Lecture 10

Proof of the Matrix-Tree Theorem

The proof here is derived from a terse account in the lecture notes from a course on Algebraic Combinatorics taught by Lionel Levine at MIT in Spring 2011. I studied them with Samantha Barlow, a former Discrete Maths student who did a third-year project with me in 2011-12.

Reading:
I don’t know of any textbook accounts of the proof given here, but the intrepid reader might like to look at the following two articles, both of which make the connection between the Principle of Inclusion/Exclusion and Tutte’s Matrix Tree theorem.


10.1 Single predecessor graphs

Before we plunge into the proof itself I’d like to define a certain family of graphs that includes, but is larger than, the family of spanning arborescences.

**Definition 10.1.** A *single predecessor graph* ("spreg") with distinguished vertex $v$ in a digraph $G(V, E)$ is a subgraph $T(V, E')$ (Note that $T$ and $G$ have the same vertex set) in which each vertex other than the distinguished vertex $v$ has exactly one predecessor while $v$ itself has no predecessors. Equivalently,

$$\deg_m(v) = 0 \quad \text{and} \quad \deg_m(u) = 1 \ \forall u \neq v \in V.$$  

Figure 10.1 includes several examples of spregs, including two that are arborescences, which prompts the following proposition:
Proposition 10.2 (Spanning arborescences are spregs). If \( T(V,E) \) is a spanning arborescence for \( G(V,E) \) with root \( v \), then it is also a spreg with distinguished vertex \( v \).

Proof. By definition, \( G \) and \( T \) share the same vertex set, so all we need check that the vertices \( u \neq v \) in \( T \) have a single predecessor. Recall that an arborescence rooted at \( v \) is a directed graph \( T(V,E) \) such that

(i) Every vertex \( u \neq v \) is accessible from \( v \). That is, there is a directed path from \( v \) to every other vertex.

(ii) \( T \) becomes an ordinary, undirected tree if we ignore the directedness of the edges.

The proposition consists of two separate claims: that \( \deg_{in}(v) = 0 \) and that \( \deg_{out}(u) = 1 \forall u \neq v \in V \). We’ll prove both by contradiction.

Suppose that \( \deg_{in}(v) > 0 \); it’s then easy to see that \( T \) must include a directed cycle. Consider one of \( v \)’s predecessors—call is \( u_0 \). It is accessible from \( v \), so there is a directed path from \( v \) to \( u_0 \). And \( u_0 \) is a predecessor of \( v \), so there is also a directed edge \((u_0,v) \in E\). If we append this edge to the end of the path, we get a directed path from \( v \) back to itself. This contradicts the second property of an arborescence and so we must have \( \deg_{in}(v) = 0 \).

The proof for the second part of the proposition is illustrated in Figure 10.2. Suppose that \( \exists u \neq v \in V \) such that \( \deg_{out}(u) \geq 2 \) and choose two distinct predecessors of \( u \); call them \( v_1 \) and \( v_2 \) and note that one of them may be the root vertex \( v \). Now consider the directed paths from \( v \) to \( v_1 \) and \( v_2 \). In the undirected version of \( T \) these paths, along with the edges \((v_1,u)\) and \((v_2,u)\), must include a cycle, which contradicts the second property of an arborescence.

Proposition 10.3 (Characterising spregs). A spreg with distinguished vertex \( v \) consists of a spanning arborescence rooted at \( v \), plus zero or more disjoint weakly connected components, each of which contains a single directed cycle.
Figure 10.2: Two examples to illustrate the second part of the proof that an arborescence is a spreg. If one ignores the directedness of the edges in the graphs above, both contain cycles.

The reasoning needed to prove this proposition is similar to that for the previous one and so is left to the Problem Sets.

This lemma is one of the key ingredients in the proof of Tutte’s Matrix-Tree Theorem. The idea is to first note that a spanning arborescence is a spreg. We then count the spanning arborescences contained in a graph by first counting all the spregs, then use the Principle of Inclusion/Exclusion to count—and subtract away—those spregs that contain one or more cycles.

