An Application of Matrix Multiplication

V Yegnanarayanan

We are well aware of the ever increasing importance of graphical and matrix representations in applications to several day-to-day real life problems. The interconnectedness of the notion of graph, matrix, probability, limits, and system of equations are visible and approachable in the use of Markov Chains. We discuss here an interesting activity that involves the above concepts in the problem of weather pattern analysis.

V Yegnanarayanan is a Seniuor Professor, Department of Mathematics at Velammal Engineering College, Chennai. His research interests include graph theory, theoretical computer science, number theory, set theory, computer vision, artificial intelligence, image processing, mathematical modeling, optimization techniques, environmental science, mathematical linguistics, graph theory applications to biological networks, social networks, electrical engineering.

Keywords
Graphs, networks, probability, weather phenomena.

1. Introduction

The development of graph theory is very similar to the development of probability theory where much of the original work was motivated by efforts to understand games of chance [1, 2]. Large portions of graph theory have been motivated by the study of games and recreational mathematics. Graph is a very convenient and natural way of representing the relationships between objects through its elements. In problems like map coloring, signal-flow graphs, tracing maze, structure of chemical molecules and in many others, such a pictorial representation is all that we look for.

1.1 *Graphs, Digraphs, Walks and Paths*

A graph G is a triple consisting of a vertex set $V(G)$, an edge set $E(G)$, and a relation that associates with each edge two vertices (not necessarily distinct) called its endpoints. A graph $G = (V, E)$ is said to be directed if the edge set is composed of ordered vertex pairs and undirected if the edge set is composed of unordered vertex pairs.A walk of length k in a graph G is a succession of k edges of G of the form $uv, vw, wx, ..., yz$. A walk becomes a path if all its vertices and hence all its edges are distinct.

1.2 *Markov Chains*

The study of Markov chains has arisen in a wide variety of areas, ranging from genetics and statistics to computing and sociology [3]. Consider the following problem, that of a drunkard standing directly between his two favorite pubs, 'The Markov Chain' and 'The Source and Sink' (*Figure* 1).

Every minute he either staggers ten meters towards the first pub (with probability $\frac{1}{2}$) or towards the second pub (with probability $\frac{1}{3}$) or he stays where he is (with probability $\frac{1}{6}$); such a procedure is called a one-dimensional random walk. We assume that the two pubs are 'absorbing', in the sense that if he arrives at either of them he stays there. Given the distance between the two pubs and his initial position, we can ask which pub he is more likely reach

Let us suppose that the two pubs are 50 meters apart, and that our friend is initially 20 meters from 'The Source and Sink'. If we denote the places at which he can stop by $E_1, ..., E_6$, where E_1 and E_6 are the two pubs, then his initial position E_4 can be described by the vector $x = (0, 0, 0, 1, 0, 0)$, in which the ith component is the probability that he is initially at E_i. Furthermore, the probabilities of his position after one minute are given by the vector $(0, 0, \frac{1}{2}, \frac{1}{6}, \frac{1}{3}, 0)$, and after two minutes by $(0, \frac{1}{4}, \frac{1}{6}, \frac{13}{36}, \frac{1}{9}, \frac{1}{9})$. It is awkward to calculate directly the probability of his being at a given place after

Development of graph theory is very similar to the development of probability theory.

The study of Markov chains has arisen in a wide variety of areas, ranging from genetics and statistics to computing and sociology

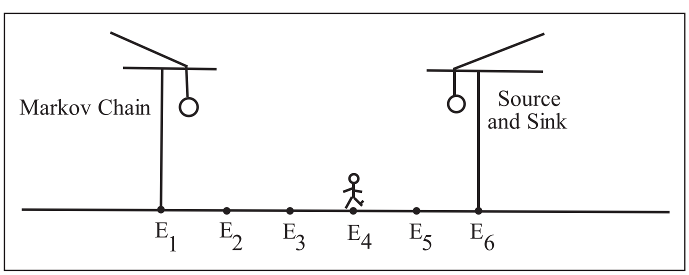


Figure 1.

We are mainly concerned with whether we can get from a given state to another state, and if so, how long it will take.

k minutes, and a more convenient way of doing this is by making use of the transition matrix, which is introduced here.

