Notes on the Matrix-Tree theorem and Cayley's tree enumerator

1. Cayley's tree enumerator

Recall that the *degree* of a vertex in a tree (or in any graph) is the number of edges emanating from it. We will determine the generating function enumerating labelled trees on the vertex set $[n] = \{1, 2, ..., n\}$, weighted by their vertex degrees. Thus we introduce variables $x_1, ..., x_n$ corresponding to the vertices, and associate to every tree T the monomial

$$x_T = x_1^{d_1(T)} \cdots x_n^{d_n(T)},$$

where $d_i(T)$ denotes the degree of vertex *i* in *T*. Note that we can also write the same monomial as the product over all edges of *T*

$$x_T = \prod_{\{i,j\} \in E(T)} x_i x_j.$$

This is the same thing because each variable x_i appears once in the above product for every edge $e \in E(T)$ that contains the vertex *i*. Our desired generating function is the sum of these monomials over all spanning trees *T* on the vertex set [n]. It is a polynomial in the variables x_i . Since each tree has n-1 edges, we see from the second formula for x_T that in fact the generating function is a homogeneous polynomial of degree 2n-2. The following remarkable formula for it was discovered by Cayley.

Theorem 1. The generating function enumerating trees on [n] by the degrees of the various vertices is given by

$$\sum_{T} x_{T} = x_{1}x_{2}\cdots x_{n}(x_{1}+x_{2}+\cdots+x_{n})^{n-2}.$$

We will now prove Cayley's formula by means of a slightly subtle generating function argrument. One nice thing about this method of proof is it can be adapted, as we will do in the next section, to prove an even more powerful formula, the matrix-tree theorem.

Let us define

$$C_n(\mathbf{x}) = \sum_T x_T$$

to be our desired generating function. We will show that the two polynomials $C_n(\mathbf{x})$ and $x_1 \cdots x_n (x_1 + \cdots + x_n)^{n-2}$ have enough properties in common to imply that they are actually equal. Let's begin by observing some of these common properties.

(1) Both polynomials are homogeneous of degree 2n - 2. We won't actually need this in the proof, but at least it is encouraging. We have already seen that this is so for $C_n(\mathbf{x})$, and it is obvious for $x_1 \cdots x_n (x_1 + \cdots + x_n)^{n-2}$.

(2) For n > 1, both polynomials are divisible by $x_1 \cdots x_n$. This is again obvious for $x_1 \cdots x_n(x_1 + \cdots + x_n)^{n-2}$. For $C_n(\mathbf{x})$ it follows from the fact that every vertex in a tree has degree at least 1, since trees are connected. Incidentally, this is not true for n = 1, when $C_n(\mathbf{x}) = 1$. However, Cayley's formula remains correct for n = 1, since it reduces to $x_1(x_1)^{-1} = 1$.

(3) For n > 1, both polynomials have the property that every term contains some variable x_i to exactly the first power. For $C_n(\mathbf{x})$, this just says that every tree has a vertex of degree 1, that is, a "leaf." For $x_1 \cdots x_n (x_1 + \cdots + x_n)^{n-2}$ it follows from the fact that the factor $(x_1 + \cdots + x_n)^{n-2}$ has degree less than the number of variables, and therefore each of its terms omits at least one variable.

This last property is the key one for our proof of the formula. To show that the two polynomials are the same, we must show that every monomial has the same coefficient in both polynomials. By property (3), it is sufficient to consider only monomials containing x_i to exactly the first power, for each i.

Now we will actually prove the theorem, by induction on n. The base case is n = 1, which we have seen holds. For n > 1, we assume the formula holds for trees on n - 1 vertices. Using property (3), we can fix an index i and consider only those terms containing x_i to the first power. By property (2), every term in both polynomials contains x_i to at least the first power, so we can extract the terms we want by first dividing by x_i and then setting $x_i = 0$ in what remains. In other words, we need to show that

(1)
$$(x_i^{-1}C_n(\mathbf{x}))_{x_i\mapsto 0} = (x_i^{-1}x_1\cdots x_n(x_1+\cdots+x_n)^{n-2})_{x_i\mapsto 0}$$

