797
      0.4203

>> x3=A*x2
x3 =
      0.5301
      0.4699

>> x4=A*x3
x4 =
      0.4865
      0.5135
```

The long term behavior

We can use Markov chain with the stochastic matrix to research the long term behavior of phenomena. Here is an example from election. Assume there are three parties: D party, R party, and L party. Suppose we record the outcome of the congressional election each time by a vector and the outcome of one election depends only on the result of the preceding election. We can setup our election matrix as follow:

$$\begin{bmatrix} 0.5 & 0.2 & 0.3 \\ 0.3 & 0.8 & 0.3 \\ 0.2 & 0 & 0.4 \end{bmatrix}$$

- The first column describes what the people voting D party will do in the next election. 50% will vote D party again, 20% will vote R party, and 30% will vote L party.
- The second column describes what the people voting R party will do in the next election. 20% will vote D party, 80% will vote R party again, and no one will vote L party.
- The third column describes what the people voting L party will do in the next election. 30% will vote D party, 30% will vote R party, and 40% will vote L party again.

We can study the long tern behavier by multiplying the election matrix many times on the vector x. Then we can get

$$x_1 = Ex, \;\; x_2 = EEx = E^2x, \;\; x_3 = EEEx = E^3x, \cdots$$

In MATLAB, we obtain

```
>> E=[0.5   0.2   0.3;  0.3   0.8   0.3  ;  0.2   0  0.4];  x=[1;0;0];
>> X=[E*x,  E^2*x,  E^3*x,  E^4*x,  E^5*x,  E^6*x,E^7*x,E^8*x...
E^9*x,  E^10*x,  E^11*x,  E^12*x,  E^13*x,  E^14*x,E^15*x,E^16*x]
```

```
X =

  Columns 1 through 7

    0.5000    0.3700    0.3290    0.3133    0.3064    0.3032    0.3016
    0.3000    0.4500    0.5250    0.5625    0.5813    0.5906    0.5953
    0.2000    0.1800    0.1460    0.1242    0.1123    0.1062    0.1031

  Columns 8 through 14

    0.3008    0.3004    0.3002    0.3001    0.3000    0.3000    0.3000
    0.5977    0.5988    0.5994    0.5997    0.5999    0.5999    0.6000
    0.1016    0.1008    0.1004    0.1002    0.1001    0.1000    0.1000

  Columns 15 through 16

    0.3000    0.3000
    0.6000    0.6000
    0.1000    0.1000
```

After certain iterations, the outcome of voting is not changing anymore. We know that there is a **steady-state vector** x such that

$$Ex = x.$$

We can find the steady-state vector of the election matrix E directly. Since $Ex = x$, we have

$$Ex - x = 0 \Rightarrow (E - I)x = 0$$

which is solving

$$\begin{bmatrix} 0.5-1 & 0.2 & 0.3 \\ 0.3 & 0.8-1 & 0.3 \\ 0.2 & 0 & 0.4-1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$

and we can get the answer by using MATLAB:

```
>>E1=[E-eye(3),[0;0;0]];
>> rref(E1)