10.2 Counting spregs with determinants

Recall that we’re trying to prove

**Theorem 1** (Tutte’s Directed Matrix-Tree Theorem, 1948). If \( G(V, E) \) is a digraph with vertex set \( V = \{v_1, \ldots, v_n\} \) and \( L \) is an \( n \times n \) matrix whose entries are given by

\[
L_{ij} = \begin{cases} 
\deg_{\text{in}}(v_j) & \text{If } i = j \\
-1 & \text{If } i \neq j \text{ and } (v_i, v_j) \in E \\
0 & \text{Otherwise}
\end{cases} \quad (10.1)
\]

then the number \( N_j \) of spanning arborescences with root at \( v_j \) is

\[
N_j = \det(\hat{L}_j)
\]

where \( \hat{L}_j \) is the matrix produced by deleting the \( j \)-th row and column from \( L \).

First note that—because we can always renumber the vertices before we apply the theorem—it is sufficient to prove the result for the case with root vertex \( v = v_n \). Now consider the representation of \( \det(\hat{L}_n) \) as a sum over permutations:

\[
\det(\hat{L}_n) \equiv \det(\mathcal{L}) = \sum_{\sigma \in S_{n-1}} \text{sgn}(\sigma) \prod_{j=1}^{n-1} L_{j, \sigma(j)}. \quad (10.2)
\]

where I have introduced the notation \( \mathcal{L} \equiv \hat{L}_n \) to avoid the confusion of having two kinds of subscripts on \( \hat{L}_n \). This means that \( \mathcal{L} \) is an \((n-1) \times (n-1)\) matrix in which

\[
\mathcal{L}_{ij} = L_{ij};
\]

where \( L_{ij} \) is the \( i, j \) entry in the matrix \( L \) defined by Eqn. [10.1] in the statement of Tutte’s theorem.
Predecessor of $v_1$ $v_2$ $v_3$ Is a spanning arboresence?
---
$v_2$ $v_1$ $v_2$ No
$v_2$ $v_3$ $v_2$ No
$v_2$ $v_4$ $v_2$ Yes
$v_4$ $v_1$ $v_2$ Yes
$v_4$ $v_3$ $v_2$ No
$v_4$ $v_4$ $v_2$ Yes

Table 10.1: Each row here corresponds to one of the spregs in Figure 10.3.

10.2.1 Counting spregs

In this section we’ll explore two examples that illustrate a connection between terms in the sum for $\det(\mathcal{L})$ and the business of counting various kinds of spregs.

The identity term: counting all spregs

In the case where $\sigma = \text{id}$, so that $\sigma(j) = j$ for all $j$, we have $\text{sgn}(\sigma) = 1$ and

$$\prod_{j=1}^{n-1} \mathcal{L}_{j\sigma(j)} = \prod_{j=1}^{n-1} \mathcal{L}_{jj} = \prod_{j=1}^{n-1} \deg_m(v_j).$$

(10.3)

This product is also equal to the total number of spregs in $G(V, E)$ that have distinguished vertex $v_n$. To see why, look back at the definition of a spreg and think about what we’d need to do if we wanted to write down a complete list of these spregs. We could specify a spreg by listing the single predecessor for each vertex other than $v_n$ in a table like the one below

<table>
<thead>
<tr>
<th>Vertex</th>
<th>$v_1$ $v_2$ $v_3$</th>
</tr>
</thead>
<tbody>
<tr>
<td>Predecessor</td>
<td>$v_2$ $v_1$ $v_2$</td>
</tr>
</tbody>
</table>

which describes one of the spregs rooted at $v_4$ contained in the four-vertex graph shown in Figure 10.3. And if we wanted to list all the four-vertex spregs contained in this graph we could start by assembling the predecessor lists of all the vertices other than the distinguished vertex,

$$P_1 = \{v_2, v_4\}, \quad P_2 = \{v_1, v_3, v_4\} \quad \text{and} \quad P_3 = \{v_2\},$$

where $P_j$ lists the predecessors of $v_j$. Then, to specify a spreg with distinguished vertex $v_4$ we would choose one entry from each of the predecessor lists, meaning that there are

$$|P_1| \times |P_2| \times |P_3| = \deg_m(v_1) \times \deg_m(v_2) \times \deg_m(v_3) = 2 \times 3 \times 1 = 6$$

such spregs in total. All six possibilities are listed in Table 10.1 and illustrated in Figure 10.3. The equation above also emphasises that $|P_j| = \deg_m(v_j)$ and so makes the connection with the product in Eqn. (10.3).
The graph $G(V, E)$ at upper left contains six spregs with distinguished vertex $v_4$, all of which are shown in the two rows below. Three of them are spanning arborescences rooted at $v_4$, while the three others contain cycles.