By symmetry, it is enough to show this for i = n, since all the variables play identical roles. Then the right hand side above simplifies to

$$x_1 \cdots x_{n-1} (x_1 + \cdots + x_{n-1})^{n-2},$$

which is equal to

$$(x_1 + \dots + x_{n-1})C_{n-1}(x_1, \dots, x_{n-1})$$

by the induction hypothesis. On the left hand side in (1) we have the sum of all terms in $C_n(\mathbf{x})$ containing x_n to the first power, except that this has been divided by x_n . Note that the terms in question are exactly those for which the vertex n is a leaf of the tree T. Hence we need to show that

(2)
$$\sum_{T: n \text{ is a leaf}} x_T = x_n (x_1 + \dots + x_{n-1}) C_{n-1} (x_1, \dots, x_{n-1}).$$

Now, this is easy to establish combinatorially. To pick a tree with vertex n as a leaf, we can first choose any tree T' on the vertex set [n-1], then connect n by an additional edge to any vertex in T'. The generating function for the possible choices of T' is $C_{n-1}(x_1, \ldots, x_{n-1})$. The generating function for the choice of an edge $\{i, n\}$ with $1 \le i \le n$ is $x_1x_n + \cdots + x_{n-1}x_n = x_n(x_1 + \cdots + x_{n-1})$. Multiplying these gives the right hand side in (2).

2. Corollaries to Cayley's formula

Setting all the variables x_i equal to 1 in Theorem 1, we get the number of labelled trees on n vertices.

Corollary 1. The number of labelled trees on n vertices is n^{n-2} .

If we want to count labelled rooted trees instead, we can just multiply by n for the possible choices of the root.

Corollary 2. The number of labelled rooted trees on n vertices is n^{n-1} .

We can also use Cayley's formula to count rooted *forests*. A forest is a graph with no cycles that is not necessarily connected. Thus its connected components are trees. It is a *rooted* forest if we specify a root in each component. Given a rooted forest F on [n], we can construct a rooted tree T on $\{0, 1, \ldots, n\}$ with 0 as the root by connecting vertex 0 by an edge to the root of each component of the forest F. Conversely, given a tree T on $\{0, 1, \ldots, n\}$, we can delete 0 and the edges containing it to get a forest F on [n], and record which vertices were connected to 0 in T by specifying them as the roots of the components of F. It is easy to see that these two constructions provide a bijection between labelled trees on $\{0, 1, \ldots, n\}$ (with fixed root at 0), and rooted forests on [n] (with any roots). Hence we have the following corollary.

Corollary 3. The number of labelled rooted forests on n vertices is $(n+1)^{n-1}$.

If we keep track of the weights in the trees-to-forests bijection we can get some additional information. For example, in the Cayley formula enumerating trees on $\{0, 1, \ldots, n\}$ suppose we set all the variables except x_0 to 1. Then we will enumerate the trees by the degree of vertex 0, which is the same as the number of components in the corresponding forest. The resulting enumerator is

$$x_0(x_0+n)^{n-1}$$

Extracting the coefficient of x_0^k using the binomial theorem, we obtain the following result.

Corollary 4. The number of labelled rooted forests on n vertices with exactly k components is

$$\binom{n-1}{k-1}n^{n-k}.$$

Note that for k = 1 this gives the same result as Corollary 2.

3. The matrix-tree theorem

In Cayley's formula, the monomial x_T keeps track of the vertex degrees in the tree T. This is quite a bit of information, but not enough to determine the tree. The matrix-tree theorem is a more refined formula that gives, in effect, the complete symbolic series for labelled trees, and more generally for labelled forests with specified roots. In this formula we will attach to each forest F a monomial that actually contains a complete description of F.

Consider the example shown here of a rooted forest on the vertex set [8].



We have directed every edge toward the root of the component containing it, that is, from child to parent. We introduce variables x_{ij} for all $i, j \in [n]$ with $i \neq j$, and define the forest monomial x_F to be the product of the variables x_{ij} for all directed edges $i \to j$ in F. Note that x_{21} and x_{12} , for example, are different variables in this setting. Thus for the example forest above, we have

$$x_F = x_{56} x_{62} x_{32} x_{18} x_{78}.$$

Note that the rooted forest F is determined by its monomial x_F . The variables occuring in x_F tell us the edges of F, and their directions tell us the roots. In fact, the roots are just the vertices ifor which no variable x_{ij} with first index i occurs in x_F , that is, the vertices with no parent in F. Given a subset $I \subseteq [n]$, we define $F_{n,I}(\mathbf{x})$ to be the generating function for all forests whose set of roots is I, that is

$$F_{n,I}(\mathbf{x}) = \sum_{F: \text{roots}(F)=I} x_F.$$

For example, by taking $I = \{1\}$ we get a generating function for spanning trees on [n], with vertex 1 fixed as the root (this only serves to determine the directions of the edges).

