Connectivity in graphs

COMS 3251 Fall 2022 (Daniel Hsu)

1 Graphs and incidence matrices

A *graph* is a tuple $(\mathcal{V}, \mathcal{E})$, where $\mathcal{V}$ is a finite set, and $\mathcal{E} \subseteq \{\{u, v\} : u, v \in \mathcal{V},\ u \neq v\}$ is a subset of unordered pairs of elements from $\mathcal{V}$. The elements of $\mathcal{V}$ are called *vertices*, and the elements of $\mathcal{E}$ are called *edges*. A graph models relationships between vertices. For example:

- The vertices could correspond to towns, and an edge could represent the presence of a road connecting two towns.
- The vertices could correspond to people, and an edge could represent a "friendship" relation on the social network *du jour*.
- The edges could represent circuit elements (e.g., resistors), and the vertices could be the junctions between these elements.

For reasons that will become clear later, we will treat edges as ordered pairs, where we arbitrarily decide on which of the two vertices in an edge comes first. So every edge is now a directed edge, where (s, t) means "$s \to t$".

We denote the number of edges by $m = |\mathcal{E}|$ and the number of vertices by $n = |\mathcal{V}|$. To simplify notation later, we let the vertices be the first n positive integers $\mathcal{V} = \{1, \ldots, n\}$, and we also arbitrarily label the edges $\{1, \ldots, m\}$.

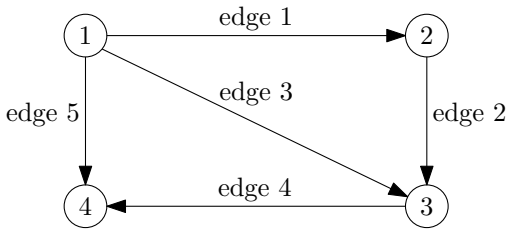
The *(edge-vertex) incidence matrix*[1] A for a graph is an $m \times n$ matrix, with rows corresponding to edges and columns corresponding to vertices, where

$$A_{e,v} = \begin{cases} -1 & \text{if edge } e \text{ starts at } v, \\ 1 & \text{if edge } e \text{ ends at } v, \\ 0 & \text{otherwise.} \end{cases}$$

We'll see that graph-theoretic connectivity properties of a graph will have linear algebraic implications for its incidence matrix. Moreover, a linear algebraic algorithm that operates on this matrix will have a graph-theoretic interpretation.

[1] Please do not confuse the incidence matrix with the adjacency matrix of a graph.

Running example, part 1. We use the following graph and incidence matrix in our running examples:



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

2 Potential differences

Suppose we assign a value—which we'll call a "potential"—to each each vertex in a graph. Such an assignment is described by an n-vector $\mathbf{x} = (x_1, \ldots, x_n)$. The matrix-vector product $\mathbf{b} = (b_1, \ldots, b_m) = A\mathbf{x}$ yields an m-vector in which b_e is the *potential difference*[2] across edge e. If edge e is the ordered pair (s, t), then

$$b_e \;=\; \sum_{v \in \mathcal{V}} A_{e,v} x_v \;=\; x_t - x_s.$$

Given the potential differences $\mathbf{b}$ and the incidence matrix A, is it possible to uniquely determine the potentials $\mathbf{x}$? No: the potential differences are the same if we increase the potential of every vertex by the same amount. In particular, if $\mathbf{x} = (c, \ldots, c)$ for any scalar c, then $A\mathbf{x} = \mathbf{0}$. This shows that the columns of the incidence matrix A are linearly dependent. In particular, A has at most $n - 1$ linearly independent columns.

Since, in any matrix, the maximum number of linearly independent columns is the same as the maximum number of linearly independent rows (which we call the rank of the matrix), we know that the incidence matrix A has at most $n - 1$ linearly independent rows. (Remember this for later!)

3 Paths and cycles

The incidence matrix A was, in a sense, defined row-wise. What can we learn about the rows of A from studying the graph structure? In this section, we

[2] In electrical circuits, differences in electrical potentials are called voltages, and the voltages between junctions are what determine the behavior (e.g., flow of electric charge) of the circuit.

look at implications of the presence of paths and cycles.

