Lecture 11: Cauchy-Binet theorem and Kirchhoff's Theorem

May 9, 2022

In this lecture, we discuss connections between linear algebra and graph theory. Using some interesting results from linear algebra, we will be able to prove the Kirchhoff's theorem about counting the number of spanning trees of a graph.

1 Cauchy-Binet

Throughout this section, we work with real matrices. However, all the linear algebra in this lecture also holds true for the complex field.

Theorem 1. Sylvester's determinant identity
Let A be an $n \times m$ matrix and B be a $m \times n$ matrix. Then,

$$\det(I_n + AB) = \det(I_m + BA)$$

Remark 1. *On wikipedia, this is called the Weinstein–Aronszajn identity.*

If $m \neq n$, one of the two sides will be a smaller determinant to calculate. This is useful for computational linear algebra.

Proof. We will first do it for $n = k$, and when at least one of A or B is an invertible matrix. Then, we will be able to generalize it for the general case.

In this special case, the identity easily follows from

$$\det(I_n + AB) = \det\left(A(I_n + BA)A^{-1}\right),$$

assuming A is invertible, but works vice-versa for the case of B being invertible.

When neither A or B are invertible, we can approximate them via invertible matrices.

Let $t \in \mathbb{R}$ be an arbitrary parameter and consider the matrix $A_t = tI_n + A$. Then, it is clear that $\det A_t$ is a polynomial in t and can be zero at most at n different values of t. Furthermore, $A_t = A$ when $t = 0$.

When A_t is invertible, we can write

$$\det(I_n + A_t B) = \det(I_n + BA_t),$$

by what we discussed. Both sides are polynomials in t and the equality holds for all but finitely many values of $t \in \mathbb{R}$. Hence, as polynomials they are equal and therefore are equal at all values of t. Therefore, equality holds when $t = 0$ and we are done.

When $n \neq m$, we can perform a "padding with 0s" to reduce it to the case of $n = m$. Suppose that $n \leq m$. Then we consider the matrices

$$B_1 = \left[\begin{array}{c|c} B & 0 \end{array}\right],$$
$$A_1 = \begin{bmatrix} A \\ - \\ 0 \end{bmatrix},$$

where we add $m - n$ columns of 0 to B to make B_1 and we add $m - n$ rows of 0 to A to make A_1. Then, observe that

$$B_1 A_1 = BA.$$

On the other hand,

$$A_1B_1 = \begin{bmatrix} AB & | & 0 \\ - & & - \\ 0 & | & 0 \end{bmatrix}.$$

So if we write

$$\det(I_m + A_1B_1) = \det(I_m + B_1A_1),$$

we can recover our required identity. $\square$

Using this identity, we will now prove the Cauchy-Binet theorem. But first, we need some notation.

Definition 1. *Suppose M is a matrix of size $m_1 \times m_2$. Let $S_1 \subseteq [m_2]$ and let $S_2 \subseteq [m_2]$. Then, we denote the $\#S_1 \times \#S_2$ size matrix M_{S_1,S_2} to be the matrix formed by taking only those entries M_{ij} of M such that $i \in S_1$ and $j \in S_2$. Such a matrix is called the minor at S_1 and S_2.*

Theorem 2. Cauchy-Binet theorem

Let A be an $n \times m$ matrix and let B be an $m \times n$ matrix. Suppose $n \leq m$. Then,

$$\begin{aligned} \det(AB) &= \sum_{S \in \binom{[m]}{n}} \det\left(B_{S,[n]}\right) \det\left(A_{[n],S}\right) \\ &= \sum_{S \in \binom{[m]}{n}} \det\left((BA)_{S,S}\right) \end{aligned}$$

Remark 2. *Observe that both sides are calculating determinants of $n \times n$ matrices.*

Proof. First of all, observe that if $i, j \in S \subseteq [m]$, then

$$((BA)_{S,S})_{ij} = \sum_{k=1}^{n} B_{ik}A_{kj} = (B_{S,[n]}A_{[n],S})_{ij}.$$

Now, let us start with a $z \in \mathbb{R} \setminus \{0\}$ and substitute $A \mapsto \frac{1}{z}A$ in Theorem 1. So we get,

$$\begin{aligned} &\det(I_n + \frac{1}{z}AB) = \det(I_m + \frac{1}{z}BA) \\ \Rightarrow &z^{m-n}\det(zI_n + AB) = \det(zI_m + BA). \end{aligned} \tag{1}$$

This is an equality of polynomials in z. Therefore, in particular, the coefficients of z^{m-n} is the same on both sides. On the left side, the coefficient of z^{m-n} is simply $\det(AB)$. On the right hand side, we will need to carefully manipulate our expression to chase the coefficient.