The matrix-tree theorem gives $F_{n,I}(\mathbf{x})$ as the determinant of a submatrix of the following $n \times n$ matrix:

$$M_{n}(\mathbf{x}) = \begin{bmatrix} (x_{12} + \dots + x_{1n}) & -x_{12} & -x_{13} & \dots & -x_{1n} \\ -x_{21} & (x_{21} + x_{23} + \dots + x_{2n}) & -x_{23} & \dots & -x_{2n} \\ \vdots & & \ddots & \vdots \\ -x_{n1} & -x_{n2} & -x_{n,n-1} & \dots & (x_{n1} + \dots + x_{n,n-1}) \end{bmatrix}$$

The rule of construction for $M_n(\mathbf{x})$ is that the off-diagonal entry in position (i, j) is $-x_{ij}$, while the diagonal entries are such that every row sums to zero. In other words, the diagonal entry in position (i, i) is the sum of the variables x_{ij} for all $j \neq i$.

Theorem 2. The generating function $F_{n,I}(\mathbf{x})$ for forests rooted at I, with edges directed towards the roots, is given by the determinant

$$F_{n,I}(\mathbf{x}) = \det M_{n,I}(\mathbf{x}),$$

where $M_{n,I}(\mathbf{x})$ is the square submatrix of the matrix $M_n(\mathbf{x})$ above, gotten by deleting the rows and columns with indices $i \in I$.

Before turning to the proof, let us clarify the meaning of this with some examples.

Example: If $I = \emptyset$ and n > 0, then $F_{n,\emptyset}(\mathbf{x}) = 0$, since every forest must have at least one root. According to the theorem, we should also have det $M_n(\mathbf{x}) = 0$, since $M_{n,\emptyset}(\mathbf{x}) = M_n(\mathbf{x})$. This is correct because the sum of the columns of $M_n(\mathbf{x})$ is zero. Recall that if the columns of a matrix are linearly dependent, then its determinant vanishes.

Example: If I = [n], there is just one forest, with no edges, so $F_{n,[n]}(\mathbf{x}) = 1$. The matrix $M_{n,[n]}(\mathbf{x})$ is the empty matrix with no rows or columns. By convention, the determinant of an empty matrix is 1. This, by the way, is the correct convention for consistency with properties of the determinant such as the Laplace expansion or the fact that the determinant is the product of the eigenvalues.

Example: If $I = \{1, 2, ..., n-1\}$, we get a forest with roots I by attaching vertex n by a single edge to one of the vertices 1 through n-1. The forest is the graph with only the one edge $\{n, j\}$, and it is directed from n to j, so its monomial is x_{nj} . Hence $F_{n,I}(\mathbf{x}) = x_{n1} + x_{n2} + \cdots + x_{n,n-1}$ in this case. To get $M_{n,I}(\mathbf{x})$ we delete all rows and columns of $M_n(\mathbf{x})$ except the last, leaving only the diagonal entry in position (n, n). This entry is indeed equal to $x_{n1} + x_{n2} + \cdots + x_{n,n-1}$. Similar reasoning with $I = [n] \setminus \{j\}$ for any j "explains" the other diagonal entries of $M_n(\mathbf{x})$.

Example: To get an example involving a non-trivial determinant we must have I nonempty and n - |I| > 1. So the first non-trivial example is n = 3, $I = \{3\}$. There are three forests, shown here.



