Due Friday, November 10, 2017 NAME(S): ____________

Math 215 Project (25 pts) : Using Linear Algebra to solve Graph Problems

1 Introduction

First let us define what we mean by a **graph**. A **graph** is a set of points (called **vertices**, or **nodes**) and a set of lines called **edges** connecting some pairs of vertices. Two vertices connected by an edge are said to be **adjacent**. Consider the graph in Figure 1. Notice that a vertex need not be connected to any other vertex (D), and that a vertex may be connected to itself (F).

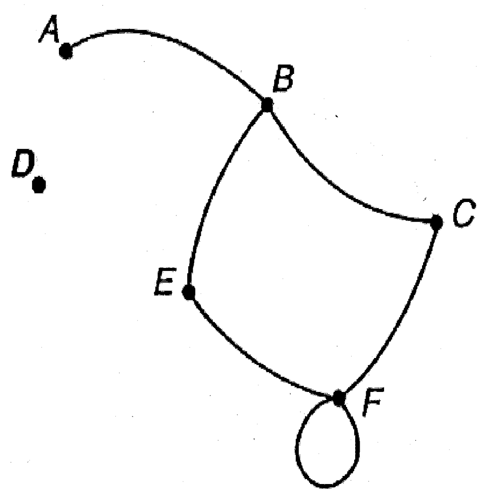


Figure 1:

An example of a graph is the route map that most airlines produce. A copy of the route map for Vanguard Airlines from 1999 is given in Figure 2. Here the vertices are the cities to which Vanguard used to fly. Two vertices are connected if a direct flight flies between them. When we are given a graph, some natural questions arise. We may want to know if two vertices are connected by a sequence of two edges, even if they are not connected by a single edge. In Figure 1, A and C are connected by a two-edge sequence. In our route map (Figure 2), Pittsburgh and Kansas City are connected by a two-edge sequence, meaning that a passenger would have to stop in Chicago while flying between those cities on Vanguard. We may want to know if it is possible to get from a vertex to another vertex. It is impossible to go from vertex D in Figure 1 to any other vertex, but a passenger on Vanguard can get from any city in their network to any other given enough flights. But how many flights are enough? This is another issue of interest: what is the minimum number of steps to get from one vertex to

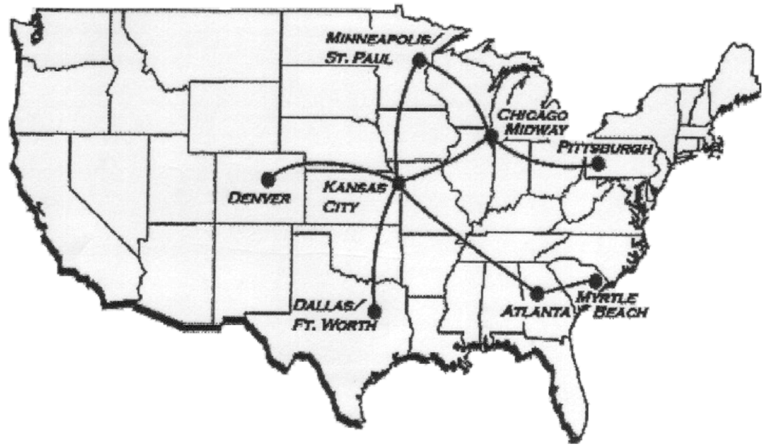


Figure 2: 1999 airline route map for Vanguard airlines.

another? What is the minimum number of steps to get from any vertex on the graph to any other? While these questions are relatively easy to answer for a small graph, as the number of vertices and edges grows, it becomes harder to keep track of all the different ways the vertices are connected. Matrix notation and computation can help us answer these questions.

The **adjacency matrix** for a graph with n vertices is an $n \times n$ matrix whose (i, j) entry is 1 if the i^{th} and j^{th} vertex are connected, and 0 if they are not. If in Figure 1 we let A be vertex one, B be vertex two, etc., we find that the adjacency matrix for this graph is,

$$M = \begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 & 1 \end{bmatrix}$$

where row one and column one of M corresponds to vertex A, row two and column two corresponds to vertex B, etc. If the vertices in the Vanguard graph (Figure 2) respectively correspond to Chicago Midway, Denver, Dallas/Ft. Worth, Minneapolis/St. Paul, Atlanta, Kansas City, Myrtle Beach, and Pittsburgh, then the adjacency matrix for Vanguard is,

$$V = \begin{bmatrix} 0 & 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

where row one and column one of V corresponds to city (vertex) Chicago Midway, row two and column two corresponds to city (vertex) Denver, etc. Vanguard graph (Figure 2) shows that Kansas City is the airlines hub (most connection). This is illustrated in the matrix row 6 and column 6 with many entries of 1. We now use adjacency matrices to address the questions we raised earlier. Which vertices are connected by a two-edge sequence? How many different two-edge sequences connect each pair of vertices?