Terms that count spregs containing a single directed cycle

In the case where the permutation $\sigma$ contains a single cycle of length $\ell$, so that

$$\sigma = (i_1, \ldots, i_\ell),$$

we have $\text{sgn}(\sigma) = (-1)^{\ell-1}$ and

$$\prod_{j=1}^{n-1} L_{\sigma(j)} = \left( \prod_{j \in \text{fix}(\sigma)} L_{jj} \right) \times \left( \prod_{k=1}^{\ell} L_{i_k i_{k+1}} \right) = \left( \prod_{j \in \text{fix}(\sigma)} \text{deg}_{\text{in}}(v_j) \right) \times \left( \prod_{k=1}^{\ell} L_{i_k i_{k+1}} \right)$$

where the indices $i_k$ are to be understood periodically, so $i_{\ell+1} = i_1$. The factors $L_{i_k i_{k+1}}$ in the second of the two products above are off-diagonal entries of $L = \hat{L}_n$ and thus satisfy

$$L_{i_k i_{k+1}} = \begin{cases} -1 & \text{if } (v_{i_k}, v_{i_{k+1}}) \in E \\ 0 & \text{otherwise.} \end{cases}$$

Thus if one or more of the edges $(v_{i_k}, v_{i_{k+1}})$ is absent from the graph we have

$$\prod_{k=1}^{\ell} L_{i_k i_{k+1}} = 0,$$

but if all the edges $(v_{i_k}, v_{i_{k+1}})$ are present we have can make the following observations:
• the graph contains a directed cycle given by the vertex sequence 
\((v_{i_1}, \ldots, v_{i_\ell}, v_{i_1});
• \(L_{i_k i_{k+1}} = -1\) for all \(1 \leq k \leq \ell\) and so we have

\[
\text{sgn}(\sigma) \prod_{j=1}^{n-1} L_{j\sigma(j)} = (-1)^{\ell-1} \left( \prod_{j \in \text{fix}(\sigma)} \deg_{\text{in}}(v_j) \right) \times \prod_{k=1}^{\ell} (-1)
= (-1)^{\ell-1} \left( \prod_{j \in \text{fix}(\sigma)} \deg_{\text{in}}(v_j) \right) (-1)^\ell
= (-1)^{2\ell-1} \prod_{j \in \text{fix}(\sigma)} \deg_{\text{in}}(v_j)
= - \prod_{j \in \text{fix}(\sigma)} \deg_{\text{in}}(v_j). \tag{10.4}
\]

Arguments similar to those in the previous section then show that the product 
\(\prod_{j \in \text{fix}(\sigma)} \deg_{\text{in}}(v_j)\) in Eqn. (10.4) counts the number of ways to choose predecessors for those vertices that aren’t part of the cycle. We can summarise all these ideas with the following pair of results:

**Proposition 10.4.** For a permutation \(\sigma \in S_{n-1}\) consisting of a single cycle 
\(\sigma = (i_1, \ldots, i_\ell)\)
define an associated directed cycle \(C_\sigma\) specified by the vertex sequence \((v_{i_1}, \ldots, v_{i_\ell}, v_{i_1})\). Then the term in \(\det(L)\) corresponding to \(\sigma\) satisfies

\[
\text{sgn}(\sigma) \prod_{j=1}^{n-1} L_{j\sigma(j)} = \begin{cases} 
- \prod_{j \in \text{fix}(\sigma)} \deg_{\text{in}}(v_j) & \text{if } C_\sigma \subseteq G(V, E) \text{ and } \text{fix}(\sigma) \neq \emptyset \\
-1 & \text{if } C_\sigma \subseteq G(V, E) \text{ and } \text{fix}(\sigma) = \emptyset \\
0 & \text{if } C_\sigma \not\subseteq G(V, E)
\end{cases}
\]

**Corollary 10.5.** For \(\sigma\) and \(C_\sigma\) as in Proposition 10.4,

\[
\prod_{j=1}^{n-1} L_{j\sigma(j)} = |\{\text{sregs containing } C_\sigma\}|.
\]