Let us denote the permutation group of a set S by

$$\mathrm{Perm}(S) = \{\sigma : S \to S \mid \sigma \text{ is a bijection}\},$$

and for a $\sigma \in \mathrm{Perm}([m])$, let

$$[m]^\sigma = \{i \in [m] \mid \sigma(i) = i\}.$$

Then, from the definition of the determinant, we have that the right-hand side is

$$\begin{aligned}\det(zI_m + BA) &= \sum_{\sigma\in\mathrm{Perm}([m])} \mathrm{sgn}(\sigma) \prod_{i\in[m]} \left(z\delta_{i\sigma(i)} + (BA)_{i\sigma(i)}\right)\\ &= \sum_{\sigma\in\mathrm{Perm}([m])} \mathrm{sgn}(\sigma) \prod_{i\in[m]\setminus[m]^\sigma} (BA)_{i\sigma(i)} \prod_{i\in[m]^\sigma} \left(z + (BA)_{i\sigma(i)}\right)\\ &= \sum_{\sigma\in\mathrm{Perm}([m])} \mathrm{sgn}(\sigma) \prod_{i\in[m]\setminus[m]^\sigma} (BA)_{i\sigma(i)} \sum_{S_1\subseteq[m]^\sigma} z^{\#S_1} \prod_{i\in[m]^\sigma\setminus S_1} (BA)_{i\sigma(i)}\\ &= \sum_{S_1\subseteq[m]} z^{\#S_1} \sum_{\substack{\sigma\in\mathrm{Perm}([m])\\ S_1\subseteq[m]^\sigma}} \mathrm{sgn}(\sigma) \prod_{i\in[m]\setminus[m]^\sigma} (BA)_{i\sigma(i)} \prod_{i\in[m]^\sigma\setminus S_1} (BA)_{i\sigma(i)}\\ &= \sum_{S_1\subseteq[m]} z^{\#S_1} \sum_{\substack{\sigma\in\mathrm{Perm}([m])\\ S_1\subseteq[m]^\sigma}} \mathrm{sgn}(\sigma) \prod_{i\in[m]\setminus S_1} (BA)_{i\sigma(i)}\\ &= \sum_{S\subseteq[m]} z^{m-\#S} \sum_{\substack{\sigma\in\mathrm{Perm}([m])\\ [m]\setminus S\subseteq[m]^\sigma}} \mathrm{sgn}(\sigma) \prod_{i\in S}(BA)_{i\sigma(i)}\end{aligned}$$

Now we will need the following lemma.

Lemma 1. *Suppose $S \subseteq [m]$. Then,*

1. *For $\sigma \in \mathrm{Perm}\,([m])$ such that $[m] \setminus S \subseteq [m]^\sigma$, we have that $\sigma|_S \in \mathrm{Perm}\,(S)$. Here $\sigma_S : S \to [m]$ is the restriction of σ to S.*
2. *For σ as above, we get $\mathrm{sgn}(\sigma|_S) = \mathrm{sgn}(\sigma)$.*
3. *The mapping $\sigma \mapsto \sigma|_S$ is a bijection between*

$$\{\sigma \in \mathrm{Perm}([m]) \mid [m] \setminus S \subseteq [m]^\sigma\} \leftrightarrow \mathrm{Perm}(S).$$

The proof of this lemma is an exercise. Using this lemma, we conclude that the coefficient of z^{m-n} in the right-hand side of Equation 1 is exactly what we need.

□

Example 1. *Let $n = 1$, $B = A^T$ and*

$$A = \begin{bmatrix} a_1 & a_2 & a_3 & \dots & a_m \end{bmatrix}.$$

Then, both sides of the Cauchy-Binet theorem give

$$\sum_{i=1}^{m} a_i^2.$$

This means that the Cauchy-Binet theorem is a vast generalization of the Pythagoras theorem.

2 Kirchhoff's theorem

Let G be a graph with n vertices and m edges. We recall the notions of $L(G), D(G), A(G)$ and $M(G, \mathcal{O})$.

Observe that if

$$v = \begin{bmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{bmatrix},$$

then $L(G)v = 0$. Hence, $\mathrm{rank}(L(G)) \leq n - 1$.