Consider Figure 1. Vertices B and F are connected by two-edge sequences in 2 different ways: B to C to F and B to E to F. Let m_{ij} be the (i, j) entry in the adjacency matrix M, and note that column 2 (associated with vertex B) of matrix M multiplied by row 6 (associated with vertex F) of M is given by

$$\begin{aligned}\begin{bmatrix} m_{61} & m_{62} & m_{63} & m_{64} & m_{65} & m_{66} \end{bmatrix} \cdot \begin{bmatrix} m_{12} \\ m_{22} \\ m_{32} \\ m_{42} \\ m_{52} \\ m_{62} \end{bmatrix} &= m_{12}m_{61} + m_{22}m_{62} + m_{32}m_{63} + m_{42}m_{64} + m_{52}m_{65} + m_{62}m_{66} \\ &= (1)(0) + (0)(0) + (1)(1) + (0)(0) + (1)(1) + (0)(1) \\ &= 2\end{aligned}$$

which is the number of two-step sequences between B and F. This calculation works because in order for a two-step sequence to occur, B and F must both be connected to an intermediate vertex. Since B connects with C ($m_{32} = 1$) and C with F ($m_{63} = 1$), we have $m_{32}m_{63} = (1)(1) = 1$ (C is associated with row 3 and column 3 entries of M); since B connects with A ($m_{12} = 1$) but F does not ($m_{61} = 0$) we have $m_{12}m_{61} = (1)(0) = 0$ (A is associated with row 1 and column 1 entries of M).

Observation The number of two-step sequences between vertex i and vertex j in a graph with adjacency matrix M is the (i, j) entry in M^2. Which can be generalized.
The number of k-step sequences between vertex i and vertex j in a graph with adjacency matrix M is the (i, j) entry in M^k.

If M is the adjacency matrix for Figure 1, we have

$$M^2 = \begin{bmatrix} 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 3 & 0 & 0 & 0 & 2 \\ 1 & 0 & 2 & 0 & 2 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 2 & 0 & 2 & 1 \\ 0 & 2 & 1 & 0 & 1 & 3 \end{bmatrix} \quad M^3 = \begin{bmatrix} 0 & 3 & 0 & 0 & 0 & 2 \\ 3 & 0 & 5 & 0 & 5 & 2 \\ 0 & 5 & 1 & 0 & 1 & 5 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 5 & 1 & 0 & 1 & 5 \\ 2 & 2 & 5 & 0 & 0 & 5 \end{bmatrix}$$

Thus there is 1 two-step sequence from C to F, and 5 three step sequences between C and F. You can verify these, by referring back to Figure 1. In your observations of figures, you might notice that some two-step and three step sequences may not be meaningful (e.g. A to B to E to B three-step sequence from A to B). On the Vanguard route (Figure 2), we see that Minneapolis is reachable in two steps from Chicago Midway, (Chicago Midway to Kansas City to Minneapolis), but in reality we would not care, since there is a direct flight between the two cities. A better question to ask of a graph might be, What is the least number of edges which must be traversed to go from vertex A to vertex B? To answer this question, consider the matrix

$$S_k = M + M^2 + M^3 + \ldots + M^k.$$

The (i, j) entry in this matrix tallies the number of ways to get from vertex i to vertex j in k steps or less. If such a trip is impossible, this entry will be zero. Thus to find the shortest number of steps between the vertices, continue to compute S_k as k increases; the first k for which the (i, j) entry in S_k is non-zero is the shortest number of steps between i and j, Note that this process is non-constructive; that is, we will know the shortest number of steps but the method does not show us what those steps are.

Another question; is possible to go from any vertex in a graph to any other? If a graph has the property that each vertex is connected to every other vertex in some number of steps, then the graph is **connected**. How can we tell if a graph is connected? This should be easy to see from a small graph, but is harder to see from the adjacency matrix of a large graph. However, there is a calculation we can do. Suppose that the graph contains n vertices, then the largest number of steps it could take to go from any vertex to any other vertex is n steps. Why? Thus $S_n = M + M^2 + M^3 + \ldots + M^n$ can help us. If there are any zeros in this matrix, it is impossible for some pair of vertices to connect in n steps or less, so this pair will never connect, and the graph is not connected.

EXERCISE 1 (4pts):
List all 5 of the three-step sequences between C and F in Figure 1.

