THE MATRIX-TREE THEOREM

1 The Matrix-Tree Theorem.

The Matrix-Tree Theorem is a formula for the number of spanning trees of a graph in terms of the determinant of a certain matrix. We begin with the necessary graph-theoretical background. Let G be a finite graph, allowing multiple edges but not loops. (Loops could be allowed, but they turn out to be completely irrelevant.) We say that G is *connected* if there exists a walk between any two vertices of G. A *cycle* is a closed walk with no repeated vertices or edges, except for the the first and last vertex. A *tree* is a connected graph with no cycles. In particular, a tree cannot have multiple edges, since a double edge is equivalent to a cycle of length two. The three nonisomorphic trees with five vertices are given by:



A basic theorem of graph theory (whose easy proof we leave as an exercise) is the following.

1.1 Proposition. Let G be a graph with p vertices. The following conditions are equivalent.

(a) G is a tree.

- (b) G is connected and has p-1 edges.
- (c) G is has no cycles and has p-1 edges.

(d) There is a unique path (= walk with no repeated vertices) between any two vertices.

A spanning subgraph of a graph G is a graph H with the same vertex set as G, and such that every edge of H is an edge of G. If G has q edges, then the number of spanning subgraphs of G is equal to 2^q , since we can choose any subset of the edges of G to be the set of edges of H. (Note that multiple edges between the same two vertices are regarded as *distinguishable*.) A spanning subgraph which is a tree is called a *spanning tree*. Clearly G has a spanning tree if and only if it is connected [why?]. An important invariant of a graph G is its number of spanning trees, called the *complexity* of G and denoted $\kappa(G)$.

1.2 Example. Let G be the graph illustrated below, with edges a, b, c, d, e.



Then G has eight spanning trees, namely, abc, abd, acd, bcd, abe, ace, bde, and cde (where, e.g., abc denotes the spanning subgraph with edge set $\{a, b, c\}$).

1.3 Example. Let $G = K_5$, the complete graph on five vertices. A simple counting argument shows that K_5 has 60 spanning trees isomorphic to the first tree in the above illustration of all nonisomorphic trees with five vertices, 60 isomorphic to the second tree, and 5 isomorphic to the third tree. Hence $\kappa(K_5) = 125$. It is even easier to verify that $\kappa(K_1) = 1$, $\kappa(K_2) = 1$, $\kappa(K_3) = 3$, and $\kappa(K_4) = 16$. Can the reader make a conjecture about the value of $\kappa(K_p)$ for any $p \ge 1$?

Our object is to obtain a "determinantal formula" for $\kappa(G)$. For this we need an important result from matrix theory which is often omitted from a beginning linear algebra course. This result, known as the Binet-Cauchy theorem (or sometimes as the Cauchy-Binet theorem), is a generalization of the familiar fact that if A and B are $n \times n$ matrices, then det(AB) = $\det(A) \det(B)$ (where det denotes determinant)¹. We want to extend this formula to the case where A and B are rectangular matrices whose product is a square matrix (so that $\det(AB)$ is defined). In other words, A will be an $m \times n$ matrix and B an $n \times m$ matrix, for some $m, n \ge 1$.

We will use the following notation involving submatrices. Suppose $A = (a_{ij})$ is an $m \times n$ matrix, with $1 \leq i \leq m, 1 \leq j \leq n$, and $m \leq n$. Given an *m*-element subset *S* of $\{1, 2, \ldots, n\}$, let A[S] denote the $m \times m$ submatrix of *A* obtained by taking the columns indexed by the elements of *S*. In other words, if the elements of *S* are given by $j_1 < j_2 < \cdots < j_m$, then $A[S] = (a_{i,j_k})$, where $1 \leq i \leq m$ and $1 \leq k \leq m$. For instance, if

$$A = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 6 & 7 & 8 & 9 & 10 \\ 11 & 12 & 13 & 14 & 15 \end{bmatrix}$$

and $S = \{2, 3, 5\}$, then

$$A[S] = \begin{bmatrix} 2 & 3 & 5 \\ 7 & 8 & 10 \\ 12 & 13 & 15 \end{bmatrix}.$$

Similarly, let $B = (b_{ij})$ be an $n \times m$ matrix with $1 \leq i \leq n, 1 \leq j \leq m$ and $m \leq n$. Let S be an m-element subset of $\{1, 2, \ldots, n\}$ as above. Then B[S] denotes the $m \times m$ matrix obtained by taking the rows of B indexed by S. Note that $A^t[S] = A[S]^t$, where t denotes transpose.

