Eigenvectors Via Graph Theory

Jennifer Harris
Advisor: Dr. David Garth
October 13, 2009

1 Introduction

There is no problem in all mathematics that cannot be solved by direct counting.
-Ernst Mach

The goal of this paper is to use counting methods to derive a natural interpretation of eigenvectors of a matrix. We will assume the reader is familiar with linear algebra. Our method will be to study the adjacency matrices of simple graphs with no loops or parallel edges. The entries of these graphs will count the number of paths of a given length between pairs of matrices. We will then describe eigenvectors using this counting interpretation. From linear algebra, we will find eigenvectors by making use of the power method. This paper will tie together the classes of discrete math, in particular graph theory, and linear algebra.

2 Definitions

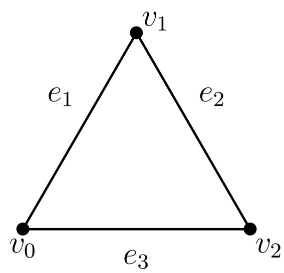
Generally speaking, a graph is a set of points, called vertices, with a set of connections between the points, called edges. We will start with a more formal definition. All of the following definitions are adapted from [3].

Definition 1. A **graph (or undirected graph)** G consists of a set V of vertices (or nodes) and a set E of edges (or arcs). Each edge e is a member of E and is associated with an unordered pair of vertices v and w. If there is a unique edge e associated with the vertices v and w, we write $e = (v, w)$, or $e = (w, v)$. In this context, (v, w) denotes an edge between v and w, and not an ordered pair. We say that e is **incident** on v and w, and v and w are **adjacent** vertices.

Definition 2. Let $G = (V, E)$ be a graph, and let $\{v_0, v_1, \ldots, v_n\} \subseteq V$. A **path** from v_0 to v_n of length n is a sequence of $n+1$ vertices and n edges beginning with vertex v_0 and ending with vertex v_n, $(v_0 e_1 v_1 e_2 v_2 \ldots v_{n-1} e_n v_n)$, in which edge e_i is incident on vertices v_{i-1} and v_i for $i = 1, \ldots, n$. In other words, start at vertex v_0; go along edge e_1 to v_1; go along edge e_2

to v_2; and so on. The **length** of a path is the number of times edges are traveled along from vertex to vertex to get from v_0 to v_n.

Example 1. Let G be the graph with vertex set $V = \{v_0, v_1, v_2\}$, and the edge set $E = \{e_1, e_2, e_3\}$, where $e_1=(v_0, v_1)$, $e_2=(v_1, v_2)$, $e_3=(v_0, v_2)$. We can associate a picture with the graph as follows.



Here we have an undirected graph $G = (V, E)$. The edge e_1 is *incident* on v_0 and v_1, and v_0 and v_1 are incident on e_1. So v_0 and v_1 are *adjacent vertices*. The edge e_2 is *incident* on v_1 and v_2, and visa-versa. Thus v_1 and v_2 are *adjacent*. Lastly e_3 is *incident* on v_0 and v_2 and visa-versa. So v_0 and v_2 are *adjacent*. A path from v_0 to v_0 can be listed in a sequence as $(v_0 e_1 v_1 e_2 v_2 e_3 v_0)$, which would have a length of 3. Likewise, a path from v_0 to v_2 could be listed as $(v_0 e_3 v_2)$, which would have a length of 1.

Definition 3. Distinct edges associated with the same pair of vertices are called **parallel edges**. An edge incident on a single vertex is called a **loop**. An **isolated vertex** is a vertex that is not incident on any edge. A graph with neither loops nor parallel edges is a **simple graph**.

Definition 4. The **complete graph** on n vertices, denoted K_n is the simple graph with n vertices in which there is an edge between every pair of distinct vertices.

Definition 5. A graph $G=(V, E)$ is **bipartite** if there exist subsets V_1 and V_2 (either possibly empty) of V such that $V_1 \bigcap V_2 \neq \emptyset$, $V_1 \bigcup V_2 = V$ and each edge in E is incident on one vertex in V_1 and one vertex in V_2.

Definition 6. The **complete bipartite graph** on m and n vertices, denoted $K_{m,n}$, is the simple graph whose vertex set is partitioned into sets V_1 with m vertices and V_2 with n vertices in which the edge set consists of all edges of the form (v_1, v_2) with $v_1 \in V_1$ and $v_2 \in V_2$.

