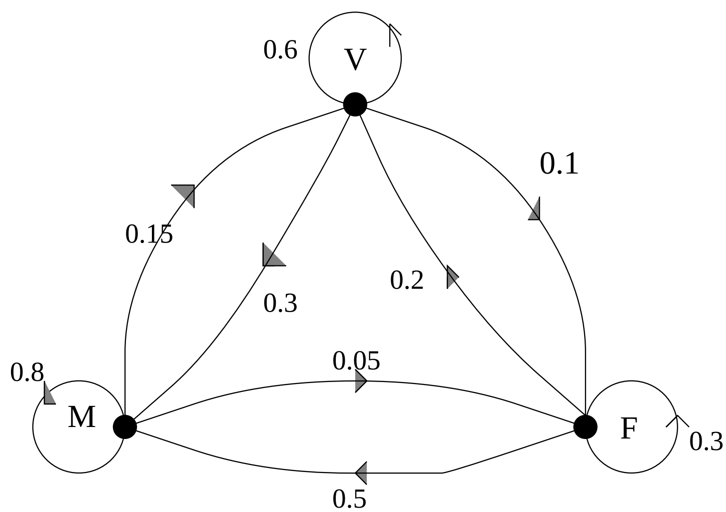
Intro. to Linear Algebra 250-C
Extra Credit Project 2 — Graphs and Markov Processes

Introduction: In this project we will learn about the connection between linear algebra and market distribution.

Suppose that, every day, there are three types of entrees that are offered at the local dining hall: Meat, Fish and Vegetarian. Students who eat at the dining hall choose any one of these entrees. Some of them choose the same type every time, while other students switch from time to time. The following directed graph below shows the percentages that stay and the percentages that switch. Here, M denotes Meat, F is for Fish and V is for (can you guess?) Vegetarian.



For example, the graph says that, of the students who order Meat on one day, 0.8 of them (that is, 80%) order Meat the next day, while 0.05 of them order Fish and 0.15 of them order Vegetarian.

1. If at Monday's lunch 50 students order Meat, 30 order Fish and 20 order Vegetarian, find out how many order each kind of entree at Tuesday's lunch. For example, the number ordering Meat will be

$$0.8(50) + 0.5(30) + 0.3(20) = 61$$

Note that the coefficients come from the arrows pointing toward the Meat vertex.

2. The above equation together with the two others you found can be written in matrix form $\mathbf{y} = A\mathbf{x}$,

$$\begin{bmatrix} 61 \\ \\ \\ \end{bmatrix} = \begin{bmatrix} 0.8 & 0.5 & 0.3 \\ & & \\ & & \end{bmatrix} \begin{bmatrix} 50 \\ 30 \\ 20 \end{bmatrix}$$

where $\mathbf{x}$ represents Monday's entree distribution, $\mathbf{y}$ represents Tuesday's entree distribution and A is the coefficient matrix. To find A and $\mathbf{y}$, you need to fill in the eight missing numbers. Note that the first component of the vectors $\mathbf{x}$ and $\mathbf{y}$ represent the number of Meat entrees, the second component represents the number of Fish and the third component represents the number of Vegetarian entrees. As a check, note that each column of A should sum up to 1 (why?). Use MATLAB to create the vectors $\mathbf{x}$, $\mathbf{y}$ and the matrix A.

3. Describe the interpretation of the $(2,3)$ entry of the 3×3 matrix A in question 2 in terms of the changing diet preference.

4. The vector $\mathbf{y}$ gives us the entree distribution for Tuesday's lunch. How do we determine Wednesday's distribution? Using the same reasoning as above, we see that this is given by $A\mathbf{y}$. Use this idea to explain what entree distribution $A^2\mathbf{x}$ gives us. (Hint: $A^2\mathbf{x} = A(A\mathbf{x})$).

5. Use MATLAB to compute A^2. What do its entries represent? (Just describe the meaning of the $(2,3)$ entry in A^2; the others are similar).

6. What are the distributions for Thursday's lunch? For Friday, Saturday and Sunday's lunches? Use MATLAB to calculate this.

7. If the dining hall manager counts 50 Meat, 30 Fish and 20 Vegetarian entrees eaten on Monday, for how many of each entree should the manager plan for the rest of the week? How do you write this in matrix form? Use MATLAB to do the calculations.

We are now interested in seeing how the entree distribution changes over a long time (over months or years). From the above problems, we have seen that, if we start with some initial entree distribution $\mathbf{x_0}$, then the next day's distribution is $\mathbf{x_1} = A\mathbf{x_0}$. The distribution for the day after that is $\mathbf{x_2} = A\mathbf{x_1}$, etc. This gives us a sequence

$$\mathbf{x_0}, \;\; \mathbf{x_1} = A\mathbf{x_0}, \;\; \mathbf{x_2} = A\mathbf{x_1}, \ldots \qquad (*)$$

8. What does $\mathbf{x_k}$ represent? Show that $\mathbf{x_k} = A^k\mathbf{x_0}$ for $k = 1, 2, 3, \ldots$.
The sequence of equations $(*)$ is called a *Markov process*, and the matrix A is called the *transition matrix* for the process.

9. Using MATLAB, find A^k for $k = 1, 2, 3, 5, 10, 20, 50, 100$. What happens as k gets bigger and bigger? Fix a matrix that approximates $\lim_{k\to\infty} A^k$ accurately to within 4 decimal places? Now compute $A^k\mathbf{x_0}$ for $k = 1, 2, 3, 5, 10, 20, 50, 100$ and $\mathbf{x_0} = [50, 30, 20]^T$. What happens to the distribution over time?

Let

$$\mathbf{x}_\infty = \lim_{k\to\infty} \mathbf{x_k} = \lim_{k\to\infty} A^k\mathbf{x_0}$$

We call $\mathbf{x}_\infty$ the *steady state vector.*

10. Suppose that we had an initial distribution where all 100 students ate a Meat entree. What would $\mathbf{x_0}$ and $\mathbf{x}_\infty$ be in this case? Suppose that initially 50 students ate Fish, 50 students ate Vegetarian and 0 students ate Meat. What would $\mathbf{x_0}$ and $\mathbf{x}_\infty$ be then? Make

up your own initial distribution vector $\mathbf{x_0}$ and find out what is $\mathbf{x}_\infty$ in this case. (Remember that there are 100 entrees total).

11. What did you notice about $\mathbf{x}_\infty$ in each of the four cases above? Show that $A\mathbf{x}_\infty = \mathbf{x}_\infty$. What is the relationship between the initial distribution and the distribution over a long period of time?

12. Use the MATLAB command **[S,E] = eig(A)** to find eigenvalues and eigenvectors of A. What are A's eigenvalues? Which one reflects the existence of a steady state vector? (Hint: look at question 11.) What is the corresponding eigenvector in the matrix **S**? How is it related to the steady state vector?

The above matrix A is an example of a *stochastic* matrix. This means that its entries are non-negative and each of its columns sum up to 1. The results that you just discovered about distribution and steady state vectors hold for all stochastic matrices. (Repeat this project for other stochastic matrices and see for yourself!).