GRAPH THEORY AND LINEAR ALGEBRA

Kyrie Johnson

May 3, 2017

Abstract

Graphs are an incredibly versatile structure insofar as they can model everything from the modernity of computer science and complexity of geography, to the intricacy of linguistic relationships and the universality of chemical structures. Representing such graphs as matrices only enhances the computational aspects of this modeling. Ultimately, this necessitates linear algebra.

This paper explores the relationships between graph theory, their associated matrix representations, and the matrix properties found in linear algebra. It explores not only the adjacency matrices of graphs, but also the more interesting examples found in incidence matrices, path matrices, distance matrices, and Laplacian matrices. Investigations include the utility of such matrix representations for various classes of graphs, including disconnected graphs, complete graphs, and trees. In order to achieve this goal, this paper presents some of the most interesting theorems regarding matrix representations of graphs, and ties these theorems back to questions in graph theory itself.

Lastly, this paper identifies certain unique properties of special classes of graphs – namely, complete graphs and acyclic graphs (trees) – and how their specialty in graph theory reflects in their matrix properties. Once again, this analysis uses linear algebra. Many of such theorems suggest special relationships for these classes, some of which have already been investigated, while others remain to be fully understood.

Contents

1 Basic Graph Theory

Graph theory investigates the structure, properties, and algorithms associated with graphs. Graphs have a number of equivalent representations; one representation, in particular, is widely used as the primary definition, a standard which this paper will also adopt.

A **graph**, denoted G, is defined as an ordered pair composed of two distinct sets:

1. A set of vertices, denoted $V(G)$
2. A set of edges, denoted $E(G)$

The **order** of a graph G refers to $|V(G)|$ and the **size** of a graph G refers to $|E(G)|$. In other words, order refers to the number of vertices and size refers to the number of edges.

In order to perform compuptations with these graphs, we utilize matrices as an incredibly valuable, alternative representation. Such representations include incidence, adjacency, distance, and Laplacian matrices.

2 Adjacency Matrices

2.1 Definition

For a graph G of order n, the adjacency matrix, denoted $A(G)$, of graph G is an n by n matrix whose (i,j)-th entry is determined as follows:

$$A_{ij} = \begin{cases} 1, & \text{if vertex } v_i \text{ is adjacent to vertex } v_j \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

Adjacency matrices not only encapsulate the structure and relationships of a graph, but also provide for an efficient method of storage and access in a computer. For this reason, adjacency matrices are one of the most common ways of representing graphs.

2.2 Distance and Powers of A

The **distance** between vertices v_i and v_j, denoted $d(i, j)$, of a graph G is defined by the path of minimum length between the two vertices. For example, take the graph in Figure 1. There are two paths of length 4 between vertices v_6 and v_8 as depicted in Figure 2.

The **adjacency matrix** of a graph provides a method of counting these paths by calculating the powers of the matrices.

Theorem 2.1. *Let G be a graph with adjacency matrix A and k be a positive integer. Then the matrix power A^k gives the matrix where A_{ij} counts the the number of paths of length k between vertices v_i and v_j.*

For example, return to the graph shown in Figure 2. Equation 2 depicts the adjacency matrix of this graph, $A(G)$, and its fourth power.