Adding up their monomials, we obtain the generating function

$$F_{3,\{3\}}(\mathbf{x}) = x_{12}x_{23} + x_{21}x_{13} + x_{13}x_{23}$$

Applying the theorem, we delete the last row and column from $M_3(\mathbf{x})$, getting the determinant

det
$$\begin{bmatrix} x_{12} + x_{13} & -x_{12} \\ -x_{21} & x_{21} + x_{23} \end{bmatrix} = (x_{12} + x_{13})(x_{21} + x_{23}) - x_{12}x_{21},$$

which agrees with the expression we found for $F_{3,\{3\}}(\mathbf{x})$.

Now we will prove Theorem 2. As with the Cayley formula, we first observe that $F_{n,I}(\mathbf{x})$ and det $M_{n,I}(\mathbf{x})$ have a property in common that will allow us to eliminate some of the variables and reduce to the n-1 case, which we can assume holds by induction. In the examples we have already seen that the theorem holds when $I = \emptyset$, so we will assume that $I \neq \emptyset$.

We claim that every term of both $F_{n,I}(\mathbf{x})$ and det $M_{n,I}(x)$ has the property that for some j, none of the variables x_{ij} with second index j occurs in that term. For $F_{n,I}(\mathbf{x})$ this is true because for the term x_F we can take j to be a leaf of one of the trees in the forest F. Then F has no edge directed into j, so no variable x_{ij} occurs in x_F . For $M_{n,I}(x)$ we observe that all the entries of the matrix are linear in the variables, and hence the determinant is a homogeneous polynomial of degree n - |I|. Since we are assuming $I \neq \emptyset$, this degree is less than n. If a monomial contains some variable x_{ij} for every index j, then its degree is at least n, so no such monomial appears in det $M_{n,I}(\mathbf{x})$.

Having established that every term in both polynomials omits all the variables x_{ij} for some j, it follows that it is enough to verify that both polynomials give the same thing on setting all $x_{ij} = 0$, for each j. By symmetry, it is enough to verify this for j = n. (Note that the statement of the theorem for a given I is not symmetric in all the indices 1 through n, since some indices belong to I and others do not. However, the statement as a whole for all I is symmetric in the vertex indices. Thus we can safely assume j = n, provided we verify the result for all I).

Setting all $x_{in} = 0$ in $F_{n,I}(\mathbf{x})$ kills off any terms x_F belonging to forests F with an edge directed into vertex n, leaving the generating function for rooted forests in which vertex n is a leaf—that is, in which vertex n has no children. If $n \in I$, then n is both a root and a leaf, so we get the enumerator for forests in which n is an isolated vertex, which is the same as the enumerator for forests on [n-1] with roots $J = I \setminus \{n\}$. If $n \notin I$, then we can choose a forest in which n is a leaf but not a root by first choosing any forest on [n-1] with roots I, and then attaching n by an edge to any vertex in [n-1]. This introduces a factor $x_{n1} + \cdots + x_{n,n-1}$, with one term x_{nj} for each vertex j that we might connect n to. To summarize, we have

$$F_{n,I}(\mathbf{x})_{x_{in}\mapsto 0} = \begin{cases} F_{n-1,J}(\mathbf{x}), & \text{if } n \in I, \text{ where } J = I \setminus \{n\} \\ (x_{n1} + \dots + x_{n,n-1})F_{n-1,I}(\mathbf{x}) & \text{if } n \notin I. \end{cases}$$

Setting all $x_{in} = 0$ in $M_n(\mathbf{x})$ gives a matrix which can be described in block form as

$$M_n(\mathbf{x})_{x_{in}\to 0} = \left[\begin{array}{c|c} M_{n-1}(\mathbf{x}) & 0\\ \hline -x_{n1} \cdots - x_{n,n-1} & (x_{n1} + \cdots + x_{n,n-1}) \end{array} \right].$$

Here we have an $(n-1) \times (n-1)$ square block equal to $M_{n-1}(\mathbf{x})$, a column of n-1 zeros to its right, and the bottom row of the original $M_n(\mathbf{x})$ is unchanged. If $n \in I$, we will delete the last row and column, giving

$$\det M_{n,I}(\mathbf{x})_{x_{in}\mapsto 0} = \det M_{n-1,J}(\mathbf{x}),$$

where $J = I \setminus \{n\}$. This is equal to $F_{n-1,J}(\mathbf{x})$ by induction, and hence to $F_{n,I}(\mathbf{x})_{x_{in}\mapsto 0}$ by the calculation above. Alternatively, if $n \notin I$, then we keep the last column in the matrix $M_n(\mathbf{x})_{x_{in}\mapsto 0}$ above. Using the Laplace expansion of the determinant along this last column, we get $(x_{n1} + \cdots + x_{n,n-1})$ times the determinant of the block $M_{n-1}(\mathbf{x})$ with its rows and columns for $i \in I$ deleted, that is,

$$\det M_{n,I}(\mathbf{x})_{x_{in}\mapsto 0} = (x_{n1} + \dots + x_{n,n-1}) \det M_{n-1,I}(\mathbf{x}).$$