One of our main tools for studying graphs will be a certain matrix associated with the graph.

Definition 7. An **adjacency matrix** is a matrix A whose rows and columns are labeled by the vertices in some order. The entry in *row* i, *column* j, denoted a_{ij}, is the number of edges incident on vertex i and vertex j.

Theorem 1. *If A is the adjacency matrix of a simple graph, the ij^{th} entry of A^n is equal to the number of paths of length n from vertex i to vertex j, $n = 1, 2, \ldots$*

The proof follows from the book by Johnsonbaugh (pages 353-354) [3]:

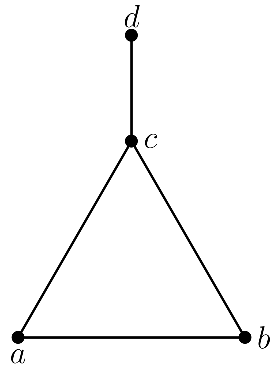
Proof. We will use induction on n. If $n = 1$, A^1 is simply A. The ij^{th} entry is 1 if there is an edge from i to j, which is a path of length 1, and 0 otherwise. Thus the theorem is true if $n = 1$. The basis step has been verified. Assume that the theorem is true for some $n \geq 1$. Let $\{v_1, \ldots, v_m\}$ be the vertices of G. Now $A_{n+1} = A^n A$ so that the ik^{th} entry in A_{n+1} is obtained by multiplying pairwise the elements in the i^{th} row $(s_1, s_2, \ldots, s_m)$ of A^n by the elements in the k^{th} column $(t_1, t_2, \ldots, t_m)^T$ of A and summing:

$$
\begin{aligned}
i^{th} \text{ row of } A^n &= (s_1, s_2, \ldots, s_m) \begin{pmatrix} t_1 \\ t_2 \\ \vdots \\ t_j \\ \vdots \\ t_m \end{pmatrix} \\
&= s_1 t_1 + s_2 t_2 + \ldots + s_j t_j + \ldots + s_m t_m \\
&= ik^{th} \text{ entry in } A^{n+1}.
\end{aligned}
$$

By induction, s_j gives the number of paths of length n from i to j in the graph G. Now, since G is a simple graph, t_j is either 0 or 1. If t_j is 0, there is no edge from j to k, so there are $s_j t_j = 0$ paths of length $n+1$ from i to k, where the last edge is (j, k). If t_j is 1, there is an edge from vertex j to vertex k. Since there are s_j paths of length n from vertex i to vertex j, there are $s_j t_j = s_j$ paths of length $n+1$ from v_i to v_k, where the last edge is (v_j, v_k). Summing over all j, we will count all paths of length $n+1$ from v_i to v_k. Thus the ik^{th} entry in A^{n+1} gives the number of paths of length $n+1$ from v_i to v_k, and the inductive step is verified. By the Principle of Mathematical Induction, the theorem is established. □

3 Examples, Discussion, and Conjectures

We will now use Theorem 1 to make some observations about the following example. Our goal is to use the counting interpretation of the adjacency matrix to derive a combinatorial interpretation of the eigenvalues of a matrix. G will refer to the following graph for the remainder of this section.



The adjacency matrix A of G is

$$A = \begin{bmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}.$$

The columns and rows are both labeled in increasing order by a, b, c, and d.

Now compute A^2 and A^3 to find the number of paths of length 2 and 3. The matrix A^2, which counts the number of paths of length 2 between pairs of vertices, is

$$A^2 = \begin{bmatrix} 2 & 1 & 1 & 1 \\ 1 & 2 & 1 & 1 \\ 1 & 1 & 3 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}.$$

Here is a table to show the paths of length 2 that have a as one of its endpoints. Since we are not dealing with a directed graph, without loss of generality we may view the paths as beginning at each of the vertices of a, b, c, d, and ending at a.

Beginning Vertex	**End Vertex**	**Paths of Length 2**	**Number of Paths of Length 2**
a	a	aba, aca	2
b	a	bca	1
c	a	cba	1
d	a	dca	1

Adding the number of paths up in column 4 of the table (column 1 of the matrix) (2+1+1+1) shows that there are 5 paths of length 2 in G that end at a. Also notice how this is the sum of the first row as well.