A word of caution: we are going extensively use row-wise matrix multiplication in this section (and the next one). Specifically, we will multiply a row vector (a $1 \times m$ matrix) by the incidence matrix (an $m \times n$ matrix) to produce another row vector (a $1 \times n$ matrix):

$$\begin{bmatrix} y_1 & \cdots & y_m \end{bmatrix} \begin{bmatrix} \leftarrow & \mathbf{a}_1^\top & \rightarrow \\ & \vdots & \\ \leftarrow & \mathbf{a}_m^\top & \rightarrow \end{bmatrix} = y_1 \, \mathbf{a}_1^\top + \cdots + y_m \, \mathbf{a}_m^\top.$$

If you prefer, you can convert all row vectors into columns, and use the usual matrix-vector multiplication that you may be more used to.

We first consider the implication of a path in the graph. Suppose there is a path from s to t in the graph, with $s \neq t$, and for now assume it respects the (arbitrarily determined) directions of the edges. Then summing the rows of A corresponding to these edges results in a row with -1 in component s, a 1 in component t, and zeros elsewhere. We specify the path using a row vector $\mathbf{y}^\top$ with the appropriate coefficients from $\{0, 1\}$ so that $\mathbf{y}^\top A$ computes the sum of rows.

Running example, part 2. Consider the path from $s = 1$ to $t = 4$ using edges 1, 2, and 4. Adding these rows yields the desired row:

$$\begin{bmatrix} 1 & 1 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ -1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} -1 & 0 & 0 & 1 \end{bmatrix}. \quad \blacksquare$$

Suppose, now, that there is a path from s to t, again with $s \neq t$, but only by using some edges in the "reverse direction". We can again obtain the same row as above by linearly combining rows of A, but this time by adding the rows for edges used in the "forward direction", and then subtracting the rows for edges used in the "reverse direction". We can think of this as scaling the rows for "reverse direction" edges by -1 before adding up the rows; the effect of the scaling is to flip the directions of these edges. We call this the "signed path sum" of the rows corresponding to the edges in the path; it is also specified by a row vector $\mathbf{y}^\top$ but now with coefficients in $\{-1, 0, 1\}$.

Running example, part 3. Consider the path from $s = 2$ to $t = 1$ using edges 2, 4, and 5. The signed path sum of these rows yields the desired row:

$$\begin{bmatrix} 0 & 1 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ -1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & -1 & 0 & 0 \end{bmatrix}. \quad \blacksquare$$

Now we consider the implication of a cycle in the graph. By a cycle, we mean a path in the graph that starts and ends at the same vertex. As above, we allow edges to be used in the "reverse direction". Then, using the signed path sum of the rows corresponding to the path, we obtain the all-zeros row. (What is the signed path sum if you remove one of the edges in the cycle?)

Running example, part 4. Consider the cycle starting at 1 using edges 1, 2, 4, and 5. The signed path sum of these rows yields the all-zeros row:

$$\begin{bmatrix} 1 & 1 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 1 & 0 \\ 0 & 0 & -1 & 1 \\ -1 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 \end{bmatrix}. \quad \blacksquare$$

So, if the graph has a cycle, then the rows of the incidence matrix are linearly dependent, because we can linearly combine their rows to produce the all-zeros row in a non-trivial way. Here is the contrapositive:

Theorem 1. *If the rows of the incidence matrix of a graph are linearly independent, then the graph is acyclic (i.e., has no cycles).*

4 CR factorization and acyclic subgraphs

Recall the algorithm that produces the CR factorization of a matrix A: it yields a matrix C whose columns are a linearly independent subset of the columns of A, of maximum size, and a matrix R that specifies how to obtain columns in A as linear combinations of columns in C, such that $A = CR$.

Let us apply the same algorithm but now to the rows of A. This means we apply the algorithm to the *transpose* of A, written $A^\intercal$ (read aloud as "A transpose"). To obtain the transpose of a matrix, turn every row of the matrix into a column, and then arrange these columns side-by-side in the same top-to-bottom order they were originally in as rows.[3] The result is a "row-wise CR factorization" of A, which we write as

$$A^\intercal \;=\; CR, \quad \text{or, equivalently,} \quad A \;=\; R^\intercal C^\intercal.$$

The rows of $C^\intercal$ are a linearly independent subset of rows of A, of maximum size, such that every row of A is a linear combination of the rows in this extracted subset, as specified by the matrix $R^\intercal$. This subset of rows in $C^\intercal$ corresponds to a subgraph of the original graph: it retains only a subset of the original edges. By Theorem 1, the subgraph is acyclic.

So, the row-wise CR factorization of the incidence matrix extracts an acyclic subgraph of the original graph with as many edges as possible.