Again, this is equal to $F_{n,I}(\mathbf{x})_{x_{in}\mapsto 0}$ by induction and the previous calculation. So we have shown that

$$\det M_{n,I}(\mathbf{x})_{x_{in}\mapsto 0} = F_{n,I}(\mathbf{x})_{x_{in}\mapsto 0}$$

in either case—whether $n \in I$ or not—and we have already seen that this is sufficient to prove the theorem.

4. Corollaries to the Matrix-Tree Theorem

We will use the matrix-tree theorem to again obtain the formula n^{n-2} for the number of labelled spanning trees on n vertices. For this purpose we take $I = \{n\}$ and set all the variables x_{ij} equal to 1. Note that any one-element set $I = \{j\}$ would give the same result, since we can root our spanning trees at any vertex. The matrix $M_{n,I}(\mathbf{x})$ now specializes to the $(n-1) \times (n-1)$ matrix with diagonal entries n-1 and off-diagonal entries -1. We can write it as

$$nI_{n-1} - J_{n-1},$$

where I_{n-1} is the identity matrix and J_{n-1} is the matrix whose entries are all equal to 1. Now recall from linear algebra that the determinant of any matrix is the product of its eigenvalues. The matrix J_{n-1} has n-2 linearly independent nullvectors, that is, vectors \mathbf{v} such that $\mathbf{v}J_{n-1} = 0$. In fact, the n-2 vectors of the form

$$\mathbf{v} = [1, 0, \dots, 0, -1, 0, \dots, 0]$$

are linearly independent nullvectors. A nullvector is an eigenvector with eigenvalue zero, so J_{n-1} has the eigenvalue zero with multiplicity n-2 (for convenience, I am using *left* eigenvectors here, that is, an eigenvector of A is a row vector **v** such that **v**A is a scalar multiple of **v**). The vector

$$\mathbf{v} = [1, \dots, 1]$$

is also an eigenvector, with eigenvalue n - 1, since $\mathbf{v}J_{n-1} = (n-1)\mathbf{v}$. So the complete list of eigenvalues of J_{n-1} consists of 0, with multiplicity n - 2, and n - 1, with multiplicity one. The eigenvalues of $-J_{n-1}$ are the negatives of these. Adding nI_{n-1} adds n to every eigenvalue, so $M_{n,I}(1)$ has eigenvalues n, repeated n - 2 times, and 1, once. Its determinant is therefore n^{n-2} .

By similar reasoning, we can derive Cayley's formula as a corollary to the matrix-tree theorem. This time, instead of setting all $x_{ij} = 1$, we introduce new variables x_1, \ldots, x_n as in Cayley's formula, and set $x_{ij} = x_i x_j$. As before, we take I to consist of any one vertex, say vertex n. The forests F rooted at n are just the spanning trees T, and our forest monomial x_F goes to x_T when we set $x_{ij} = x_i x_j$. The matrix $M_{n,I}(\mathbf{x})$ now specializes to

$$\begin{bmatrix} x_1(x_2 + \dots + x_n) & -x_1x_2 & \dots & -x_1x_{n-1} \\ -x_2x_1 & x_2(x_1 + x_3 + \dots + x_n) & \dots & -x_2x_{n-1} \\ \vdots & & \ddots & \vdots \\ -x_{n-1}x_1 & -x_{n-1}x_2 & \dots & x_{n-1}(x_1 + \dots + x_{n-2} + x_n) \end{bmatrix}$$

which can also be written in the form

$$M_{n,I}(x_i x_j) = D((x_1 + \dots + x_n)I_{n-1} - J_{n-1}D),$$

where D is the diagonal matrix with diagonal entries $x_1, x_2, \ldots, x_{n-1}$. The determinant of D is $x_1 \cdots x_{n-1}$, so the determinant of $M_{n,I}(x_i x_j)$ is given by

$$x_1 \cdots x_{n-1} \det((x_1 + \cdots + x_n)I_{n-1} - J_{n-1}D).$$

Just as before, the n-2 vectors $[1, 0, \ldots, 0, -1, 0, \ldots, 0]$ are independent left nullvectors of $J_{n-1}D$.