Now look at A^3, which counts the paths of length 3 ending at a.

$$A^3 = \begin{bmatrix} 2 & 3 & 4 & 1 \\ 3 & 2 & 4 & 1 \\ 4 & 4 & 2 & 3 \\ 1 & 1 & 3 & 0 \end{bmatrix}$$

The following table lists out the actual paths of length 3 in G.

Beginning Vertex	**End Vertex**	**Paths of Length 3**	**Number of Paths of Length 3**
a	a	*abca*, *acba*	2
b	a	*baca*, *baba*, *bcba*	3
c	a	*caba*, *caca*, *cdca*, *cbca*	4
d	a	*dcba*	1

Here there are 10 paths of length 3 that end at a. Similarly, there are 10 that end at b, 13 at c, and 5 at d, which we obtain by adding the entries in the corresponding columns of the adjacency matrix.

Let $\mathbf{e}_i$ be the i^{th} standard basis vector. From linear algebra we know that $A^n\mathbf{e}_i$ is the i^{th} column of A^n. These observations lead to the following corollary of Theorem 1. The proof of the corollary comes automatically from the theorem.

Corollary 1. *In a simple graph, the number of paths of length n having vertex i as an endpoint is the sum of the entries of $A^n\mathbf{e}_i$.*

Now let's consider the question: What proportion of paths of length n in G begin at each of the vertices a, b, c, and d and end at a?

To do this, first add the entries in column one of A^n. This gives the total number of paths of length n in G that have a as an endpoint. Then divide each entry by the total, which results in the proportions of paths that have each vertex as the other endpoint. The following table shows the proportions for A^3.

Beginning Vertex	**Proportion**
a	0.20
b	0.30
c	0.40
d	0.10

We will now continue doing this for $A^n \mathbf{e}_1$ arriving at the following table.

Beginning Vertex	n=**5**	n=**6**	n=**7**	n=**8**
a	0.25	0.280374	0.26158	0.273453
b	0.270835	0.271028	0.267544	.0271457
c	0.354167	0.28972	0.3333	0.303393
d	0.125	0.158879	0.135965	0.151697

Beginning Vertex	n=**9**	n=**10**	n=**30**	n=**50**
a	0.267161	0.271143	0.269595	0.269594
b	0.268089	0.270718	0.269595	0.269594
c	0.323748	0.309817	0.315466	0.315449
d	0.141002	0.148321	0.145364	0.145362

From the table it appears that the proportions converge to a limiting frequency. If they truly do converge, then the limiting frequencies represent the proportion of paths of infinite length in G that begin at each vertex. So looking at all of the paths of infinite length in G, about 27 percent of them begin at a and b, about 32 percent begin at c, and about 15 percent begin at d.

We will leave this for the moment and find the eigenvalues of A, and an eigenvector associated with with each one. We used Mathematica to do this.

Eigenvalue	**Eigenvector**
2.17009	[1.85464, 1.85464, 2.17009, 1]
-1.48119	[0.566968, 0.596968, -1.48119, 1]
-1	[-1, 1, 0, 0]
0.31108	[-0.451606, -0.451606, 0.311108, 1]

Notice that the eigenvectors and eigenvalues are all real. This is the consequence of the spectral threorem of linear algebra [5] since A is a symmetric matrix. We will discuss this more later.

Now we will take the eigenvector corresponding to the largest eigenvalue

$$[1.85465, 1.85454, 2.71009, 1]$$

and sum up the entries:

$$1.85465 + 1.85454 + 2.71009 + 1 = 6.87936.$$

Next we divide the entries of the eigenvector by this total resulting with:

$$0.2695, 0.2695, 0.3154, 0.1453$$

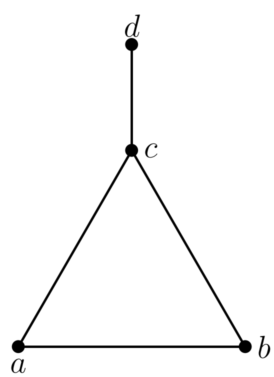
Notice that these appear to be the limiting frequencies from the table. This leads us to make the following conjecture.

Conjecture The entries of the eigenvector that corresponds to the largest eigenvalue, normalized so the sum of the entries is 1, are the proportion of the paths of infinite length in the graph that begin at each of the vertices in the graph.

In the next section we make the conjecture more precise and prove it as a theorem.