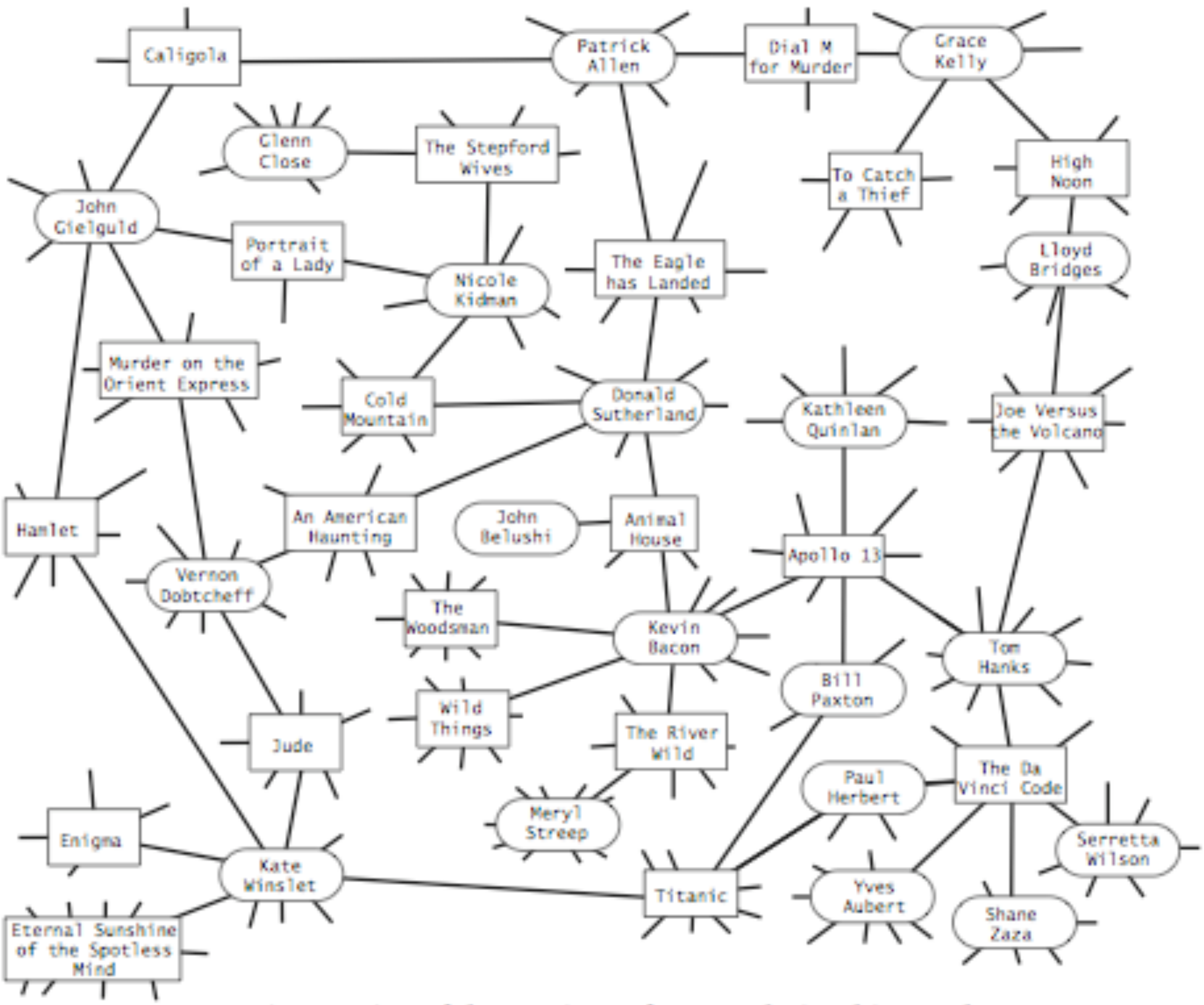
EXERCISE 2 (4pts):
Using the matrix V on page 2, which Vanguard cities may be reached by a two flight sequence from Chicago Midway? Which may be reached by a three flight sequence? For full credit you must use the matrix V on page 2.

EXERCISE 3 (4pts):
Which trip(s) in the Vanguard network take the greatest number of connections (i.e. flights) between cities? For full credit you must use the matrix V on page 2.

EXERCISE 4 (4pts):
(a) Is the graph in Figure 1 connected?
(b) Is the Vanguard graph connected?
Must show work (involves using matrices M and V) for full credit.

EXERCISE 5 (9 pts):
Degrees Separation of Kevin Bacon. Use the actors from the figure and table given on the next page.
a.) Create the adjacency matrix (cities are replaced with actors). If the actors were in the same movie together, place a 1 in the matrix entry. Hint: Spreadsheet program (e.g. Excel) can be useful in setting up the matrix for part a and getting an output that can be used in parts b, c, and d.
(Do NOT use a spreadsheet program for computations in parts b, c, and d!)
b.) Using your matrix from part a.), how many degrees of separation is Grace Kelly from Kevin Bacon? Assume same movie is 1 degree of separation, etc. Must show work using your matrix and must use a computer program like Octave or MatLab
c.) Is the graph in Figure 3 connected? Must show work using your matrix.
d.) What is the largest degree of separation (assume same movie is 1 degree of separation) for the graph in Figure 3 between Kevin Bacon and any other actor? Must show work using your matrix.

Kevin Bacon	Glenn Close
Bill Paxton	Nicole Kidman
Tom Hanks	John Gielgud
Donald Sutherland	Vernon Dobtcheff
John Belushi	Paul Herbert
Meryl Streep	Yves Aubert
Kate Winslet	Shane Zaza
Patrick Allen	Seretta Wilson
Grace Kelly	Kathleen Quinlan
Lloyd Bridges	



A tiny portion of the movie-performer relationship graph

Figure 3: Small performer-movie graph, taken from the website: http://introcs.cs.princeton.edu/java/45graph/