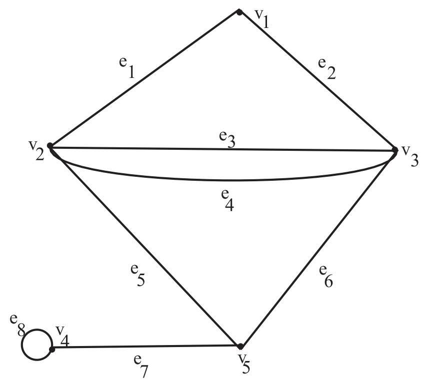
Let p_{ij} be the probability that he moves from E_i to E_j in one minute; for example $p_{23} = \frac{1}{3}$ and $p_{24} = 0$. These probabilities p_{ij} are called the transition probabilities, and the 6×6 matrix $P = (p_{ij})$ is the transition matrix:

$$P = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{6} & \frac{1}{3} & 0 & 0 & 0 \\ 0 & \frac{1}{2} & \frac{1}{6} & \frac{1}{3} & 0 & 0 \\ 0 & 0 & \frac{1}{2} & \frac{1}{6} & \frac{1}{3} & 0 \\ 0 & 0 & 0 & \frac{1}{2} & \frac{1}{6} & \frac{1}{3} \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}.$$

Note that each entry of P is non-negative and that the sum of the entries in each row is 1. If x is the initial row vector defined above, then the probabilities of his position after one minute are given by the row vector xP, and after k minutes by the vector xP^k. In other words, the ith component of xP^k represents the probability that he is at E_i after k minutes have elapsed. In general, we define a probability vector to be a row vector whose entries are all non-negative and have sum 1, and a transition matrix to be a square matrix, each of whose rows is a probability vector. We then define a finite Markov chain (or simply a chain) to consist of an $n \times n$ transition matrix P and a $1 \times n$ row vector x. The positions E_i are the states of the chain and our aim is to describe ways of classifying them. We are mainly concerned with whether we can get from a given state to another state, and if so, how long it will take.

2. Matrix Association with Graphs

We denote by $G(p,q)$ a graph G with p vertices and q edges.



	v_1	v_2	v_3	v_4	v_5
v_1	0	1	1	0	0
v_2	1	0	2	0	1
v_3	1	2	0	0	1
v_4	0	0	0	2	1
v_5	0	1	1	1	0

Figure 2.

2.1 *Adjacency Matrix*

Let G be the (p, q) graph shown in *Figure* 2. The adjacency matrix of a labeled graph G denoted by $A(G) = (a_{ij})$ is a $p \times p$ matrix defined by a_{ij} = the number of edges joining the vertex v_i to the vertex v_j, with the usual convention regarding loops. That is, the diagonal entry a_{ii} is twice the number of loops at vertex i. Therefore it is easy to see that A is a symmetric matrix, and if G is a simple graph, then all the entries of the main diagonal of A are 0; and if G has no multiple edges, then all the entries of A are either 1 or 0.

2.2 *Incidence Matrix*

Consider again the (p, q) graph G shown in *Figure* 2. The incidence matrix denoted by $M(G) = (m_{ij})$ is a $p \times q$ matrix defined by m_{ij} = the number of times the vertex v_i is incident with the edge e_j. That is, $m_{ij} = 0$, 1, 2, if e_j is not incident on v_i, e_j is an edge, e_j is a loop, respectively. (See *Figure* 3).

	e_1	e_2	e_3	e_4	e_5	e_6	e_7	e_8
v_1	1	1	0	0	0	0	0	0
v_2	1	0	1	1	1	0	0	0
v_3	0	1	1	1	0	1	0	0
v_4	0	0	0	0	0	0	1	2
v_5	0	0	0	0	1	1	1	0

Figure 3.

Our main concern is not with the actual probabilities p_{ij}, but with when they are non-zero.

2.3 *Markov Chains and Graphs*

Our main concern is not with the actual probabilities p_{ij}, but with when they are non-zero. To decide this, we represent the situation by a digraph whose vertices correspond to the states and whose arcs tell us whether we can go from one state to another in one minute. Thus, if each state E_i is represented by a vertex v_i, then we obtain the required digraph by drawing an arc from v_i to v_j if and only if $p_{ij} \neq 0$. Alternatively, we can define the digraph in terms of its adjacency matrix by replacing each non-zero entry of the matrix P by 1. We refer to this digraph as the associated digraph of the Markov chain.

If we are given a Markov chain whose transition matrix is

$$\begin{pmatrix} 0 & \frac{1}{4} & \frac{1}{2} & 0 & 0 & \frac{1}{4} \\ 0 & 1 & 0 & 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{3} & 0 & \frac{1}{12} & 0 & \frac{1}{12} \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix},$$

then its associated adjacency matrix is

$$\begin{pmatrix} 0 & 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 \end{pmatrix}.$$

The digraph is as shown in *Figure* 4.