4 Proof of the Main Theorem

Before we state and prove the conjecture of the last section as a theorem, we need to formally define some of our concepts. First we define the notion of an infinite path in a graph. Then we define what we mean by the proportion or frequency of infinite paths in a graph beginning at a particular vertex. We will restrict our attention to simple graphs.



In definition 3 we defined a *finite path* from v_0 to v_n as a sequence of n+1 vertices. We will now define an *infinite path*. It seems natural for us to define an infinite path in a graph in the following way.

Definition 8. Let G be a simple graph with vertex set V and edge set E. An **infinite path** beginning at vertex v_1 is an infinite string of symbols $v_1 v_2 v_3 \ldots$, where each $v_i \in V$, and for all i, $(v_i,\, v_{i+1}) \in E$.

As an example, $ababab\ldots$ is an infinite path in the above graph, but $adadad\ldots$ is not. We now define *frequency* of an infinite path.

Definition 9. Let G be a simple graph with verticies $\{y_1, \ldots, y_m\}$. Let A be the adjacency matrix of G. For $n \geq 1$, let S_n be the sum of the entries of the vector $A^n\mathbf{e}_1$, where $\mathbf{e}_1$ is the first standard basis vector. If the limit

$$\lim_{n\to\infty} \frac{1}{S_n} A^n \mathbf{e}_1$$

exists, then the i^{th} coordinate of the limit is the **frequency, or proportion, of infinite paths in G that begin at vertex** i. If the limit does not exist, we say that the frequency of infinite paths beginning at v_i does not exist.

The last section showed a graph where the frequencies of infinite paths beginning at each vertex existed. In the next section, we will give an example where they do not exist.

We are now ready to state the conjecture of the previous section as a theorem. In order to be able to prove the theorem, we need to impose certain conditions on the eigenvalues and eigenvectors of the adjacency matrix A. We will show how these conditions can follow from restrictions imposed on the graph G in the next section.

Theorem 2. *Let G be a simple graph with adjacency matrix A. Let $\{y_1, \ldots, y_m\}$ be vertices of G. Assume that A has a real eigenvalue λ that is larger than any other eigenvalue. Assume also that the eigenvectors $\{v_1, \ldots, v_m\}$ of A form a basis for $\mathbb{R}^m$. Assume also that the proportion of paths of infinite length in G beginning at each vertex exists. Then the eigenvector of A corresponding to λ, normalized so that the sum of its coordinates is 1, represents the proportion of paths of infinite length in G beginning at each vertex.*

As far as we know, this theorem is original. However, it is analogous to theorem 8.4.6 of [1] that deals with a different topic, not graphs.

Proof. Assume frequencies exist. Let S_n be the sum of entries of $A^n\mathbf{e}_1$. So there exists a vector $\mathbf{w}$ such that

$$\frac{1}{S_n} A^n \mathbf{e}_1 \to \mathbf{w} \text{ as } n \to \infty.$$

So if $n \to \infty$, the paths beginning at y_i and never ending. The goal is to show that $\mathbf{w}$ is the the eigenvector of A corresponding to λ. Since

$$\frac{1}{S_n} A^n \mathbf{e}_1 \to \mathbf{w}$$

it follows that

$$A\left(\frac{1}{S_n} A^n \mathbf{e}_1\right) \to A\mathbf{w}. \tag{1}$$

Notice that

$$A\left(\frac{1}{S_n} A^n \mathbf{e}_1\right) = \frac{1}{S_n} A^{n+1}\mathbf{e}_1 = \frac{S_{n+1}}{S_n} \cdot \frac{1}{S_{n+1}} A^{n+1}\mathbf{e}_1. \tag{2}$$

We will show that $\frac{S_{n+1}}{S_n} \to \lambda$. Suppose for the moment that this is true. Then by (2), since $\frac{1}{S_{n+1}}A^{n+1}\mathbf{e}_1 \to \mathbf{w}$, we have that $A(\frac{1}{S_n}A^n\mathbf{e}_1) \to \lambda\mathbf{w}$. It will follow then from (1) that $A\mathbf{w} = \lambda\mathbf{w}$. Therefore, all we have to do is prove that $\frac{S_{n+1}}{S_n} \to \lambda$.