1.4 Theorem. (the Binet-Cauchy Theorem) Let $A = (a_{ij})$ be an $m \times n$ matrix, with $1 \leq i \leq m$ and $1 \leq j \leq n$. Let $B = (b_{ij})$ be an $n \times m$ matrix with $1 \leq i \leq n$ and $1 \leq j \leq m$. (Thus AB is an $m \times m$ matrix.) If m > n, then det(AB) = 0. If $m \leq n$, then

$$\det(AB) = \sum_{S} (\det A[S])(\det B[S]),$$

¹In the "additional material" handout (Theorem 2.4) there is a more general determinantal formula without the use of the Binet-Cauchy theorem. However, the use of the Binet-Cauchy theorem does afford some additional algebraic insight.

where S ranges over all m-element subsets of $\{1, 2, \ldots, n\}$.

Before proceeding to the proof, let us give an example. We write $|a_{ij}|$ for the determinant of the matrix (a_{ij}) . Suppose

$$A = \begin{bmatrix} a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}, \quad B = \begin{bmatrix} c_1 & d_1 \\ c_2 & d_2 \\ c_3 & d_3 \end{bmatrix}.$$

Then

 $\det(AB) = \begin{vmatrix} a_1 & a_2 \\ b_1 & b_2 \end{vmatrix} \cdot \begin{vmatrix} c_1 & d_1 \\ c_2 & d_2 \end{vmatrix} + \begin{vmatrix} a_1 & a_3 \\ b_1 & b_3 \end{vmatrix} \cdot \begin{vmatrix} c_1 & d_1 \\ c_3 & d_3 \end{vmatrix} + \begin{vmatrix} a_2 & a_3 \\ b_2 & b_3 \end{vmatrix} \cdot \begin{vmatrix} c_2 & d_2 \\ c_3 & d_3 \end{vmatrix}.$

Proof of Theorem 1.4 (sketch). First suppose m > n. Since from linear algebra we know that rank $(AB) \leq \operatorname{rank}(A)$ and that the rank of an $m \times n$ matrix cannot exceed n (or m), we have that rank $(AB) \leq n < m$. But AB is an $m \times m$ matrix, so det(AB) = 0, as claimed.

Now assume $m \leq n$. We use notation such as M_{rs} to denote an $r \times s$ matrix M. It is an immediate consequence of the definition of matrix multiplication (which the reader should check) that

$$\begin{bmatrix} R_{mm} & S_{mn} \\ T_{nm} & U_{nn} \end{bmatrix} \begin{bmatrix} V_{mn} & W_{mm} \\ X_{nn} & Y_{nm} \end{bmatrix} = \begin{bmatrix} RV + SX & RW + SY \\ TV + UX & TW + UY \end{bmatrix}.$$
 (1)

In other words, we can multiply "block" matrices of suitable dimensions as if their entries were numbers. Note that the entries of the right-hand side of (1) all have well-defined dimensions (sizes), e.g., RV + SX is an $m \times n$ matrix since both RV and SX are $m \times n$ matrices.

Now in equation (1) let $R = I_m$ (the $m \times m$ identity matrix), S = A, $T = O_{nm}$ (the $n \times m$ matrix of 0's), $U = I_n$, V = A, $W = O_{mm}$, $X = -I_n$, and Y = B. We get

$$\begin{bmatrix} I_m & A \\ O_{nm} & I_n \end{bmatrix} \begin{bmatrix} A & O_{mm} \\ -I_n & B \end{bmatrix} = \begin{bmatrix} O_{mn} & AB \\ -I_n & B \end{bmatrix}.$$
 (2)

Take the determinant of both sides of (2). The first matrix on the left-hand side is an upper triangular matrix with 1's on the main diagonal. Hence its

determinant is one. Since the determinant of a product of square matrices is the product of the determinants of the factors, we get

$$\begin{vmatrix} A & O_{mm} \\ -I_n & B \end{vmatrix} = \begin{vmatrix} O_{mn} & AB \\ -I_n & B \end{vmatrix}.$$
 (3)