If G is simple, then the entries on the diagonals of both MM^{T} and A^2 are the degrees of the vertices of G. (M^{T} is the transpose of M). To see this let $M = (a_{ij})$,

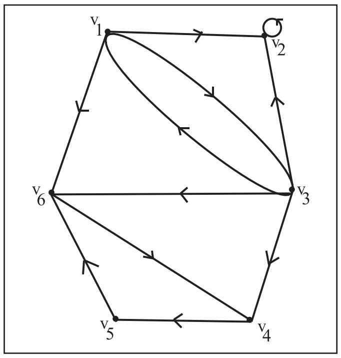


Figure 4.

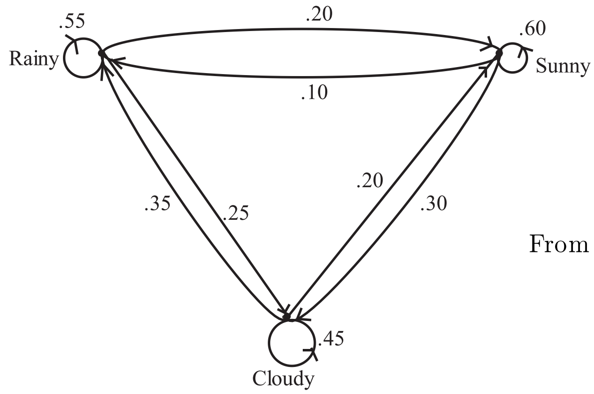
$M^{\mathrm{T}} = (b_{ij})$ and $A = (c_{ij})$. Then, the (i, i)th entry of $MM^{\mathrm{T}} = \sum_{j=1}^{q} a_{ij}b_{ji} = \sum_{j=1}^{q} a_{ij}a_{ij} = \sum_{j=1}^{q} a_{ij}^2 =$ degree of v_i (since G is simple, $a_{ij} = 1$ or 0). Next, the (i, i)th entry of $A^2 = \sum_{j=1}^{p} c_{ij}c_{ji} = \sum_{j=1}^{p} c_{ij}^2 =$ degree of v_i(since $c_{ij} = 1$ or 0).

If G is a labelled graph with adjacency matrix A then the number of v_iv_j walks of length k in G is the (i, j)th entry of A^k. This can be seen easily by induction on k. Since $A^1 = A$, the result is true for $k = 1$. Assume that the result is true for any walk of length less than k. Now $(A^k)_{ij} = \sum_{r=1}^{p} (A^{k-1})_{ir}a_{rj}$. Let us consider a typical term $(A^{k-1})_{ir}a_{rj}$ of the summation on the RHS. By induction hypothesis, $(A^{k-1})_{ir}$ gives the number of v_iv_r walks of length $k - 1$. If $a_{rj} = 1$, then this will also give the number of v_iv_r walks of length k in which v_r occurs as the last but one vertex. If $a_{rj} = 0$, there will be no such walk. Thus, the sum on the RHS gives the number of v_iv_j walks of length k.

A model which gives the probability of the occurrence of various weather phenomena several days into the future.

3. An Application of Matrix Multiplication

We construct a model which gives the probability of the occurrence of various weather phenomena several days into the future, based on today's weather conditions and the probability of the occurrence of various weather phenomena tomorrow.



$$\text{From} \begin{array}{c} \\ \end{array} \overset{\text{To}}{\begin{pmatrix} & C & S & R \\ C & .45 & .20 & .35 \\ S & .30 & .60 & .10 \\ R & .25 & .20 & .55 \end{pmatrix}} = W$$

Figure 5.

In trailing weather patterns and while collecting data on weather phenomena, probabilities of transition from one state to a finite number of other states emerge. For simplicity, we will consider three phenomena: cloudy days, C; rainy days, R; and sunny days, S. The representative digraph (directed graph) provides transition probabilities, day-to-day, from one weather state to another (see *Figure* 5). Each of the transition probabilities corresponds to an edge going from one type of weather to another. For example, the edge from cloudy to rainy (C, R), labeled .35, can be interpreted as '35% of cloudy days are followed by a rainy day'. Then the edge from rainy to sunny (R, S), labeled .20, means that 20% of rainy days are followed by a sunny day. Suppose today is rainy. What will the weather 'most likely' be two days from now?

From the graphical representation we compile the probabilities into the matrix W, with the 'from' states forming the rows and the 'to' states forming the columns (i.e., the entry in the ith row and the jth column gives the probability of transition *from* the weather state in the i^{th} row *to* the weather state in the j^{th} column.)