The vector $\mathbf{v} = [1, \ldots, 1]$ satisfies $\mathbf{v}DJ_{n-1} = (x_1 + \cdots + x_{n-1})\mathbf{v}$, so it is a left eigenvector of DJ_{n-1} with eigenvalue $(x_1 + \cdots + x_{n-1})$. Now D and J_{n-1} are symmetric matrices, so DJ_{n-1} is the transpose of $J_{n-1}D$. Since a matrix at its transpose have the same eigenvalues, it follows that $(x_1 + \cdots + x_{n-1})$ is also an eigenvalue of $J_{n-1}D$. Since we saw previously that $J_{n-1}D$ has n-2 independent left nullvectors, it follows that its eigenvalues are zero, with multiplicity n-2, and $(x_1 + \cdots + x_{n-1})$, with multiplicity one. Negating these eigenvalues and adding $(x_1 + \cdots + x_n)$ to all of them, we see that the eigenvalues of $(x_1 + \cdots + x_n)I_{n-1} - J_{n-1}D$ are $(x_1 + \cdots + x_n)$,

with multiplicity n-2, and x_n , with multiplicity one. Finally, multiplying the eigenvalues together along with the factor $x_1 \cdots x_{n-1}$ yields Cayley's formula

$$x_1 x_2 \cdots x_n (x_1 + x_2 + \cdots + x_n)^{n-2}$$

for the generating function.

One advantage of the matrix-tree theorem is that we can use it to count forests as well as trees. Fix a set of roots I with k elements. To count all forests with exactly k components and the specified roots I, we can set the variables x_{ij} equal to 1 in $M_{n,I}(\mathbf{x})$ just as we did above. In this case we get the matrix

$$M_{n,I}(1) = nI_{n-k} - J_{n-k}.$$

The eigenvalues of $-J_{n-k}$ are n-k-1 zeroes and an n-k, by the same reasoning as before, so the eigenvalues of $M_{n,I}(1)$ are n, repeated n-k-1 times, and k, once. This yields the following result.

Corollary 5. The number of labelled rooted forests on n vertices with exactly k components, and with roots specified in advance, is

 kn^{n-k-1} .

It is instructive to compare this with Corollary 4, where we computed the corresponding number with unspecified roots. To compute the number in Corollary 4 we could first pick the k roots, in $\binom{n}{k}$ ways, then choose a forest with these specified roots. The resulting number $\binom{n}{k}kn^{n-k-1}$ agrees with Corollary 4 after a little algebraic manipulation. Neither Corollary is helpful if we want to count *unrooted* forests, however. Eventually we will need exponential generating functions to accomplish that.

To close, here is a curious application of the matrix-tree theorem to deduce a general theorem of linear algebra. If we set $x_{ij} = x_{ji}$ for each $\{i, j\}$, the matrix $M_n(\mathbf{x})$ becomes a symmetric matrix, with all row (and column) sums equal to zero. What's more, since there is a separate variable x_{ij} for each off-diagonal position (up to transpose), we can get any symmetric matrix A whatsoever, provided it has row and column sums zero, by setting the variables x_{ij} to the corresponding entries of A. For a one-element set of roots $I = \{j\}$, the determinant of $M_{n,\{j\}}(\mathbf{x})$ enumerates spanning trees on [n], with their edges directed towards vertex j. But once we set $x_{ij} = x_{ji}$, the directions of edges don't matter, so this determinant does not depend on which j we take. This proves the following result.

Corollary 6. If A is a symmetric matrix with all row and column sums equal to zero, and A_j denotes the submatrix gotten by deleting row and column j, then det A_j is the same for every j.

By the way, for any symmetric matrix A, without restriction on the row and column sums, the determinants det A_i are called *principal minors* of A.