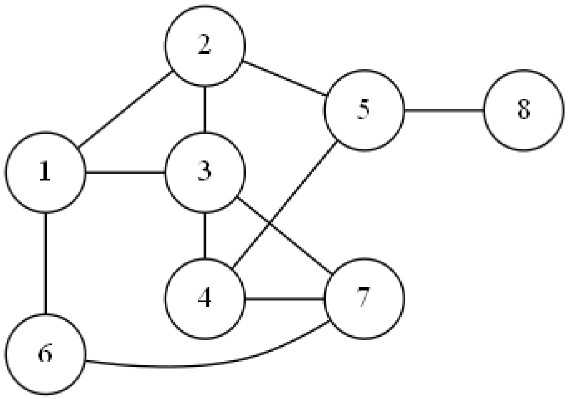

Figure 1: Graph of Order 8

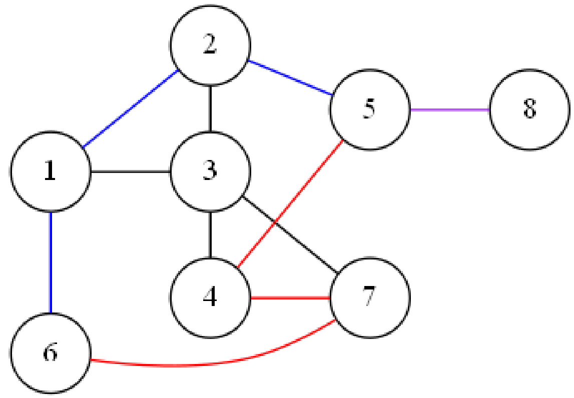

Figure 2: Depiction of 2 Paths Between Vertices 6 and 8

$$A(G) = \begin{bmatrix} 0 & 1 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{bmatrix} \quad A^4(G) = \begin{bmatrix} 17 & 11 & 13 & 11 & 10 & 4 & 16 & 2 \\ 11 & 18 & 13 & 17 & 4 & 9 & 11 & 6 \\ 13 & 13 & 28 & 13 & 16 & 14 & 13 & 2 \\ 11 & 17 & 13 & 18 & 4 & 9 & 11 & 6 \\ 10 & 4 & 16 & 4 & 15 & 4 & 10 & 0 \\ 4 & 9 & 14 & 9 & 4 & 10 & 4 & 2 \\ 16 & 11 & 13 & 11 & 10 & 4 & 17 & 2 \\ 2 & 6 & 2 & 6 & 0 & 2 & 2 & 3 \end{bmatrix} \tag{2}$$

As expected, the (6,8)-entry of $A^4(G)$ counts the two paths between vertices v_6 and v_8 – by symmetry, the same is true of the (8,6)-entry. While this result provides a useful method of counting the number of paths between vertices in a graph, it can be extended to a similarly interesting and logical corollary.

Corollary 2.1. *Let G be a graph with adjacency matrix A and k be a positive integer. Then the sum S_n defined by*

$$S_n = A + A^2 + ... + A^k \tag{3}$$

is the n x n matrix whose (i,j)-entry counts the number of paths of length k or less between vertices v_i and v_j.

Herein matrices once again provide a computationally useful tool for identifying and organizing the properties of a graph.

2.3 Eigenvalues and Complete Graphs

A **complete graph** G of order n, denoted K_n, includes an edge between every pair of vertices. For example, K_5 is depicted in Figure 3.

A **complete bipartite graph** G of order $n = p + q$, denoted $K_{p,q}$, consists of two sets $U, V \subset E(G)$ for which every vertex of U is adjacent to every vertex of V, and no two vertices within the same set are adjacent. For example, $K_{2,3}$ is depicted in Figure 4.

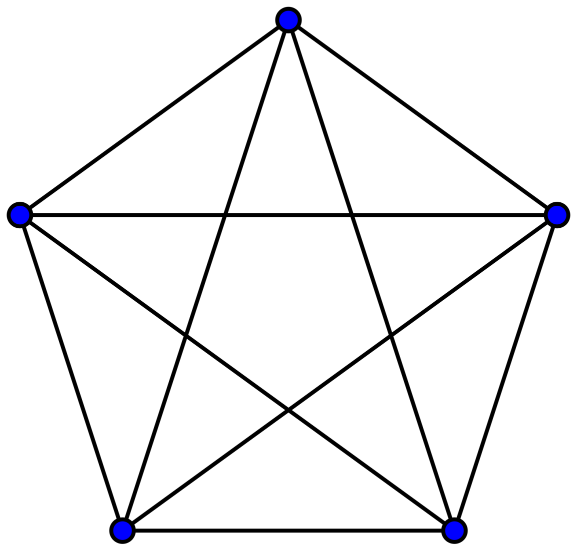

Figure 3: Complete Graph K_5

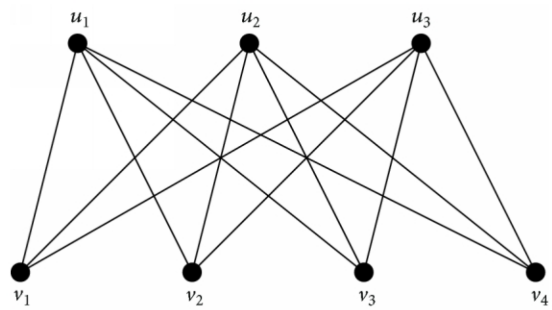

Figure 4: Complete Bipartite Graph $K_{3,4}$

The **eigenvalues of a graph** G are the eigenvalues of its adjacency matrix. In the case of complete graphs – both complete and complete bipartite – some interesting patterns emerge.