It is easy to see [why?] that the determinant on the right-hand side of (3) is equal to $\pm \det(AB)$. So consider the left-hand side. A nonzero term in the expansion of the determinant on the left-hand side is obtained by taking the product (with a certain sign) of m + n nonzero entries, no two in the same row and column (so one in each row and each column). In particular, we must choose m entries from the last m columns. These entries belong to m of the bottom n rows [why?], say rows $m + s_1, m + s_2, \ldots, m + s_m$. Let $S = \{s_1, s_2, \ldots, s_m\} \subseteq \{1, 2, \ldots, n\}$. We must choose n - m further entries from the last n rows, and we have no choice but to choose the -1's in those rows m+i for which $i \notin S$. Thus every term in the expansion of the left-hand side of (3) uses exactly n - m of the -1's in the bottom left block $-I_n$.

What is the contribution to the expansion of the left-hand side of (3) from those terms which use exactly the -1's from rows m + i where $i \notin S$? We obtain this contribution by deleting all rows and columns to which these -1's belong (in other words, delete row m + i and column i whenever $i \in$ $\{1, 2, \ldots, n\} - S$), taking the determinant of the $2m \times 2m$ matrix M_S that remains, and multiplying by an appropriate sign [why?]. But the matrix M_S is in block-diagonal form, with the first block just the matrix A[S] and the second block just B[S]. Hence det $M_S = (\det A[S])(\det B[S])$ [why?]. Taking all possible subsets S gives

$$\det AB = \sum_{\substack{S \subseteq \{1,2,\dots,n\}\\|S|=m}} \pm (\det A[S])(\det B[S]).$$

It is straightforward but somewhat tedious to verify that all the signs are +; we omit the details. This completes the proof. \Box

Let G be a graph with vertices v_1, \ldots, v_p . The *adjacency matrix* of G is the $p \times p$ matrix $\mathbf{A} = \mathbf{A}(G)$, over the field of complex numbers, whose (i, j)-entry a_{ij} is equal to the number of edges incident to v_i and v_j . Thus

A is a real symmetric matrix (and hence has real eigenvalues) whose trace is the number of loops in G.

We now define two matrices related to A(G). Assume for simplicity that G has no loops. (This assumption is harmless since loops have no effect on $\kappa(G)$.)

1.5 Definition. Let G be as above. Give G an orientation \mathfrak{o} , i.e, for every edge e with vertices u, v, choose one of the ordered pairs (u, v) or (v, u). (If we choose (u, v), say, then we think of putting an arrow on e pointing from u to v; and we say that e is directed from u to v, that u is the *initial vertex* and v the *final vertex* of e, etc.)

(a) The *incidence matrix* M(G) of G (with respect to the orientation \mathfrak{o}) is the $p \times q$ matrix whose (i, j)-entry M_{ij} is given by

$$\boldsymbol{M}_{ij} = \begin{cases} 1, & \text{if the edge } e_j \text{ has initial vertex } v_i \\ -1, & \text{if the edge } e_j \text{ has final vertex } v_i \\ 0, & \text{otherwise.} \end{cases}$$

(b) The laplacian matrix L(G) of G is the $p \times p$ matrix whose (i, j)-entry L_{ij} is given by

 $\boldsymbol{L}_{ij} = \begin{cases} -m_{ij}, & \text{if } i \neq j \text{ and there are } m_{ij} \text{ edges between } v_i \text{ and } v_j \\ \deg(v_i), & \text{if } i = j, \end{cases}$

where $\deg(v_i)$ is the number of edges incident to v_i . (Thus L(G) is symmetric and does not depend on the orientation \mathfrak{o} .)

Note that every column of M(G) contains one 1, one -1, and q-20's; and hence the sum of the entries in each column is 0. Thus all the rows sum to the 0 vector, a linear dependence relation which shows that rank(M(G)) < p. Two further properties of M(G) and L(G) are given by the following lemma.

1.6 Lemma. (a) We have $MM^t = L$.

(b) If G is regular of degree d (i.e., every vertex of G has degree d), then L(G) = dI - A(G), where A(G) denotes the adjacency matrix of G. Hence if G (or $\mathbf{A}(G)$) has eigenvalues $\lambda_1, \ldots, \lambda_p$, then $\mathbf{L}(G)$ has eigenvalues $d - \lambda_1, \ldots, d - \lambda_p$.