Running example, part 5. Let the rows of A be denoted by $\mathbf{a}_1^\intercal, \ldots, \mathbf{a}_5^\intercal$:

$$A^\intercal \;=\; \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \mathbf{a}_3 & \mathbf{a}_4 & \mathbf{a}_5 \end{bmatrix} \;=\; \begin{bmatrix} -1 & 0 & -1 & 0 & -1 \\ 1 & -1 & 0 & 0 & 0 \\ 0 & 1 & 1 & -1 & 0 \\ 0 & 0 & 0 & 1 & 1 \end{bmatrix}.$$

The execution of the CR factorization algorithm on $A^\intercal$ is as follows.

- Initially: C is the empty list.
- Iteration $k = 1$: $\mathbf{a}_1 \notin \mathsf{CS}(C)$, so $\mathbf{a}_1$ is appended to C.
- Iteration $k = 2$: $\mathbf{a}_2 \notin \mathsf{CS}(C)$, so $\mathbf{a}_2$ is appended to C.
- Iteration $k = 3$: $\mathbf{a}_3 = \mathbf{a}_1 + \mathbf{a}_2$, so there is no change to C.
- Iteration $k = 4$: $\mathbf{a}_4 \notin \mathsf{CS}(C)$, so $\mathbf{a}_4$ is appended to C.
- Iteration $k = 5$: $\mathbf{a}_3 = \mathbf{a}_1 + \mathbf{a}_2 + \mathbf{a}_4$, so there is no change to C.

[3] Or: turn every column of the matrix into a row, and then arrange these rows from top-to-bottom in the same left-to-right order they were originally in as columns.

The final row-wise CR factorization of A is $A = R^\intercal C^\intercal$, where

$$R^\intercal = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 1 & 1 \end{bmatrix}, \quad C^\intercal = \begin{bmatrix} \leftarrow & \mathbf{a}_1^\intercal & \rightarrow \\ \leftarrow & \mathbf{a}_2^\intercal & \rightarrow \\ \leftarrow & \mathbf{a}_4^\intercal & \rightarrow \end{bmatrix} = \begin{bmatrix} -1 & 1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix}. \quad \blacksquare$$

5 Connectivity

Finally, let us consider the implication of the connectedness of the graph. Here, by connected, we mean that there is a path between every pair of vertices. Again, we allow paths to use edges in the "reverse direction".

If the graph is connected, then there is a path from vertex 1 to every other vertex. This means that each of the $n-1$ rows of the matrix Z, defined by

$$Z = \begin{bmatrix} -1 & 1 & 0 & \cdots & 0 \\ -1 & 0 & 1 & \cdots & 0 \\ & & & \vdots & \\ -1 & 0 & 0 & \cdots & 1 \end{bmatrix},$$

can be obtained by a signed path sum of the rows of A. Therefore, connectivity implies that there is an $(n-1) \times m$ matrix Y such that

$$YA = Z. \tag{1}$$

The ith row of Y corresponds to a path that starts at vertex 1 and ends at vertex $i+1$. It is clear that the $n-1$ rows of this matrix Z are linearly independent. We are able to produce $n-1$ linearly independent rows using linear combinations of the rows of A.

Running example, part 6. We show how to reach all other vertices from vertex 1:

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

(By no means were these the only ways to reach the other vertices from 1.) ∎

Let us substitute the row-wise CR factorization of A into (1):

$$YR^\intercal C^\intercal \;=\; Z, \tag{2}$$

and think of $YR^\intercal$ together as a single matrix, so the equation shows how to linearly combine the rows of $C^\intercal$ to produce the rows of Z.

If k is the number of rows in $C^\intercal$, then (2) shows that the $n-1$ linearly independent rows of Z are in the span of the k linearly independent rows of $C^\intercal$. By the Exchange Theorem, we must have $n-1 \leq k$. On the other hand, as you recall from Section 2, there are at most $n-1$ linearly independent rows of A. Since rows of $C^\intercal$ are a subset of the rows of A, we have $k \leq n-1$. We conclude $k = n-1$.

To summarize: if a graph is connected, then applying the row-wise CR factorization algorithm to its incidence matrix extracts an acyclic subgraph of the original graph of size exactly $n-1$. Such a subgraph is a tree, and it is called a *spanning tree* of the original graph. And the rank of the incidence matrix is exactly $n-1$ in this case.