Note that each entry represents the probability of a transition from one state to another. Consequently, the entries are non-negative. Also note that since the probability of the occurrence of each weather pattern appears exactly once in each row, the sum of the entries in each row equals 1.0. A matrix having these characteristics is referred to as a stochastic matrix. We attempt to predict what will happen two days after a rainy day; we can follow the directed edges of the digraph and account for all possible compound events. For example, the event of Rainy → Cloudy → Sunny can be extracted from the graph as Rainy → Cloudy = .25, times the probability of the event from Cloudy → Sunny = .20. The probability of that 'path' occurring is .05. There do, however, exist other paths that begin with a Rainy day and conclude with a Sunny day two days later. Each one of these needs to be accounted for before predictions can be made. Through the use of matrix multiplication, all such possibilities can be accounted for.

We attempt to predict what will happen two days after a rainy day; we can follow the directed edges of the digraph and account for all possible compound events.

$$W^2 = W.W = \begin{pmatrix} .35 & .28 & .37 \\ .34 & .44 & .22 \\ .31 & .28 & .41 \end{pmatrix}.$$

The interpretation of the matrix W^2 is shown below:

$$\begin{array}{c} \text{To} \\ \text{(Two Days from now)} \end{array}$$

$$\text{From (Today)} \begin{pmatrix} & C & S & R \\ C & .35 & .28 & .37 \\ S & .34 & .44 & .22 \\ R & .31 & .28 & .41 \end{pmatrix} = W^2$$

Note that the entries in each row of W^2 still add up to 1.0 and that W^2 is also a stochastic matrix. The entries in the third row represent a transition from a Rainy day (today) to one of three weather states (indicated by their respective column headings) two days from now. For example, the entry $W_{32} = .28$ implies that if today is Rainy, then there is a 28% probability of a Sunny day occurring two days from now. This process

The matrix W^k gives the probabilities of the occurrence of each of the three weather states, k days in the future, given any of the three weather states occurring today.

can be repeated to determine whether it will be C or S or R, k days from the present. The matrix W^k gives the probabilities of the occurrence of each of the three weather states, k days in the future, given any of the three weather states occurring today. We illustrate this concept with a problem. Given that today is Monday and that it is Cloudy, what will the weather most likely be on Thursday, based on this model? To answer this question, we use the fact that we are interested in the weather three days from now. So we must compute the matrix W^3.

$$W^3 = W.W.W = \begin{pmatrix} .334 & .312 & .354 \\ .340 & .376 & .284 \\ .326 & .312 & .362 \end{pmatrix}.$$

The interpretation of the matrix W^3 is shown below:

$$\begin{array}{c} \\ \\ \text{From (Today)} \end{array} \begin{array}{c} \text{To} \\ \text{(Three Days from now)} \\ \begin{pmatrix} & C & S & R \\ C & .334 & .312 & .354 \\ S & .340 & .376 & .284 \\ R & .326 & .312 & .362 \end{pmatrix} = W^3 \end{array}$$

From row 1 (representing an initial state of Cloudy) we see that, on Thursday, there is a 33.4% chance that it will be Cloudy, a 31.2% chance that it will be Sunny and a 35.4% chance that it will be Rainy. An interesting phenomenon occurs as we project further into the future, applying the initial probabilities to indicate transitions. The rows of the matrix W^{30}, for example, are almost identical.

$$W^{30} = \begin{pmatrix} .\overline{3} & .\overline{3} & .\overline{3} \\ .\overline{3} & .\overline{3} & .\overline{3} \\ .\overline{3} & .\overline{3} & .\overline{3} \end{pmatrix}.$$

This is indicative of a stabilization occurring over the long term and a convergence of probabilities of states that can be seen in many physical phenomena, as in our case, weather on a particular day. Although weather is dependent on many factors, this model is based solely on using the first state as an indicator of a transition state. That is, we have made the assumption that proceeding from one weather state to the next is dependent only on the current state. This characteristic, along with the existence of a finite number of states, represents a Markov Chain.

Although weather is dependent on many factors, this model is based solely on using the first state as an indicator of a transition state.

Suggested Reading

[1] J M Aldous and R J Wilson, *Graphs and Applications*., Springer-Verlag, London, England, 2000.

[2] J A Bondy and U S R Murty, *Graph Theory with Applications*, Elsevier Science, North Holland, 1982.

[3] G Chartrand, *An Introductory Graph Theory*, Dover, New York, 1985.

[4] P G Hoel, C J Stone, S C Port, *Introduction to Stochastic Processes*, Waveland Press, Inc, USA, 1986.

Address for Correspondence
V Yegnanarayanan
Department of Mathematics
Velamman Engineering College
Chennai 600 066, India.
Email: prof.yegna@gmail.com