Theorem 2.2. *For any positive integer n, the eigenvalues of K_n are $n-1$ with multiplicity 1, and -1 with multiplicity n - 1. For any positive integer p, q, the eigenvalues of $K_{p,q}$ are $\sqrt{pq}$, $-\sqrt{pq}$, and 0 with multiplicity $p+q-2$.*

While this result is interesting in its own right, this theorem can be used to interweave a basic result from graph theory with one in linear algebra.

In graph theory, the removal of any vertex – and its incident edges – from a complete graph of order n results in a complete graph of order $n-1$. Combining this fact with the above result, this means that every $n-k+1$ square submatrix, $1 \leq k \leq n$, of $A(K_n)$ possesses the eigenvalue -1 with multiplicity k and the eigenvalue $n-k+1$ with multiplicity 1. Notably, this produces a natural bound on the eigenvalues of the submatrices of $A(K_n)$:

$$-1 \leq \lambda_i \leq n \ , \quad 1 \leq i \leq n$$

This result also emerges in linear algebra, as a general property of symmetric matrices: *interlacing eigenvalues* occurs for all symmetric matrices, a linear algebra equivalent to this natural result in graph theory.

3 Incidence Matrices

3.1 Definition

For a graph G of order n and size m, the incidence matrix, denoted $Q(G)$, of G is the n by m matrix whose (i,j)-th entry is determined as follows:

$$Q_{ij} = \begin{cases} 1, & \text{if vertex } v_i \text{ is incident to edge } e_j \\ 0, & \text{otherwise} \end{cases} \tag{4}$$

Although incidence matrices are not as computationally useful as adjacency matrices, they retain certain interesting properties using linear algebra.

3.2 Incidence and Rank

First, an observation: the column sums of $Q(G)$ are all zero, and hence the rows of $Q(G)$ are linearly depedent, because the vector consisting of all 1 spans the null space of Q. Therefore, the rank of the incidence matrix Q for any graph must be less than the order n. It turns out, however, that for any graph G, only one of the columns is a linear combination of the others:

Lemma 3.1. *If G is a connected graph on n vertices, then rank $Q(G) = n - 1$.*

This lemma, however, applies only to connected graphs, in which there exists a path between any pair of vertices. A **disconnected graph** is a graph in which at least one pair of vertices does not have a path between them. This describes the intuitive idea that not all vertices need to lie on the same part of the graph, as demonstrated by the example in Figure 5.

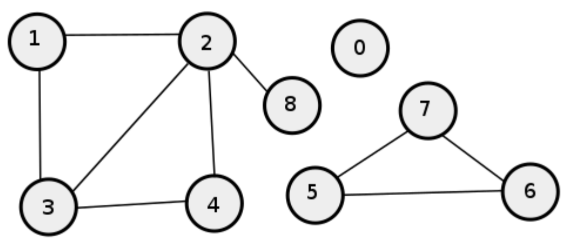

Figure 5: A disconnected graph of order 9

Each connected subsection of a graph G is called a **component** G. For connected graphs, there is only one component. Using the previous lemma, we can produce a more general result for any graph.

Theorem 3.1. *If G is a graph on n vertices and has k connected components then rank $Q(G) = n - k$.*

Thus, incidence matrices capture a different aspect of graphs. While adjacency matrices capture the density of a graph and allow for computations on relationships between vertices, incidence matrices account for the edges' relationships with the vertices, and therefore relate to properties such as components.