### 10.2.2 An example

Before pressing on to generalise the results of the previous section to arbitrary permutations, let’s see what Corollary 10.5 allows us to say about the graph in Figure 10.3. There \(G(V, E)\) is a digraph on four vertices, so the determinant that comes into Tutte’s theorem is that of \(L_4\), a three-by-three matrix. We’ve already seen that if \(\sigma = \text{id}\) the product \(\prod_{j \in \text{fix}(\sigma)} \deg_{\text{in}}(v_j)\) gives six, the total number of sregs contained in the graph. The results for the remaining elements of \(S_3\) are listed in Table 10.2 and all are covered by Corollary 10.5 as all non-identity elements of \(S_3\) are single cycles.
| $\sigma$   | $C_\sigma$ | $C_\sigma \in G$? | $\left| \prod_{j=1}^{n-1} L_{j\sigma(j)} \right|$ | Number of spregs containing $C_\sigma$ |
|------------|------------|--------------------|---------------------------------|----------------------------------|
| (1,2)      | $(v_1, v_2, v_1)$ | Yes | $\deg_{\text{in}}(v_3) = 1$ | 1 |
| (1,3)      | $(v_1, v_3, v_1)$ | No | $\deg_{\text{in}}(v_2) \times 0$ | 0 |
| (2,3)      | $(v_2, v_3, v_2)$ | Yes | $\deg_{\text{in}}(v_1) = 2$ | 2 |
| (1,2,3)    | $(v_1, v_2, v_3, v_1)$ | No | 0 | 0 |
| (1,3,2)    | $(v_1, v_3, v_2, v_1)$ | No | 0 | 0 |

Table 10.2: The results of using Corollary 10.5 to count spregs containing the various cycles $C_\sigma$ associated with the non-identity elements of $S_3$. The right column lists the number one gets by direct counting of the spregs shown in Figure 10.3.

### 10.2.3 Counting spregs in general

Here we generalise the results from Section 10.2.1 to permutations that are the products of arbitrarily many cycles.

**Lemma 10.6** (Counting spregs containing cycles).

Suppose $\sigma \in S_{n-1}$ is the product of $k > 0$ disjoint cycles

$$\sigma = (i_{1,1}, \ldots, i_{1,\ell_1}) \cdots (i_{k,1}, \ldots, i_{k,\ell_k}),$$

where $\ell_j$ is the length of the $j$-th cycle. Associate the directed cycle $C_j$ defined by the vertex sequence $(v_{i_{j,1}}, \ldots, v_{i_{j,\ell}}, v_{i_{j,1}})$ with the $j$-th cycle in the permutation and define

$$C_\sigma = \bigcup_{j=1}^{k} C_j.$$

Then the term in $\det(L)$ corresponding to $\sigma$ satisfies

$$\text{sgn}(\sigma) \prod_{j=1}^{n-1} L_{j\sigma(j)} = \begin{cases} (-1)^k \prod_{j \in \text{fix(}\sigma)} \deg_{\text{in}}(v_j) & \text{if } C_\sigma \subseteq G(V, E) \text{ and } \text{fix}(\sigma) \neq \emptyset \\ (-1)^k & \text{if } C_\sigma \subseteq G(V, E) \text{ and } \text{fix}(\sigma) = \emptyset \\ 0 & \text{if } C_\sigma \not\subseteq G(V, E) \end{cases}$$

Further,

$$\left| \prod_{j=1}^{n-1} L_{j\sigma(j)} \right| = \left| \left\{ \text{spregs containing } C_\sigma = \bigcup_{j=1}^{k} C_j \right\} \right|. \quad (10.5)$$

The proof of this result requires reasoning much like that used in Section 10.2.1 and so is left to the reader.

### 10.3 Proof of Tutte’s theorem

Throughout this section I will continue to write $L$ in place of $\hat{L}_n$ to avoid confusing welter of subscripts.
Proof. As we argued at the beginning of Section 10.2, it is sufficient to prove that \( \det(\hat{L}_n) = \det(L) \) is the number of spanning arborescences rooted at \( v_n \). We’ll do this with the Principle of Inclusion/Exclusion and so, to begin, we need to specify the universal set \( U \) and the subsets \( X_j \). Begin by considering the set \( \mathcal{C} \) of all possible directed cycles involving the vertices \( v_1 \ldots v_{n-1} \). It’s clearly a finite set and so we can declare that it has \( M \) elements and imagine that we’ve chosen some (arbitrary) numbering scheme so that we can list the set of cycles as

\[
\mathcal{C} = \{C_1, \ldots, C_M\}.
\]