Proof. (a) This is immediate from the definition of matrix multiplication. Specifically, for $v_i, v_j \in V(G)$ we have

$$(\boldsymbol{M}\boldsymbol{M}^t)_{ij} = \sum_{e_k \in E(G)} \boldsymbol{M}_{ik} \boldsymbol{M}_{jk}.$$

If $i \neq j$, then in order for $M_{ik}M_{jk} \neq 0$, we must have that the edge e_k connects the vertices v_i and v_j . If this is the case, then one of M_{ik} and M_{jk} will be 1 and the other -1 [why?], so their product is always -1. Hence $(MM^t)_{ij} = -m_{ij}$, as claimed.

There remains the case i = j. Then $M_{ik}M_{ik}$ will be 1 if e_k is an edge with v_i as one of its vertices and will be 0 otherwise [why?]. So now we get $(MM^t)_{ii} = \deg(v_i)$, as claimed. This proves (a).

(b) Clear by (a), since the diagonal elements of $\boldsymbol{M}\boldsymbol{M}^t$ are all equal to d.

Now assume that G is connected, and let $M_0(G)$ be M(G) with its last row removed. Thus $M_0(G)$ has p-1 rows and q columns. Note that the number of rows is equal to the number of edges in a spanning tree of G. We call $M_0(G)$ the reduced incidence matrix of G. The next result tells us the determinants (up to sign) of all $(p-1) \times (p-1)$ submatrices N of M_0 . Such submatrices are obtained by choosing a set S of p-1 edges of G, and taking all columns of M_0 indexed by the edges in S. Thus this submatrix is just $M_0[S]$.

1.7 Lemma. Let S be a set of p-1 edges of G. If S does not form the set of edges of a spanning tree, then det $M_0[S] = 0$. If, on the other hand, S is the set of edges of a spanning tree of G, then det $M_0[S] = \pm 1$.

Proof. If S is not the set of edges of a spanning tree, then some subset R of S forms the edges of a cycle C in G. Consider the submatrix $M_0[R]$ of $M_0[S]$ obtained by taking the columns indexed by edges in R. Suppose that the cycle C defined by R has edges f_1, \ldots, f_s in that order. Multiply the column of $M_0[R]$ indexed by f_j by 1 if in going around C we traverse

 f_i in the direction of its arrow; otherwise multiply the column by -1. These column multiplications will multiply the determinant of $\boldsymbol{M}_0[R]$ by ± 1 . It is easy to see (check a few small examples to convince yourself) that every row of this modified $\boldsymbol{M}_0[R]$ has the sum of its elements equal to 0. Hence the sum of all the columns is 0. Thus in $\boldsymbol{M}_0[S]$ we have a set of columns for which a linear combination with coefficients ± 1 is 0 (the column vector of all 0's). Hence the columns of $\boldsymbol{M}_0[S]$ are linearly dependent, so det $\boldsymbol{M}_0[S] = 0$, as claimed.

Now suppose that S is the set of edges of a spanning tree T. Let e be an edge of T which is connected to v_p (the vertex which indexed the bottom row of \mathbf{M} , i.e., the row removed to get \mathbf{M}_0). The column of $\mathbf{M}_0[S]$ indexed by e contains exactly one nonzero entry [why?], which is ± 1 . Remove from $\mathbf{M}_0[S]$ the row and column containing the nonzero entry of column e, obtaining a $(p-2) \times (p-2)$ matrix \mathbf{M}'_0 . Note that $\det(\mathbf{M}_0[S]) = \pm \det(\mathbf{M}'_0)$ [why?]. Let T' be the tree obtained from T by contracting the edge e to a single vertex (so that v_p and the remaining vertex of e are merged into a single vertex u). Then \mathbf{M}'_0 is just the matrix obtained from the incidence matrix $\mathbf{M}(T')$ by removing the row indexed by u [why?]. Hence by induction on the number p of vertices (the case p = 1 being trivial), we have $\det(\mathbf{M}'_0) = \pm 1$. Thus $\det(\mathbf{M}_0[S]) = \pm 1$, and the proof follows. \Box

We have now assembled all the ingredients for the main result of this section (due originally to Borchardt). Recall that $\kappa(G)$ denotes the number of spanning trees of G.