6 Disconnected graphs

Suppose a graph has p connected components, $p > 1$. Say the components are $\mathcal{V}_1, \ldots, \mathcal{V}_p$, where $\mathcal{V}_i \cap \mathcal{V}_j = \emptyset$ and $\mathcal{V}_1 \cup \cdots \cup \mathcal{V}_p = \{1, \ldots, n\}$. For simplicity, assume $\mathcal{V}_1$ is the first $|\mathcal{V}_1|$ positive integers, $\mathcal{V}_2$ is the next $|\mathcal{V}_2|$ positive integers, and so on. Then, it is possible to label the edges so that the incidence matrix A for the graph has the "block diagonal" form

$$A \;=\; \left[\begin{array}{c|c|c} A_1 & \text{zeros} & \text{zeros} \\ \hline \text{zeros} & \ddots & \text{zeros} \\ \hline \text{zeros} & \text{zeros} & A_p \end{array}\right].$$

Here, the submatrix A_i is the incidence matrix for the ith component, and as argued above, it has $|\mathcal{V}_i| - 1$ linearly independent rows. Note: it is possible that some of these blocks are empty (corresponding to components with a single vertex and no edges). So the overall matrix A has $\sum_{i=1}^{p}(|\mathcal{V}_i|-1) = n-p$ linearly independent rows. This is the rank of the matrix A.

A Fundamental subspaces

This appendix can be read after studying the concepts of column space, nullspace, row space, and left nullspace.

Let A be the edge-vertex incidence matrix for a graph $(\mathcal{V}, \mathcal{E})$. We consider the four fundamental subspaces of A.

A.1 Column space

Recall from Section 2 that the matrix-vector product $A\mathbf{x}$ for any vector $\mathbf{x}$ of "potentials" yields a vector of potential differences across edges. So the column space $\mathsf{CS}(A)$ contains vectors specifying potential differences across edges realizable by some assignment of potentials to vertices.

If r is the rank of A, then is a subset of r vertices in $\mathcal{V}$ such that the corresponding columns of A form a basis for $\mathsf{CS}(A)$. This means that any vector of realizable potential differences across edges can be obtained by assigning potential values to this set of vertices, and assigning 0 potential to the remaining vertices.

A.2 Nullspace

Starting with a vector $\mathbf{x}$ that assigns potentials to every vertex in the graph, adding the same amount c to the potentials at every vertex does not change the potential differences. In other words, $A(\mathbf{x} + c\mathbf{1}) = A\mathbf{x}$, where $\mathbf{1} = (1, \ldots, 1)$ is the all-ones vector. This implies that $\mathbf{1} \in \mathsf{NS}(A)$, so the dimension of the nullspace is at least 1.

If the graph is connected, then we saw in Section 5 that A has $n-1$ linearly independent rows; in this case, $\dim(\mathsf{NS}(A)) = 1$. From Section 6, we see that if the graph has p connected components $\mathcal{V}_1, \ldots, \mathcal{V}_p$, then A has $n-p$ linearly independent rows, i.e., $\mathrm{rank}(A) = n - p$. This means $\dim(\mathsf{NS}(A)) = p$. For each connected component $\mathcal{V}_i$, let $\mathbf{v}_i$ be the vector with 1's in components $j \in \mathcal{V}_i$ and 0's elsewhere. So $\mathbf{v}_i$ is the "all ones" vector for just the ith connected component. Then $\{\mathbf{v}_1, \ldots, \mathbf{v}_p\}$ is a basis for $\mathsf{NS}(A)$.

In summary, the nullspace is the space of vectors that assign the same "potential" all vertices in the same connected component.

A.3 Row space

If $\mathbf{p}$ is an n-vector with -1 in component s and a 1 in component t, and zeros elsewhere, then $\mathbf{p}$ is in the row space $\mathsf{CS}(A^\mathsf{T})$ of A if there is a path from s to t in the graph. Linear combinations $\mathbf{f} = (f_1, \ldots, f_n)$ of such vectors are also in $\mathsf{CS}(A^\mathsf{T})$. These vectors correspond to *flows* in the graph, where (say) water is injected into some vertices (vertices i such that $f_i < 0$) and extracted from other vertices (vertices i such that $f_i > 0$); $|f_i|$ is the amount of (or, technically, the rate at which) water is injected or extracted.

Linearly independent subsets of rows correspond to acyclic subgraphs. If the graph is connected, then we can obtain $n - 1$ linearly independent rows of A, a basis for the row space, which corresponds to a spanning tree.