Let $\lambda_1, \ldots, \lambda_m$ be the eigenvalues of A, where the largest eigenvalue λ is λ_1. Let $\mathbf{v}_1, \ldots, \mathbf{v}_m$ be the corresponding eigenvectors. Since the eigenvectors are a basis of $\mathbb{R}^m$, we can write

$$\mathbf{e}_1 = c_1\mathbf{v}_1 + \ldots + c_m\mathbf{v}_m.$$

Thus, we have the formula

$$A^n\mathbf{e}_1 = c_1\lambda_1^n\mathbf{v}_1 + \ldots + c_m\lambda_m^n\mathbf{v}_m. \tag{3}$$

which is the same as the equation 10.98 of [6]. We will essentially use the power method to find λ. See [6] for a discussion of the power method. So S_n is the sum of the coordinates of $(c_1\lambda_1^n\mathbf{v}_1 + \ldots + c_m\lambda_m^n\mathbf{v}_m)$. Notice that the right hand side of (3) is

$$\lambda_1^n \left[c_1\mathbf{v}_1 + c_2\left(\frac{\lambda_2}{\lambda_1}\right)^n \mathbf{v}_2 + \ldots + c_m\left(\frac{\lambda_m}{\lambda_1}\right)^n \mathbf{v}_m \right]$$

Since λ_1 is the largest eigenvalue, we know that

$$\left(\frac{\lambda_i}{\lambda_1}\right)^n \to 0 \text{ for } 2 \le i \le m.$$

Let E_n be the sum of the coordinates of $c_1\mathbf{v}_1 + c_2(\frac{\lambda_2}{\lambda_1})^n\mathbf{v}_2 + \ldots + c_m(\frac{\lambda_m}{\lambda_1})^n\mathbf{v}_m$ and let E be the sum of the coordinates of $c_1\mathbf{v}_1$. Thus $E_n \to E$ as $n \to \infty$. Hence

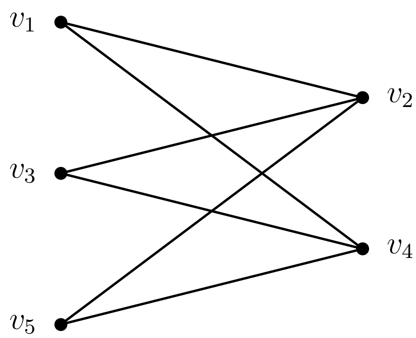
$$\frac{S_{n+1}}{S_n} = \frac{\lambda_1^{n+1}E_{n+1}}{\lambda_1^n E_n} \to \lambda$$

Thus we have shown that $A\mathbf{w} = \lambda\mathbf{w}$ where $\mathbf{w}$ is the eigenvector of A that corresponds to λ.
□

Remark: We could have chosen $\mathbf{e}_2, \ldots, \mathbf{e}_m$ in the proof of the previous theorem. In each case, $\frac{1}{S_n}A^n\mathbf{e}_i \to \mathbf{w}$.

5 Refinement of the Hypothesis of the Main Theorem

We will now work on stating the hypothesis of the theorem in terms of the graph G rather than the adjacency matrix of the graph. In order to get a better insight into this goal, we investigate some graphs for which the results of the theorem are not true. For example, look at the following complete bipartite graph with vertex sets $V_1 = \{v_1, v_3, v_5\}$ and $V_2 = \{v_2, v_4\}$.



The adjacency matrix is

$$A = \begin{bmatrix} 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

Notice the powers of A.

$$A^2 = \begin{bmatrix} 2 & 0 & 2 & 0 & 2 \\ 0 & 3 & 0 & 3 & 0 \\ 2 & 0 & 2 & 0 & 2 \\ 0 & 3 & 0 & 3 & 0 \\ 2 & 0 & 2 & 0 & 2 \end{bmatrix} \qquad A^3 = \begin{bmatrix} 0 & 6 & 0 & 6 & 0 \\ 6 & 0 & 6 & 0 & 6 \\ 0 & 6 & 0 & 6 & 0 \\ 6 & 0 & 6 & 0 & 6 \\ 0 & 6 & 0 & 6 & 0 \end{bmatrix} \qquad A^4 = \begin{bmatrix} 12 & 0 & 12 & 0 & 12 \\ 0 & 18 & 0 & 18 & 0 \\ 12 & 0 & 12 & 0 & 12 \\ 0 & 18 & 0 & 18 & 0 \\ 12 & 0 & 12 & 0 & 12 \end{bmatrix}$$