1.8 Theorem. (the Matrix-Tree Theorem) Let G be a finite connected graph without loops, with laplacian matrix $\mathbf{L} = \mathbf{L}(G)$. Let \mathbf{L}_0 denote \mathbf{L} with the last row and column removed (or with the ith row and column removed for any i). Then

$$\det(\boldsymbol{L}_0) = \kappa(G).$$

Proof. Since $L = MM^t$ (Lemma 1.6(a)), it follows immediately that $L_0 = M_0 M_0^t$. Hence by the Binet-Cauchy theorem (Theorem 1.4), we have

$$\det(\boldsymbol{L}_0) = \sum_{S} (\det \boldsymbol{M}_0[S]) (\det \boldsymbol{M}_0^t[S]), \qquad (4)$$

where S ranges over all (p-1)-element subsets of $\{1, 2, \ldots, q\}$ (or equivalently, over all (p-1)-element subsets of the set of edges of G). Since in general

 $A^t[S] = A[S]^t$, equation (4) becomes

$$\det(\boldsymbol{L}_0) = \sum_{S} (\det \boldsymbol{M}_0[S])^2.$$
(5)

According to Lemma 1.7, $\det(\mathbf{M}_0[S])$ is ± 1 if S forms the set of edges of a spanning tree of G, and is 0 otherwise. Therefore the term indexed by S in the sum on the right-hand side of (5) is 1 if S forms the set of edges of a spanning tree of G, and is 0 otherwise. Hence the sum is equal to $\kappa(G)$, as desired. \Box

The operation of removing a row and column from L(G) may seem somewhat contrived. We would prefer a description of $\kappa(G)$ directly in terms of L(G). Such a description will follow from the next lemma.

1.9 Lemma. Let M be a $p \times p$ matrix with real entries such that the sum of the entries in every row and column is 0. Let M_0 be the matrix obtained from M by removing the last row and last column (or more generally, any row and any column). Then the coefficient of x in the characteristic polynomial $\det(M - xI)$ of M is equal to $-p \cdot \det(M_0)$. (Moreover, the constant term of $\det(M - xI)$ is 0.)

Proof. The constant term of det(M - xI) is det(M), which is 0 since the rows of M sum to 0.

For simplicity we prove the rest of the lemma only for removing the last row and column, though the proof works just as well for any row and column. Add all the rows of M - xI except the last row to the last row. This doesn't effect the determinant, and will change the entries of the last row all to -x(since the rows of M sum to 0). Factor out -x from the last row, yielding a matrix N(x) satisfying det $(M-xI) = -x \det(N(x))$. Hence the coefficient of x in det(M - xI) is given by $- \det(N(0))$. Now add all the columns of N(0)except the last column to the last column. This does not effect det(N(0)). Because the columns of M sum to 0, the last column of N(0) becomes the column vector $[0, 0, \ldots, 0, p]^t$. Expanding the determinant by the last column shows that det $(N(0)) = p \cdot \det(M_0)$, and the proof follows. \Box

1.10 Corollary. (a) Let G be a connected (loopless) graph with p vertices. Suppose that the eigenvalues of L(G) are $\mu_1, \ldots, \mu_{p-1}, \mu_p$, with $\mu_p =$

 $0. \ Then$

$$\kappa(G) = \frac{1}{p}\mu_1\mu_2\cdots\mu_{p-1}$$

(b) Suppose that G is also regular of degree d, and that the eigenvalues of $\mathbf{A}(G)$ are $\lambda_1, \ldots, \lambda_{p-1}, \lambda_p$, with $\lambda_p = d$. Then

$$\kappa(G) = \frac{1}{p}(d - \lambda_1)(d - \lambda_2)\cdots(d - \lambda_{p-1}).$$

Proof. (a) We have

$$\det(\boldsymbol{L} - xI) = (\mu_1 - x) \cdots (\mu_{p-1} - x)(\mu_p - x)$$

= $-(\mu_1 - x)(\mu_2 - x) \cdots (\mu_{p-1} - x)x.$

Hence the coefficient of x is $-\mu_1\mu_2\cdots\mu_{p-1}$. By Lemma 1.9, we get

$$-\mu_1\mu_2\cdots\mu_{p-1}=p\cdot\det(\boldsymbol{L}_0).$$

By Theorem 1.8 we have $det(\mathbf{L}_0) = \kappa(G)$, and the proof follows.

