Random Walks on Graphs

1. Introduction to Graph Theory

The intuitive notion of a **graph** is a figure consisting of points and lines adjoining these points. More precisely, we have the following definition: A graph is a set of objects called **vertices** along with a set of **unordered** pairs of vertices called **edges**. Note that each edge in a graph has no direction associated with it. If we wish to specify a direction, then we use the notion of a **directed graph** or **digraph**. The definition of a digraph is the same as that of a graph, except the edges are **ordered** pairs of edges. If (u, v) is an edge in a digraph, then we say that (u, v) is an edge from u to v. We also say that if (u, v) is an edge in a graph or digraph then u is **adjacent** to v (and v is adjacent from u in a digraph). The **degree** of a vertex v in a graph is the number $\deg(v)$ of edges which contain v. In a digraph we define the **out degree** of v to be the number $\deg^+(v)$ of v to be the number of edges which start at v and the **in degree** of v to be the number $\deg^-(v)$ of edges which end at v.

Below are some examples of graphs and digraphs:

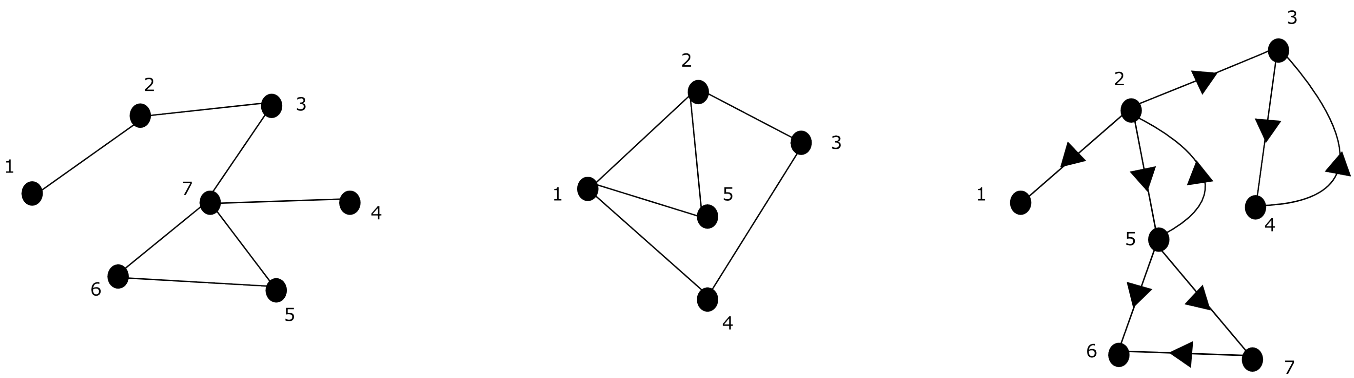

Figure 1

Note that it is possible for edges to cross even though this is not the case with the figures above. If two edges cross, then their intersection is a vertex only if indicated, otherwise their intersection is not a vertex.

A **walk** in a graph or digraph is a sequence of vertices $v_1, v_2, \ldots, v_{k+1}$, not necessarily distinct, such that (v_i, v_{i+1}) is an edge in the graph or digraph. The **length** of a walk is number of edges in the path, equivalently it is equal to k.

2. Random Walks on Graphs

Let G be a graph or digraph with the additional assumption that if G is a digraph, then $\deg^+(v) > 0$ for every vertex v. Now consider an object placed at vertex v_j. At each stage the object must move to an adjacent vertex. The probability that it moves to the vertex v_i is $\frac{1}{deg(v_j)}$ or $\frac{1}{deg^+(v_j)}$ if (v_j, v_i) is an edge on G and G is a graph or digraph, respectively. Otherwise the probability is 0. Therefore if we define

$$m_{ij} = \begin{cases} \frac{1}{deg(v_j)} \text{ if } (v_j, v_i) \text{ is an edge in the graph } G \\ \frac{1}{deg^+(v_j)} \text{ if } (v_j, v_i) \text{ is an edge in the digraph } G \\ 0 \text{ otherwise} \end{cases}$$

Then $M = (m_{ij})$ is a Markov matrix. Note that the roles of i and j are reversed as we need the columns of M to sum to 1. As each stage occurs, a sequence of adjacent vertices is produced. This sequence represents the position of the object at a given stage. Moreover this sequence is a walk in the graph. We call such a walk a **random walk** on the graph or digraph G. Using the Markov matrix, we see that the i, j entry of M^k represents the probability that a random walk of length k starting at vertex v_j, ends at the vertex v_i. The steady-state vector will correspond to the probability of being at a given vertex after a sufficient number of random walks.

Example:

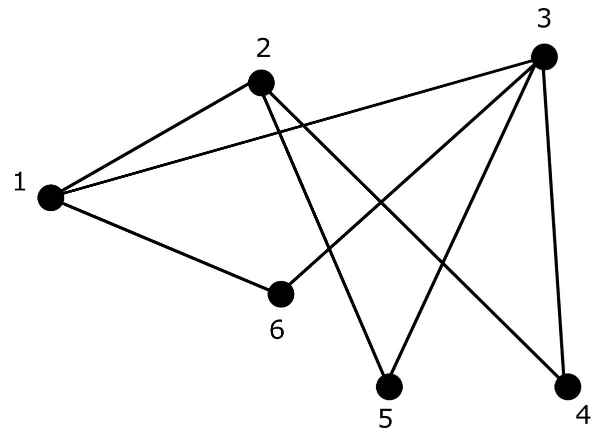

FIGURE 2

For the graph in Figure 2.

$$M = \begin{pmatrix} 0 & 1/3 & 1/4 & 0 & 0 & 1/2 \\ 1/3 & 0 & 0 & 1/2 & 1/2 & 0 \\ 1/3 & 0 & 0 & 1/2 & 1/2 & 1/2 \\ 0 & 1/3 & 1/4 & 0 & 0 & 0 \\ 0 & 1/3 & 1/4 & 0 & 0 & 0 \\ 1/3 & 0 & 1/4 & 0 & 0 & 0 \end{pmatrix}$$

Now we compute M^5 which represents corresponds to random walks of length 5 along this graph:

$$M^5 = \begin{pmatrix} 0.1192 & 0.2858 & 0.2582 & 0.0911 & 0.0911 & 0.1940 \\ 0.2858 & 0.0370 & 0.0723 & 0.3395 & 0.3395 & 0.1921 \\ 0.3442 & 0.0965 & 0.1273 & 0.4063 & 0.4063 & 0.2717 \\ 0.0608 & 0.2263 & 0.2032 & 0.0243 & 0.0243 & 0.1144 \\ 0.0608 & 0.2263 & 0.2032 & 0.0243 & 0.0243 & 0.1144 \\ 0.1293 & 0.1281 & 0.1359 & 0.1144 & 0.1144 & 0.1134 \end{pmatrix}$$

So the probability that an object which starts at vertex 3, makes 5 random moves and ends at vertex 1 is given by the $3, 1$ entry of M^5 which is 0.2582.

The steady-state vector is given by using:

$$\mathbf{x} = \mathrm{null}(\mathrm{M} - \mathrm{eye}(6), '\mathrm{r}')$$

$$\mathbf{x}_\mathrm{s} = \mathbf{x}/\mathrm{sum}(\mathbf{x})$$

which gives $\mathbf{x}_s = \begin{pmatrix} 0.1875 \\ 0.1875 \\ 0.2500 \\ 0.1250 \\ 0.1250 \\ 0.1250 \end{pmatrix}$ This indicates that no matter what the distribution of objects is on the graph; over time, 1/4 of them will accumulate at vertex 3.