## Proof of the Cauchy-Binet Theorem and the Matrix Tree Theorem

Cauchy-Binet Theorem: Assume $p \leq q$. Let $A=\left(a_{i j}\right)$ be an $p \times q$ matrix, let $B=\left(b_{i j}\right)$ be a $q \times p$ matrix, and write $A B=C=\left(c_{i j}\right)$. Then

$$
\begin{aligned}
& \operatorname{det}(A B)=\operatorname{det}\left(C_{1}, \ldots, C_{p}\right)=\operatorname{det}\left(\sum_{i=1}^{q} b_{i 1} A_{i}, \ldots, \sum_{i=1}^{q} b_{i p} A_{i}\right)= \\
& \sum_{1 \leq i_{1}, \ldots, i_{p} \leq q} b_{i_{1} 1} \cdots b_{i_{p} p} \operatorname{det}\left(A_{i_{1}}, \ldots, A_{i_{p}}\right)= \\
& \sum_{1 \leq i_{1}<i_{2}<\cdots<i_{p} \leq q} \sum_{\sigma \in \mathcal{S}_{p}} b_{i_{\sigma(1)} 1} \cdots b_{i_{\sigma(p) p} p} \operatorname{det}\left(A_{i_{\sigma(1)}}, \ldots, A_{i_{\sigma(p)}}\right)= \\
& \sum_{1 \leq i_{1}<i_{2}<\cdots<i_{p} \leq q} \sum_{\sigma \in \mathcal{S}_{p}} b_{i_{\sigma(1)}} \cdots b_{i_{\sigma(p) p} p} \operatorname{sgn}(\sigma) \operatorname{det}\left(A_{i_{1}}, \ldots, A_{i_{p}}\right)= \\
& \sum_{I \in\binom{[q])}{p}} \operatorname{det}\left(A_{I}\right) \operatorname{det}\left(B_{I}\right)
\end{aligned}
$$

where for a subset $I$ of $[q]$ of size $p, A_{I}$ is the submatrix of $A$ using the $p$ columns from $I$ and $B_{I}$ is the submatrix of $B$ using the $p$ rows from $I$.

Counting spanning trees: Let $G$ be a graph with vertex set $V=\left\{x_{1}, \ldots, x_{n}\right\}$ and edge set $E=\left\{e_{1}, \ldots, e_{m}\right\}$ where $m \geq n-1$. Then the number of spanning trees of $G$ is

$$
\sum_{H \in\binom{E}{n-1}} \chi(H \text { is a spanning tree })
$$

Given the resemblance of this formula to the Cauchy-Binet Theorem, it should not be surprising that there is a determinant formula for this expression.

Matrix-Tree Theorem: Let

$$
C=\left((-1)^{\chi\left(x_{i}=\min e_{j}\right)} \chi\left(x_{i} \in e_{j}\right)\right)
$$

where $1 \leq i \leq n-1$ and $1 \leq j \leq m$. Then the number of spanning trees is $\operatorname{det}\left(C C^{T}\right)$.

## Example:

$$
\begin{aligned}
& G=3-4 \xrightarrow{2} \begin{array}{l}
\text { 2 } \\
C=\left[\begin{array}{cccccc}
-1 & 0 & 0 & -1 & 0 & 0 \\
1 & -1 & 0 & 0 & -1 & 0 \\
0 & 0 & -1 & 1 & 1 & 0 \\
0 & 1 & 1 & 0 & 0 & -1
\end{array}\right] \\
\operatorname{det}\left(C C^{T}\right)=8
\end{array} \\
& \hline
\end{aligned}
$$

Spanning trees:


Proof of the Matrix-Tree Theorem: We have

$$
\operatorname{det}\left(C C^{T}\right)=\sum_{I \in\left(\begin{array}{c}
{[m]} \\
n-1 \\
)
\end{array}\right.} \operatorname{det}\left(C_{I}\right) \operatorname{det}\left(C_{I}^{T}\right)=\sum_{I \in\binom{[m]}{n-1}} \operatorname{det}\left(C_{I}\right)^{2} .
$$

We will prove that

$$
\operatorname{det}\left(C_{I}\right)^{2}=\chi\left(\left\{e_{i}: i \in I\right\} \text { is a spanning tree }\right)
$$

for each $I \in\binom{[m]}{n-1}$.

Let $I \in\binom{[m]}{n-1}$ be given. Name the corresponding edges $f_{1}, \ldots, f_{n-1}$. Then the $i j$-entry of $C_{I}$ is 0 if $x_{i} \notin f_{j}$ and is $\pm 1$ if $x_{i} \in f_{j}$. These edges form a spanning tree if and only if they are connected and encompass $n$ vertices.
Case 1: $\left\{f_{1}, \ldots, f_{n-1}\right\}$ does not incorporate all $n$ vertices. If $x_{n}$ is isolated then each column of $C_{I}$ has a 1 and a -1 in it, so the sum of its columns is the 0 vector, so its columns are linearly dependent and $\operatorname{det}\left(C_{I}\right)=0$. If some other vertex $x_{k}$ is isolated then row $k$ in $C_{I}$ is the 0 vector, which again implies $\operatorname{det}\left(C_{I}\right)=0$.
Case 2: $\left\{f_{1}, \ldots, f_{n-1}\right\}$ encompasses all $n$ vertices but is not connected. Each component has at least two vertices. The sum of all the rows corresponding to vertices in a component not containing $n$ is 0 , hence the rows are not linearly independent and $\operatorname{det}\left(C_{I}\right)=0$.
Case 3: $\left\{f_{1}, \ldots, f_{n-1}\right\}$ incorporates all $n$ vertices and is connected. The collection of edges forms a spanning tree. Clipping leaf vertices and edges, we can permute the rows and columns of $C_{I}$ to produce a lower-triangular matrix with $\pm 1$ in each diagonal entry. This implies $\operatorname{det}\left(C_{I}\right)= \pm 1$.