(b) Immediate from (a) and Lemma 1.6(b). \Box

Let us look at a couple of examples of the use of the Matrix-Tree Theorem.

1.11 Example. Let $G = K_p$, the complete graph on p vertices. Now K_p is regular of degree d = p - 1, and

$$\boldsymbol{A}(K_p) + I = J,$$

the $p \times p$ matrix of all 1's. Note that rank(J) = 1 [why?], so p-1 eigenvalues of J are equal to 0. Since trace(J) = p and the sum of the eigenvalues equals the trace, the remaining eigenvalue of J is p. Thus the eigenvalues of $A(K_p) = J-I$ are -1 (p-1 times) and p-1 (once). Hence from Corollary 1.10 there follows

$$\kappa(K_p) = \frac{1}{p}((p-1) - (-1))^{p-1} = p^{p-2}.$$

This surprising result is often attributed to Cayley, who stated it without proof in 1889 (and even cited Borchardt explicitly). However, it was in fact stated by Sylvester in 1857, while a proof was published by Borchardt in 1860. It is clear that Cayley and Sylvester could have produced a proof if asked to do so. There are many other proofs known, including elegant combinatorial arguments due to Prüfer, Joyal, Pitman, and others.

1.12 Example. The *n*-cube C_n is the graph with vertex set \mathbb{Z}_2^n (the set of all *n*-tuples of 0's and 1's), and two vertices u and v are connected by an edge if they differ in exactly one component. Now C_n is regular of degree n, and it can be shown that its eigenvalues are n - 2i with multiplicity $\binom{n}{i}$ for $0 \le i \le n$. (See the solution to Exercise 10.18(d) of the text.) Hence from Corollary 1.10(b) there follows the amazing result

$$\kappa(C_n) = \frac{1}{2^n} \prod_{i=1}^n (2i)^{\binom{n}{i}} \\ = 2^{2^n - n - 1} \prod_{i=1}^n i^{\binom{n}{i}}$$

To my knowledge a direct combinatorial proof is not known.

18.314 (Fall 2011)

Problems on the Matrix-Tree Theorem

- 1. The complete bipartite graph K_{rs} has vertex set $A \cup B$, where #A = r, #B = s, and $A \cap B = \emptyset$. There is an edge between every vertex of A and every vertex of B, so rs edges in all. Let $\mathbf{L} = \mathbf{L}(K_{rs})$ be the laplacian matrix of K_{rs} .
 - (a) Find a simple upper bound on $\operatorname{rank}(\boldsymbol{L} rI)$. Deduce a lower bound on the number of eigenvalues of \boldsymbol{L} equal to r.
 - (b) Assume $r \neq s$, and do the same as (a) for s instead of r.
 - (c) Find the remaining eigenvalues of L. (HINT. Use the fact that the rows of L sum to 0, and compute the trace of L.)
 - (d) Use (a)–(c) to compute $\kappa(K_{rs})$, the number of spanning trees of K_{rs} .
 - (e) (optional) Give a combinatorial proof of the formula for $\kappa(K_{rs})$.
- 2. Let V be the subset of $\mathbb{Z} \times \mathbb{Z}$ on or inside some simple closed polygonal curve whose vertices belong to $\mathbb{Z} \times \mathbb{Z}$, such that every line segment that makes up the curve is parallel to either the x-axis or y-axis. Draw an edge e between any two points of V at distance one apart, provided e lies on or inside the boundary curve. We obtain a planar graph G, an example being



Let G' be the dual graph G^* with the "outside" vertex deleted. (The vertices of G^* are the regions of G. For each edge e of G, say with regions R and R' on the two sides of e, there is an edge of G^* between R and R'.) For the above example, G' is given by



Let $\lambda_1, \ldots, \lambda_p$ denote the eigenvalues of G' (i.e., of the adjacency matrix A(G')). Show that

$$\kappa(G) = \prod_{i=1}^{p} (4 - \lambda_i).$$

18.314 Combinatorial Analysis Fall 2014

For information about citing these materials or our Terms of Use, visit: http://ocw.mit.edu/terms.