ans =

    1.0000         0   -3.0000         0
         0    1.0000   -6.0000         0
         0         0         0         0
```

Then $x_1 = 3x_3$, $x_2 = 6x_3$, and x_3 is free. Since $\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ is a probability vector, $x_1+x_2+x_3 = 1$, so we have

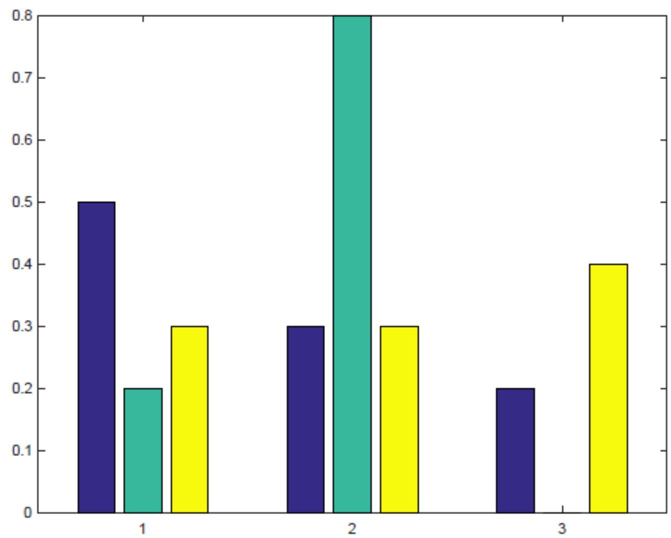
$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 0.3 \\ 0.6 \\ 0.1 \end{bmatrix}.$$

Matlab charts plotting

- `bar( )`

  The command `bar(y)` creates a bar graph with one bar for each element y. If y is a matrix, then bar groups the bars according to the rows in y. For example, we can use `bar(E)` to analysis the election matrix E:

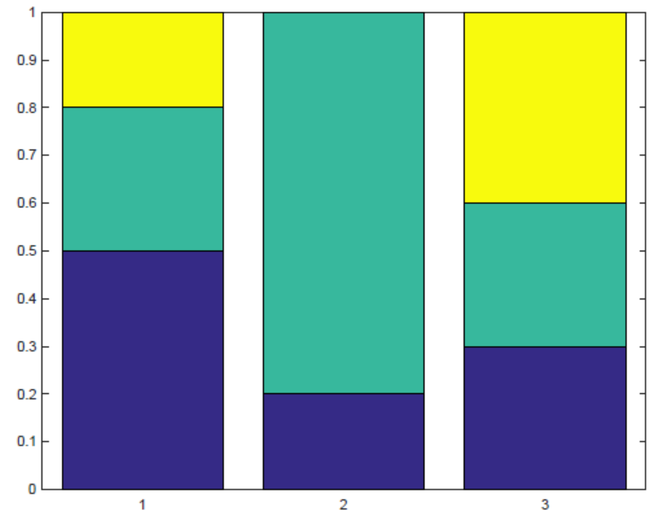
```
>>bar(E)
```



The label 1's data is from row one which is about the voting probabilities next time for the one who vote D party this time. The fisrt bar of label 1 is the percentage or probabilities of the voters who will still vote D party next time. The second bar is the probability of voting R party and third bar is the probability of voting L party.

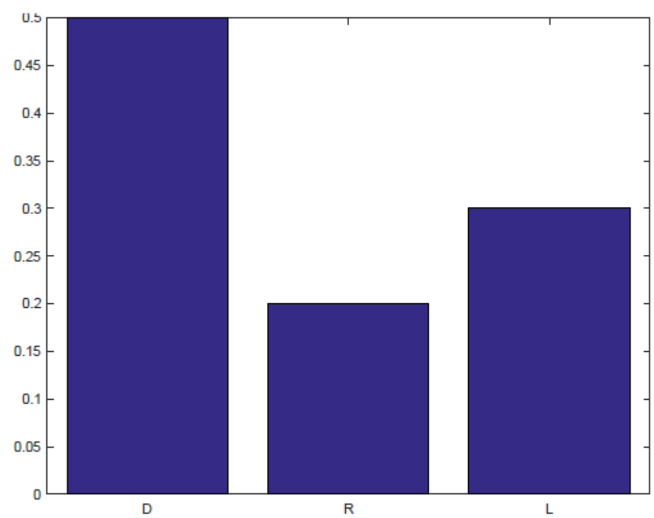
We also can "stark" the bars together:

```
>> bar(E','stack');
```



If we just want to focus on the D party case and label each bar, we can use the following commands:

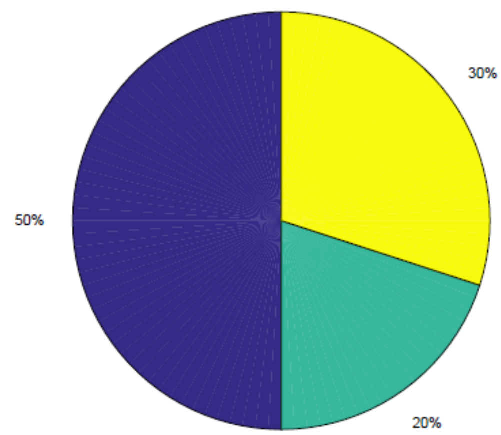
```
>> x=1:3;
bar(x,E(1,:));
set(gca,'xticklabel',{'D','R','L'});
```



- `pie( )`

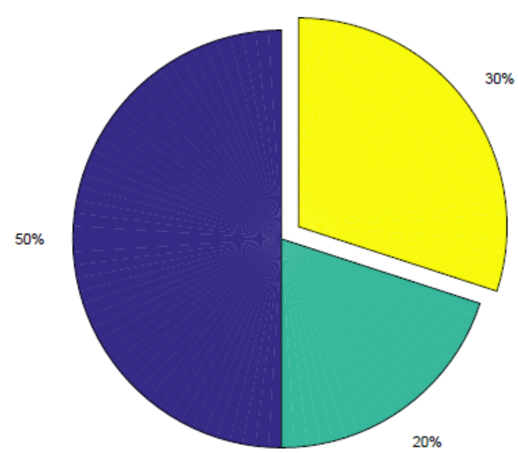
  Besides `bar( )`, we also can use pie chart `pie(E(1,:))`:

  ```
  >> pie(E(1,:));
  ```



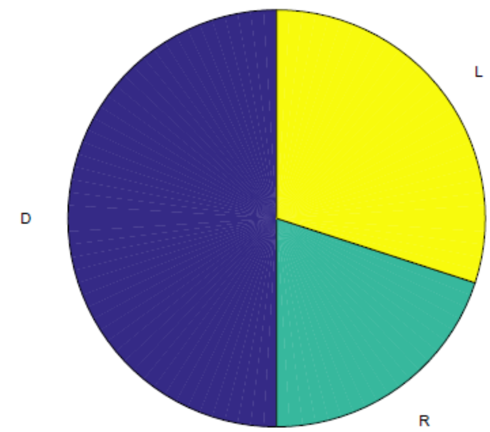
We can offset the third pie slice by setting the corresponding explode element to 1:

```
explode = [0 0 1];
pie(E(1,:),explode)
```



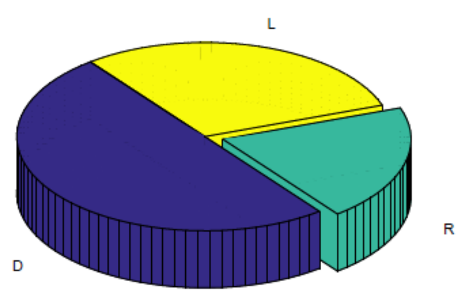
We can label each slice of a pie chart:

```
labels = {'D','R','L'};
pie(E(1,:),labels)
```



We also can try the 3D pie chart `pie3( )` and offset one slice:

```
explode = [0 1 0];
label ={'D','R','L'};
pie3(E(1,:), explode, label);
```



Absorbing Markov Chains

Let's talk about a special topic in Markov chains. It is called absorbing Markov chains. Before that, I would like introduce some names.

- stochastic matrix

  We have introduced stochastic matrix last time. The definition from previous lecture is the sum of the entries in each **column** of stochastic matrix is 1. However you may find out that the matrices in project3 is each **row** summing to 1, not column.

  In mathematics, there are several different definitions and types of stochastic matrices:

  A **right stochastic matrix** is a real square matrix, with each **row** summing to 1.

  A **left stochastic matrix** is a real square matrix, with each **column** summing to 1.

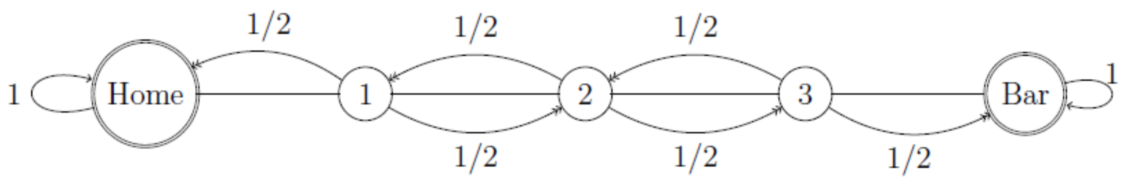
  If an $n \times n$ left stochastic matrix A and a probability vector $x = \begin{bmatrix} x1 \\ x2 \\ \vdots \\ x_n \end{bmatrix}$ are given. Then $x, Ax, A^2x, A^3x, \cdots$ is a Markov chain.

  If an $n \times n$ right stochastic matrix A and a probability vector $x = \begin{bmatrix} x1 \\ x2 \\ \vdots \\ x_n \end{bmatrix}$ are given. Then $x^t, x^tA, x^tA^2, x^tA^3, \cdots$ is a Markov chain where x^t is the transpose of x.

  In linear algebra, we use left stochastic matrix and $A^m x$. But it is equivalent to use right stochastic matrix and $x^t A^m$ for some integer m.

Now, let's see an example of absorbing Markov chains.

- Drunkards Walk



A man walks along a four-block stretch of Westheimer Road. If he is at corner 1, 2, or 3, then he walks to the left or right with equal probability. He continues until he reaches corner 4, which is a bar, or corner 0, which is his home. If he reaches either home or the bar, he stays there.

Then the transition martix of this problem is

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

1. Give a probability vector x which represent the status that the man is at home.
2. Calculate $x^t P$. Explain the result you got.
3. Give a probability vector x which represent the status that the man is at corner 2.
4. Calculate $x^t P$. Explain the result you got.

For an absorbing Markov chain, renumber the states so that the transient states come first. If there are r absorbing states and t transient states, the transition matrix will have the following canonical form

$$P \to \begin{bmatrix} 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 1/2 & 0 \\ 0 & 0 & 1/2 & 0 & 1/2 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 0 & 1/2 & 0 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix}$$

First we change the row1 to be row 4. Second we change the column 1 to be column 4.

$$P_1 = \left[\begin{array}{ccc|cc} 0 & 1/2 & 0 & 1/2 & 0 \\ 1/2 & 0 & 1/2 & 0 & 0 \\ 0 & 1/2 & 0 & 0 & 1/2 \\ \hline 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{array}\right] = \begin{bmatrix} Q & R \\ 0_{2\times 3} & I_{2\times 2} \end{bmatrix}$$

where Q is the transient matrix

$$\begin{bmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 0 \end{bmatrix}.$$

5. Find P_1^2 and Q^2.
6. Find P_1^4 and Q^4.
7. Based on 5. and 6., could you find out some patterns?

Theorem 0.1. *For an absorbing Markov chain the matrix $I - Q$ has an inverse N and*

$$N = I + Q + Q^2 + Q^3 + \cdots$$

For an absorbing Markov chain P , the matrix $N = (I - Q)^{-1}$ is called the fundamental matrix for P. The entry n_{ij} of N gives the expected number of times that the process is in the transient state s_j if it is started in the transient state s_i.

Since

$$Q = \begin{bmatrix} 0 & 1/2 & 0 \\ 1/2 & 0 & 1/2 \\ 0 & 1/2 & 0 \end{bmatrix},$$

then

$$N = (I - Q)^{-1} = \begin{bmatrix} 1 & -1/2 & 0 \\ -1/2 & 1 & -1/2 \\ 0 & -1/2 & 1 \end{bmatrix}^{-1},$$

5. Find N.
6. What is expected number of times that the man arrive corner 3 if he started from corner 1?
7. What is expected number of times that the man arrive corner 2 if he started from corner 1?