ADJACENCY MATRICES, MARKOV CHAINS

BENEDICT MORRISSEY

Contents

These are notes for part of L4 of Math 312 Linear algebra, Summer 2019.

Definition 0.1. A *directed graph* is a set $S_{Vertices}$ of vertices, and a set S_{edges} of edges, together with two functions $f_{out}, f_{in} : S_{edges} \to S_{vertices}$.

Graphically we can represent this by drawing the vertices as points, and and edge $e \in S_{edges}$ as an arrow from $f_{out}(e)$ to $f_{in}(e)$.

Informally we say that an edge e starts at $f_{out}(e)$ and ends at $f_{in}(e)$. We also say that an edge e is between the vertices $f_{out}(e)$ and $f_{in}(e)$

Example 0.2. See Figure 1.

Example 0.3.

- Consider the internet, form a graph where vertices are webpages, edges are hyperlinks.
- Consider a circuit, form a graph where vertices are junctions, edges are paths current can flow along.
- Consider a country, form a graph where vertices are cities, edges are roads. (NB. This is not really directed, but for each road that is not one way one can just put an edge in each direction).

Remark 0.4. The above examples suggest that we will also want additional information on top of just a graph.

Definition 0.5 (Adjacency Matrix of a directed graph)**.** For a directed graph with vertices labelled by the numbers $1, ..., n$, the *Adjacency matrix* of the graph is the matrix is given by

$$A = (a_{ij}) = \{\text{The number of edges starting at vertex } i \text{ and ending at vertex } j.\}$$

Example 0.6. Consider the graph of figure 0.6.

The adjacency matrix for this graph is

$$A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 3 & 0 & 0 \end{pmatrix}$$

Question 0.7. *How do we count the number of paths of length m from vertex i to vertex j?*

Let us first consider the case of length 2. A path of length 2 from vertex i to vertex j must be (for some vertex k) an edge from vertex i to vertex k followed by an edge from vertex k to vertex j. Fixing a vertex k the number of such paths is $a_{ik}a_{kj}$. Hence the total number of such paths is

$$\sum_{\text{vertices } k} a_{ik}a_{kj} = (A^2)_{ij}$$

That is the number of paths of length 2 from vertex i to vertex j is the element in the i^{th} row and j^{th} column of A^2.

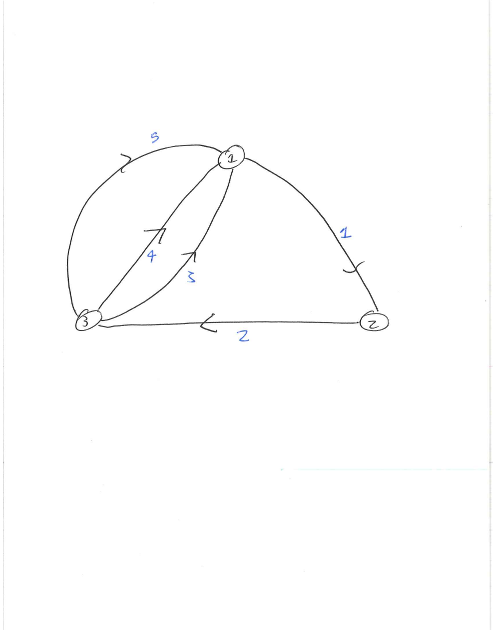


FIGURE 1. Graph 1: Edges are numbered in Blue

Claim 0.8. *The number of paths of length m from vertex i to vertex j is $(A^m)_{ij}$ – the entry in the i^{th} row and j^{th} column of A^m.*

Firstly note that this is true for $m = 0$ (As $A^0 = Id$), $m = 1$ and $m = 2$. Suppose that it is true for $m = k$, we show that it is true for $m = k + 1$:

A path of length m must be (for some vertex k) a path of length k from vertex i to vertex k followed by an edge from vertex k to vertex j. Fixing a vertex k the number of such paths is $(A)_{ik}a_{kj}$. Hence the total number of such paths is

$$\sum_{\text{vertices } k} (A)_{ik}a_{kj} = (A^{k+1})_{ij}.$$

Hence as we know it is true for $m = 2$, it is true for $m = 3$, hence it is also true for $m = 4$, hence it is true for $n = 5$,... Hence it is true for every natural number m.

Remark 0.9. This is an example of a *proof by induction.*

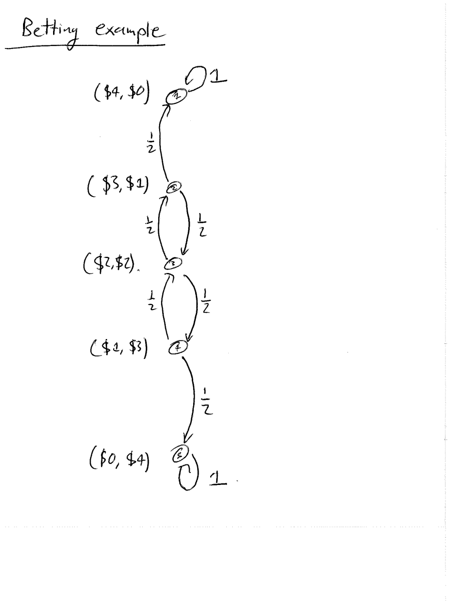
Definition 0.10 (Weighted Graph)**.** A *weighted directed graph*, is a directed graph $(S_{Vertices}, S_{edges}, f_{in}, f_{out})$, together with a function $w : S_{edges} \to \mathbb{R}$.