We’ll then choose the sets \( U \) and \( X_j \) as follows:

- \( U \) is the set of all spregs with distinguished vertex \( v_n \). That is, \( U \) is the set of subgraphs of \( G(V, E) \) in which
  \[
  \deg_{in}(v_n) = 0 \quad \text{and} \quad \deg_{in}(v_j) = 1 \quad \text{for } 1 \leq j \leq (n-1).
  \]

- \( X_j \subseteq U \) is the subset of \( U \) consisting of spregs containing the cycle \( C_j \). This subset may, of course, be empty.

Proposition 10.3—the one about characterising spregs—tells us that a spreg that has distinguished vertex \( v_n \) is either a spanning arborescence rooted at \( v_n \) or a graph that contains one or more disjoint cycles. This means that

\[
N_n = |\{\text{spanning arborescences rooted at } v_n\}| = |U| - \left| \bigcup_{j=1}^{M} X_j \right|.
\]

and the Principle of Inclusion/Exclusion then says

\[
N_n = |U| - \left( \sum_{I \subseteq \{1, \ldots, M\}, I \neq \emptyset} (-1)^{|I|-1} \left| \bigcap_{j \in I} X_j \right| \right) = |U| + \sum_{I \subseteq \{1, \ldots, M\}, I \neq \emptyset} (-1)^{|I|} \left| \bigcap_{j \in I} X_j \right| \quad (10.6)
\]

As we know that spregs contain only disjoint cycles, we can say

\[
|X_j \cap X_k| = 0 \quad \text{unless } C_j \cap C_k = \emptyset
\]

and so can eliminate many of the terms in the sum over intersections in Eqn. (10.6), rewriting it as a sum over collections of disjoint cycles:

\[
N_n = |U| + \sum_{I \subseteq \{1, \ldots, M\}, I \neq \emptyset, \forall j, k \in I} (-1)^{|I|} \left| \bigcap_{j \in I} X_j \right| . \quad (10.7)
\]

10.8
Then we can use the lemma from the previous section—Lemma \[10.6\] which relates non-identity permutations to numbers of spregs containing cycles—to rewrite Eqn. \((10.7)\) in terms of permutations. First note that Eqn. \((10.5)\) allows us to write

\[
\left| \bigcap_{j \in I} X_j \right| = \left| \{ \text{spregs containing } \bigcup_{j \in I} C_j \} \right| = \prod_{k=1}^{n-1} L_{k \sigma(k)}.
\]

Here \(\sigma_I \in S_{n-1}\) is the permutation

\[
\sigma_I = \prod_{j \in I} \sigma_{C_j}
\]

whose cycle representation is the product of the permutations corresponding to the directed cycles \(C_j\) for \(j \in I\). In the product above \(\sigma_{C_j}\) is the cycle permutation corresponding to the directed cycle \(C_j\). The correspondence here comes from the bijection between permutations and unions of directed cycles that we discussed in Section \[8.2\].

Now, again using Lemma \[10.6\], we have

\[
N_n = |U| + \sum_{I \subseteq \{1, \ldots, M\}, I \neq \emptyset} (-1)^{|I|} \left| \prod_{j=1}^{n-1} L_{j \sigma_I(j)} \right|
\]

As the sum in Eqn. \((10.8)\) ranges over all collections of disjoint cycles, the permutations \(\sigma_I\) range over all non-identity permutations in \(S_{n-1}\) and so we have

\[
N_n = |U| + \sum_{\sigma \neq \text{id}} \text{sgn}(\sigma) \prod_{j=1}^{n-1} L_{j \sigma(j)}.
\]

Finally, from Eqn. \((10.3)\) we know that

\[
|U| = |\{ \text{spregs containing all } v \in V \text{ with distinguished vertex } v_n \}| = \prod_{j=1}^{n-1} \text{deg}_{in}(v_j)
\]

which is the term in \(\det(L)\) corresponding to the identity permutation. Combining this observation with Eqn. \((10.9)\) gives us

\[
N_n = \sum_{\sigma \in S_{n-1}} \text{sgn}(\sigma) \prod_{j=1}^{n-1} L_{j \sigma(j)} = \det(L),
\]

which is the result we sought.