3.3 Path Matrices and Incidence Matrices

Let G be a graph with size m and $u, v \in V(G)$ (that is, any two vertices u and v of G). The path matrix for vertices u and v – denoted $P(u, v)$ – is the q by m matrix, where q is the number of different paths between u and v, defined as follows:

$$P(u,v)_{ij} = \begin{cases} 1, & \text{if edge } e_j \text{ lies in the } i-th \text{ path}, v_j \\ 0, & \text{otherwise.} \end{cases} \tag{5}$$

In other words, a path matrix is defined for a certain pair of vertices in a graph G in which

- the rows of $P(u, v)$ correspond to the different paths from vertex u to vertex v (which, for example, could be counted using the m-th power of the adjacency matrix $A^m(G)$, as described in Section 2.2)
- the columns of $P(u, v)$ correspond each to an edge in graph G

Incredibly, there exists a connection between the incidence matrix of a graph and its path matrix.

Theorem 3.2. *Let G be a graph of order n, and at least two vertices $u, v \in V(G)$. Arrange the columns of its incidence matrix Q(G) such that they align with the columns of the path matrix P(u,v). Then,*

$$Q(G) * P^T(u,v) \ (mod\ 2) = M, \tag{6}$$

where M is the matrix having all ones in two rows u and v, and zeros in the remaining n - 2 rows.

Herein incidence matrices again encapsulate the edge and vertex relationships, rather than just the vertex relationships found in adjacency matrices. One of the most powerful matrix representations, however, has yet to come.

4 Laplacian Matrices

4.1 Definition

The **degree** of a vertex v_i, denoted d_i, is the total number of other vertices to which vertex i is adjacent. That is, the degree of each vertex is how many edges are incident from that vertex. This provides us with a basis of knowledge for defining the Laplacian matrix, which at first glance seems entirely arbitrary, but has a number of incredible properties. The Laplacian Matrix of a Graph G, denoted $L(G)$, is the n by n matrix defined as follows:

$$L_{ij} = \begin{cases} -1, & \text{if vertex } v_i \text{ is adjacent to vertex } v_j \\ 0, & \text{if vertex } v_i \text{ is not adjacent to vertex } v_j \\ d_i, & \text{if } i = j \end{cases} \tag{7}$$

Thus, a Laplacian matrix has the degrees of each of its vertices along its diagonal, and a -1 elsewhere if two vertices are adjacent. Otherwise, every other entry consists of all zeros.

If we define $D\prime(G)$ to be a diagonal matrix whose entries are the degrees of each vertex, then the Laplacian Matrix of a graph G can be equivalently defined as

$$L(G) = D\prime(G) - A(G) \tag{8}$$

Remarkably, Laplacian Matrices can also be expressed in terms of Incidence Matrices. Assigning any arbitrary orientation to a graph G – that is, assigning a fixed, arbitrary direction to each edge – the Laplacian matrix of G can be expressed in terms of the incidence matrix of the oriented representation:

$$L(G) = Q(G)Q(G)^T \tag{9}$$

For example, take the basic graph in Figure 6:

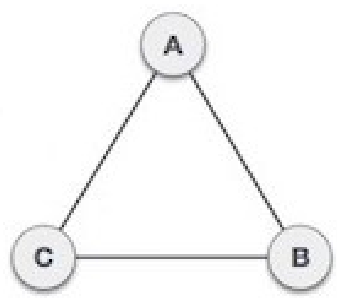

Figure 6: Laplacian Matrix Example

The Laplacian Matrix for this graph, including its relationship to the adjacency and inciddence matrices, is shown in Equation 10:

$$L(G) = \begin{bmatrix} 2 & -1 & -1 \\ -1 & 2 & -1 \\ -1 & -1 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} - \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 1 \\ -1 & -1 & 0 \\ 0 & 1 & -1 \end{bmatrix} * \begin{bmatrix} 1 & -1 & 0 \\ 0 & -1 & 1 \\ 1 & 0 & -1 \end{bmatrix} \tag{10}$$

4.2 The Matrix Tree Theorem

A **spanning tree** of a graph G is a tree which is a subgraph of G. That is, a spanning tree is a tree on the same vertex set as G which uses some subset of $E(G)$ such that every single vertex is adjacent to at least one edge. As one can imagine, counting all of the spanning subtrees of a graph G can rapidly become an arduous task. However, there exists an effective solution, and it makes use of Laplacian Matrices.