Intuitively this means it is a directed graph with each edge labelled by some real number $\mathbb{R}$.

Example 0.11 (Markov Chain). We consider some system with a set of states, that discretely changes from one state to another state. The vertices are the possible states. The edges are the possible transitions between states, labelled with the probability of this transition happening on a given timestep. When we say labelled we mean that the function w evaluated on an edge gives the probability of that transition occuring.

Example 0.12 (Betting example). Suppose there are two people C and B, betting (when they have a non-zero amount of money) on the outcomes of a series of coin tosses. If they have \$4 between them, the possible states of the system are $\{(\$0, \$4), (\$1, \$3), (\$2, \$2), (\$3, \$1), (\$4, \$1)\}$, where the state $(\$c, \$b)$ is the state of the system where person C has $\$c$, and person B has $\$b$. The time steps correspond to the coin tosses.

Figure 0.12 shows the weighted graph for this Markov chain:



Chain.pdf

FIGURE 2. Betting Markov Chain

Definition 0.13 (Transition Matrix)**.** The Transition matrix for a Markov Chain is the matrix

$$(T)_{ij} = \{\text{Probability of transitioning from state } j \text{ to state } i \text{ in a given time step}\}.$$

Warning 0.14. *Note that for an Adjacency matrix A the entry A_{ij} is related to edges from i to j, while for the Transition matrix T the entry T_{ij} is related to the edge from j to i in the weighted graph associated to the Markov chain.*

Example 0.15. For the Markov chain of example 0.12, the transition matrix is

$$T = \begin{pmatrix} 1 & \frac{1}{2} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} & 0 & 0 \\ 0 & 0 & 0 & \frac{1}{2} & 1 \end{pmatrix}$$

Question 0.16. *How does the* probability *of the system being in a given state evolve over timesteps?*

Let's consider the case of the Markov chain of example 0.12. Suppose that the system starts in state $(\$2, \$2)$. That is to say with probability one it is in the state $(\$2, \$2)$. After one coin has been tossed there is a probability of $1/2$ of being in each of the states $(\$1, \$3)$ and $(\$3, \$1)$.

More generally suppose that after m coin tosses the probability that the system is in state $(\$c, \$b)$ is $p_{(\$c,\$b).m}$. We can then describe this probability distribution over states via the vector

$$\vec{p_m} \begin{pmatrix} p_{(\$0,\$4),m} \\ p_{(\$1,\$3),m} \\ p_{(\$2,\$2),m} \\ p_{(\$3,\$1),m} \\ p_{(\$4,\$0),m} \end{pmatrix}.$$

We can now ask what the probability of the system being in the state $(\$c, \$b)$ after $m+1$ coin tosses (which we denote by $p_{(\$c,\$b),m+1}$).

- The probability of being in state $(\$0, \$4)$ after $m+1$ coin tosses is the sum of the probability of being in state $(\$0, \$4)$ after m coin tosses plus one half[1] times the probability of being in state $(\$1, \$3)$ after m coin tosses. As a formula:
$$p_{(\$0,\$4),m+1} = p_{(\$0,\$4),m} + \frac{1}{2} p_{(\$1,\$3),m}$$
- The probability of being in state $(\$1, \$3)$ after $m+1$ coin tosses is one half[2] time the probability of being in state $(\$2, \$2)$ after m coin tosses., That is
$$p_{(\$1,\$3),m+1} = \frac{1}{2} p_{(\$2,\$2),m}$$

Exercise 0.17. *Do the same computation for the remaing three states.*

This shows that the probability distribution evolves as follows:

$$\vec{p_{m+1}} = T\vec{p_m}.$$

This is true for a general Markov chain.

Question 0.18. *What is the long term behaviour of the probability distribution?*

We will only be able to make a partial answer to this with our current knowledge about linear algebra, but will return to this question next week after considering eigenvalues and eigenvectors.

Suppose that as m (the number of time steps, in our example the number of coin tosses) goes to infinity the probability distribution (over the different states) converges to a given probability distribution $\vec{p}$.

Then the above rule for evolution of probability distribution tells us that $\vec{p} = T\vec{p}$. This motivates the following definition:

Definition 0.19 (Steady State Probability Distribution)**.** A *Steady State probability distribution* (for a Markov chain with transition matrix T) is a vector $\vec{p}$ (with entries p_i) such that

- $$\vec{p} = T\vec{p},$$
- $$\sum_i p_i = 1,$$
- $$0 \leq p_i \leq 1$$

  for all entries p_i of the vector p.

Remark 0.20. Note that the last two conditions are ensuring that $\vec{p}$ is giving a probability distribution over states.

Remark 0.21. Note that we can solve for steady state probability distributions by solving the equation $(T - Id)\vec{p} = \vec{0}$.

Exercise 0.22. *Show that in example 0.12 the steady state probability distributions are given by*

$$\vec{p} = \begin{pmatrix} a \\ 0 \\ 0 \\ 0 \\ b \end{pmatrix}$$

for $a + b = 1$, $0 \leq a, b \leq 1$.

[1]The probability of person C winning the $(m+1)^{st}$ coin toss

[2]Again the probability of person C winning the $(m+1)^{st}$ coin toss.

Remark 0.23. At the moment we do *not* really understand the long term behaviour of a Markov chain. This is because if we are given an initital probability distribution we do not know

- If the probability distribution after m time steps converges as $m \to \infty$.
- If if does converge and there are multiple possible steady state probability distributions we don't know which one it converges to.