If the graph has $p > 1$ connected components, we can find $n - p$ linearly independent rows of A (again, a basis for the row space), and they correspond to a "spanning forest" composed of p trees, one in each component.

A.4 Left nullspace

If $\mathbf{y}^\mathsf{T}$ specifies the signed path sum for a cycle in the graph, then $\mathbf{y}^\mathsf{T} A = \mathbf{0}^\mathsf{T}$. The span of all such vectors $\mathbf{y}$ is the left nullspace $\mathsf{NS}(A^\mathsf{T})$ of A.

If the graph is connected, then a basis for $\mathsf{NS}(A^\mathsf{T})$ can be obtained as follows. Let $\mathcal{T}$ denote the edges of any spanning tree for the graph. Note that $|\mathcal{T}| = n - 1$. For each edge $e = (i, j)$ in $\mathcal{E} \setminus \mathcal{T}$, consider the cycle that contains i and edges from $\mathcal{T}$. This cycle is unique, because there is exactly one path between the endpoints of e using edges in $\mathcal{T}$. Let $\mathbf{y}_e^\mathsf{T}$ be the row vector that specifies the signed path sum for this cycle. Notice that the set $\mathcal{B} = \{\mathbf{y}_e : e \in \mathcal{E} \setminus \mathcal{T}\}$ of such vectors is linearly independent: $\mathbf{y}_e$ is the only vector with a non-zero value in the component corresponding to edge e. The cardinality of this set is $|\mathcal{E} \setminus \mathcal{T}| = m - (n-1)$. By the Dimension Theorem and the fact $\mathrm{rank}(A^\mathsf{T}) = \mathrm{rank}(A) = n - 1$, we have $\dim(\mathsf{NS}(A^\mathsf{T})) = m - (n - 1)$. Hence, $\mathcal{B}$ is a basis for the left nullspace of A.

If the graph has $p > 1$ connected components, the construction from above applied to each component separately will, again, yield a basis for $\mathsf{NS}(A^\mathsf{T})$.

So the left nullspace is spanned by vectors corresponding to particular sets cycles in the graph. Sometimes these cycles are called "small loops", since they appear by adding just a single edge to a spanning tree (or forest).

B Orthogonality

This appendix can be read after studying the concept of orthogonal subspaces.

Let A be the edge-vertex incidence matrix for a graph $(\mathcal{V}, \mathcal{E})$. We interpret what it means for certain pairs of fundamental subspaces of A to be orthogonal.

B.1 Column space and left nullspace

The column space $\mathsf{CS}(A)$ contains vectors $\mathbf{b}$ specifying potential differences across edges realizable by some assignment of potentials to vertices. The left nullspace $\mathsf{NS}(A^\mathsf{T})$ is spanned by vectors $\mathbf{y}$ that specify signed path sums for cycles in the graph. Since $\mathsf{CS}(A)$ and $\mathsf{NS}(A^\mathsf{T})$ are orthogonal, the vectors $\mathbf{b}$ and $\mathbf{y}$ must be orthogonal:

$$\langle \mathbf{b}, \mathbf{y} \rangle \;=\; \sum_{e \in \mathcal{E}} b_e y_e \;=\; 0.$$

Notice that $y_e = 0$ if edge e is not in the cycle. The equation says that the (signed) sum of potential differences around the cycle must be zero. In the theory of electrical networks, this is called *Kirchhoff's Voltage Law (KVL)*.

B.2 Row space and nullspace

Assume that the graph is connected, so the nullspace of A is $\mathsf{NS}(A) = \mathrm{span}(\{\mathbf{1}\})$. A vector $\mathbf{f} = (f_1, \ldots, f_n)$ in the row space $\mathsf{CS}(A^\mathsf{T})$ of A represents a flow in the graph: when f_i is negative, we inject $|f_i|$ units of water into vertex i, and when f_i is positive, we extract $|f_i|$ units of water from the vertex i. Since $\mathsf{CS}(A^\mathsf{T})$ and $\mathsf{NS}(A)$ are orthogonal, the vector $\mathbf{f}$ satisfies

$$\langle \mathbf{f}, \mathbf{1} \rangle \;= \sum_{i \in \mathcal{V}} f_i \;=\; 0.$$

This says that the total amount of water injected in must equal the total amount of water extracted out. In the theory of electrical networks, this is a consequence of *Kirchhoff's Current Law (KCL)*.

If there are $p > 1$ connected components, then the sum of the f_i's within each component must also be zero.