It is not hard to see that in general

$$A^{2n+1} = \begin{bmatrix} 0 & 6^n & 0 & 6^n & 0 \\ 6^n & 0 & 6^n & 0 & 6^n \\ 0 & 6^n & 0 & 6^n & 0 \\ 6^n & 0 & 6^n & 0 & 6^n \\ 0 & 6^n & 0 & 6^n & 0 \end{bmatrix}$$

and

$$A^{2n} = \begin{bmatrix} 2(6)^{n-1} & 0 & 2(6)^{n-1} & 0 & 2(6)^{n-1} \\ 0 & 3(6)^{n-1} & 0 & 3(6)^{n-1} & 0 \\ 2(6)^{n-1} & 0 & 2(6)^{n-1} & 0 & 2(6)^{n-1} \\ 0 & 3(6)^{n-1} & 0 & 3(6)^{n-1} & 0 \\ 2(6)^{n-1} & 0 & 2(6)^{n-1} & 0 & 2(6)^{n-1} \end{bmatrix}.$$

If P_n is the proportion of paths of length n in G that are between $\mathbf{v}_1$ and $\mathbf{v}_2$ we have the following table.

n	P_n
1	$\frac{1}{2}$
2	0
3	$\frac{6}{12}$
4	0
5	$\frac{36}{72}$

So in general, $P_n = \frac{1}{2}$ if n is odd, and $P_n = 0$ if n is even. Thus the frequencies do not converge.

These matrices show that for odd lengths of paths, odd powers of adjacency matrices (A, A^3, ...), there will not be any paths from vertices in V_1 back to any vertices in V_1, but there will be paths from vertices in V_1 to vertices in V_2. The even powers of adjacency matrices (A^2, A^4, ...) for even lengths of paths show the opposite; there are no paths from vertices in V_1 to V_2, but there are paths from vertices in V_1 to V_1. This shows that the frequencies of paths do not converge in a complete bipartite graph because each time the coordinates of $A\mathbf{e}_1$ alternate between zero and nonzero values.

We want a condition on the graph which guarantees that the theorem holds. From this example it seems we need, for n large enough, the graph to contain a path of length n between any pair of vertices. Therefore, we need to modify the theorem. We must add the assumption that G has a path of length n between every pair of vertices, for n large enough. The adjacency matrix A^n will not have any zeros when this happens. This condition guarantees that for n large enough, A^n will have only positive entries. We will now give some definitions and theorems from matrix theory that give the same result. These theorems and defintions are taken from [2] and [5].

For a matrix A, the spectrum of A is the set of eigenvalues of A, denoted $\sigma(A)$.

Definition 10. The algebraic multiplicity of λ is the number of times it is repeated as a root of the characteristic polynomial. In other words, $alg\ mult_A(\lambda_i) = a_i$ if and only if $(x - \lambda_1)^{a_1}$ ... $(x - \lambda_s)^{a_s}$ is the characteristic equation for A. When $alg\ mult_A(\lambda) = 1$, λ is called an **algebraically simple eigenvalue**.

Definition 11. The geometric multiplicity of λ is $\dim N(A - \lambda\mathbf{I})$. In other words, $geomult_A\lambda$ is the maximal number of linearly independent eigenvectors associated with λ. When $geomult_A\lambda = 1$, λ is called a **geometrically simple eigenvalue**.

Definition 12. A **permutation matrix** P is a square matrix, all of whose entries are 0 or 1; in each row and column of P there is precisely one 1.

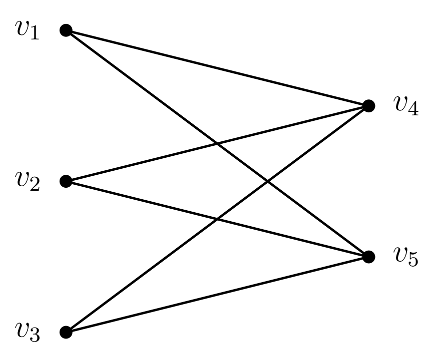
P is obtained from I by permuting its rows. The transformation P^TAP interchanges the i^{th} and j^{th} columns as well as the i^{th} and j^{th} rows of A. [2] For example, if

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

and A is the adjacency matrix for the complete bipartite graph on 2 and 3 vertices mentioned earlier,

$$P^TAP = \begin{bmatrix} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 \end{bmatrix}.$$

Notice that this is the same adjacency matrix we would get if we relabled the vertices as



.