Namely, one of the most useful, elegant theorems in graph theory is that of The Matrix Tree Theorem.

Theorem 4.1 (**Matrix Tree Theorem**). *Let G be a graph of order n. Then the cofactor of any element of L(G) equals the number of spanning trees of G.*

That is, all cofactors of a Laplacian matrix are the same, and that shared value counts the number of spanning trees of the graph G!

Return to the above example. Taking the (2,3)-cofactor of the Laplacian Matrix of the triangle:

$$C_{23} = (-1)^{2+3} * \begin{vmatrix} 2 & -1 \\ -1 & -1 \end{vmatrix} = (-1) * (-2 - 1) = 3 \tag{11}$$

which corresponds to the 3 spanning trees of the graph, depicted in Figure 7.

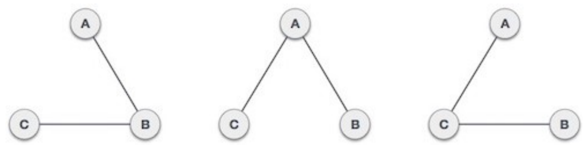

Figure 7: The Three Spanning Trees of K_3

It can be easily checked that every other cofafctor of the Laplacian matrix produces the same result.

5 Distance Matrices

5.1 Definition

For a graph G of order n, the distance matrix, denoted $D(G)$, of graph G is an n by n matrix whose (i,j)-th entry is the distance between vertices v_i and v_j. Naturally, this means that $D(G)$ is a symmetric matrix whose diagonal consists entirely of zeros.

5.2 Distance Matrices and Trees

When distance matrices are applied to trees, an interesting generality emerges:

Theorem 5.1. *Let T be a tree with order n. Then the determinant of $D(T)$ is given by*

$$det\ D(T) = (-1)^{n-1}(n-1) * 2^{n-2} \tag{12}$$

While this may seem a trivial result, it has profound implications. First, as a direct connection to linear algebra, because the determinant is nonzero for all trees of order $n \geq 3$, the distance matrix of all trees is nonsingular. That is, all distance matrices of trees have an inverse

Second, because the determinant depends only on the size of the matrix, and not on the placement of entries, the result depends only on the size of a given tree, independent of its structure or shape. That is, of the n^{n-2} trees of order n, every single one of their distance matrices has the same determinant. This suggests an underlying similarity between all trees, a similarity which has persisted in myriad other results regarding trees, and an even greater diversity of conjectures.

5.3 Connection Between Distance and Laplacian Matrices

Finally, there exists a connection between the distance and Laplacian matrices of a tree:

Theorem 5.2. *Let T be a tree of order n with Laplacian matrix L and distance matrix D. Then*

$$LDL = -2L$$

Once again, there exists a pattern for trees which does not occur for other graphs, and it emerges for all trees of a given order, regardless of their structure or shape. It suggests an underlying similarity of great interest.

References

[1] Graph Matrices. http://compalg.inf.elte.hu/~tony/Oktatas/TDK/FINAL/Chap%2010.PDF, 2010. [Online; accessed 16-03-2017].

[2] R.B. Bapat. *Graphs and Matrices*. Hindustan Book Agency, New Delhi, India, 2010.

[3] Gary Chartrand and Ping Zhang. *A First Course in Graph Theory*. Dover Publications, Boston, 2012.

[4] Douglas E. Ensley and J. Winston Crawley. *Discrete Mathematics*. John Wiley and Sons, Inc., Hoboken, New Jersey, 2006.

[5] Laszlo Lovasz. Eigenvalues of Graphs. http://www.cs.elte.hu/~lovasz/eigenvals-x.pdf, 2007. [Online; accessed 17-03-2017].

[6] Damir Vukiccevic and Ivan Gutman. Laplacian Matrix and Distance Matrices in Trees. http://elib.mi.sanu.ac.rs/files/journals/kjm/26/d003download.pdf, 2004. [Online; accessed 17-03-2017].