Definition 13. $A_{n \times n}$ is said to be a **reducible matrix** when there exists a permutation matrix P such that

$$P^TAP = \begin{pmatrix} X & Y \\ \mathbf{0} & Z \end{pmatrix}, \text{ where } X \text{ and } Z \text{ are both square.}$$

Otherwise A is said to be an **irreducible matrix**.

If a graph is disconnected, we can label the vertices in such a way that the adjacency matrix has the form

$$\begin{pmatrix} \mathrm{X} & \mathrm{Y} \\ \mathbf{0} & \mathrm{Z} \end{pmatrix}.$$

So the adjacency matrix of a disconnected graph is reducible. If the graph is connected, the adjacency matrix will be irreducible.

Definition 14. For square matrices A, the number $\rho(A) = \max_{\lambda \in \sigma(A)} |\lambda|$ is called the **spectral radius**.

$A \geq \mathbf{0}$ means that all the entries in the matrix A are nonnegative.

Theorem 3. Perron-Frobenius Theorem [5]

If $A_{n \times n} \geq \mathbf{0}$ is irreducible, then each of the following is true.

- $\rho(A) \in \sigma(A)$ *and* $\rho(A) > 0$.
- *alg* $mult_A(\rho(A)) = 1$
- *There exists an eigenvector* $\mathbf{x} > \mathbf{0}$ *such that* $A\mathbf{x} = \rho(A)\mathbf{x}$.

From the Perron-Frobenius Theorem we can see that there is a unique real eigenvalue of largest absolute value.

Definition 15. A is **primitive** if it is irreducible, nonnegative, and has only one eigenvalue of maximum modulus. [6]

Theorem 4. *A is primitive if and only if all the entries of A^m are positive for some $m \geq 1$.[6]*

So if G has the property that for n large enough there exists a path of length n between any pair, then $A^n > \mathbf{0}$ for n large enough, and thus A is irreducible. A graph with this property is called strongly connected [2]. Thus the hypothesis that A has a largest eigenvector in Theorem 2 will be satisfied. Note that the adjacency matrix of the complete bipartite graph is irreducible, but not primitive.

Another hypothesis from Theorem 2 that we must now also address is that the eigenvectors of A form a basis for $\mathbb{R}^m$. Notice that it follows directly from the definition of an adjacency matrix that A is a symmetric matrix such that $A^T = A$. This is because entry a_{ij} of A counts the number of paths between vertex i and vertex j in A. Clearly this is the same as a_{ij}. Thus, $a_{ij} = a_{ij}$ in A, so $A^T = A$.

It follows directly from the spectral theorem from linear algebra that the eigenvectors form a basis for $\mathbb{R}^m$.

Theorem 5. Spectral Theorem [4]
An $n \times n$ symmetric matrix A has the following properties:
- *A has n real eigenvalues, counting multiplicities.*
- *The dimension of the eigenspace for each eigenvalue λ equals the multiplicity of λ as a root of the characteristic equation, i.e. the geometric multiplicity of each multiplicity equals its algebraic multiplicity.*

Now we are ready to restate Theorem 2.

Theorem 6. *Let G be a simple graph with adjacency matrix A. Assume that G is strongly connected. Also assume the proportions of paths beginning at each vertex exist. Then these proportions are given by the eigenvector of A corresponding to λ, the largest eigenvalue of A normalized so that the sum of its coordinates is 1.*

References

[1] Allouche JP. and Shallit J., *Automatic Sequences, Theory, Applications, and Generalizations*, Cambridge University Press, Cambridge, 2003.

[2] Horn R. and Johnson C., *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[3] Johnsonbaugh, R., *Discrete Mathematics, Sixth Edition*, Pearson, Prentice Hall, Upper Saddle River, New Jersey, 2005.

[4] Lay, David, *Linear Algebra and its Applications, Third Ed.*, Addison Wesley, 2003.

[5] Meyer, C., *Matrix Analysis and Applied Linear Algebra*, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2000.

[6] Olver P. and Shakiban C., *Applied Linear Algebra*, Prentice Hall, Upper Saddle River, NJ, 2006.