Theorem 3. Kirchhoff
Let G be a connected graph with n vertices. Then $\operatorname{rank}(L(G)) = n-1$. *Furthermore, if* $\lambda_1, \lambda_2, \ldots, \lambda_{n-1}$ *are the non-zero roots of* $L(G)$, *then*

$$\frac{1}{n}\lambda_1\lambda_2 \ldots \lambda_n = \# \textit{ of spanning trees of } G.$$

Proof. Fix an orientation $\mathcal{O}$ on G. Suppose that $V(G) = [n]$.

Let $L_0(G) = L(G)_{[n-1],[n-1]}$ and $M_0(G, \mathcal{O}) = M(G, \mathcal{O})_{[n-1],E(G)}$. Then, check that (see previous lecture)

$$L_0(G) = M_0(G, \mathcal{O})M_0(G, \mathcal{O})^T.$$

Also, check that (this is an exercise)

$$\det(L_0(G)) = \frac{1}{n}\lambda_1\lambda_2\lambda_3 \ldots \lambda_{n-1}.$$

Then, from the Cauchy-Binet theorem, we get that

$$\det(L_0(G)) = \sum_{S \in \binom{E(G)}{n-1}} \det\left(M_0(G, \mathcal{O})_{[n-1],S}\right)^2.$$

Hence, on the right-hand side, we will have a sum of all possible subgraphs of size $n-1$ of G. To finish this proof, will now show that the

$$\det\left(M_0(G, \mathcal{O})_{[n-1],S}\right)^2 = \begin{cases} 1 & \text{if } S \text{ are the edges forming a spanning tree} \\ 0 & \text{otherwise} \end{cases}$$

Let us try to observe this for the case of the oriented graph $G = ([4], E)$, where the oriented edges are

$$E(G) = \{(1,2), (1,3), (2,3), (2,4), (3,4)\}.$$

Then,

$$M(G, \mathcal{O}) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 & 1 \\ 0 & 0 & 0 & -1 & -1 \end{bmatrix},$$

so

$$M_0(G, \mathcal{O}) = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 \\ -1 & 0 & 1 & 1 & 0 \\ 0 & -1 & -1 & 0 & 1 \end{bmatrix}.$$

We know that the edge set $\{(1,2), (1,3), (2,3)\}$ don't make a spanning tree. This is reflected in that there is a non-zero vector in the kernel of the matrix.

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

Indeed, this vector in the kernel can be constructed in general. If S don't form a spanning tree in the vertices, then it is possible to find a cycle among edges $R \subseteq S$. Choose a cyclic orientation on R (clockwise, or counter clockwise) and then define the vector $v = (v_e)_{e \in S} \in \mathbb{R}^S$ as

$$v_e = \begin{cases} 1 & e \in R \text{ and orientation } \mathcal{O} \text{ agrees with orientation on } R \\ -1 & e \in R \text{ and orientation } \mathcal{O} \text{ disagrees with orientation on } R \\ 0 & e \notin R \end{cases}.$$

With this vector, we see that for any vertex $i \in [n-1]$, the sum

$$\begin{aligned}(M_0(G,\mathcal{O})_{[n-1],S} \cdot v)_i &= \sum_{e \in S} M_0(G,\mathcal{O})_{ie} v_e \\ &= \begin{cases} +1-1 & \text{if } i \text{ is in the cycle formed by } R \\ 0 & \text{if } i \text{ is not in the cycle formed by } R \end{cases}.\end{aligned}$$

This means that we can always find a non-trivial vector in the kernel of $M_0(G,\mathcal{O})_{[n-1],S}$ as long as S contains a cycle and therefore the determinant of this matrix is zero.

On the other hand, we know that $S = \{(1,2),(2,3),(2,4)\}$ make a spanning tree. The matrix corresponding to it has an inverse matrix completely made of integer entries.

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

It is true in general that if a matrix has integer entries and has an inverse that also has integer entries, then the matrix has a determinant of ± 1 (indeed, take the determinant of this equation). Hence, all we need to show is that $M_0(G,\mathcal{O})_{[n-1],S}$ has an inverse made of integer entries whenever S are the edge-set of a spanning tree.

This can be achieved using the following construction. Let $T = ([n], S)$ denote the tree formed by the edges in S. We know that in a connected tree, every pair of distinct vertices are connected by a unique path. Imagine now, that current is flowing down in the tree along paths down to the vertex n, corresponding to the deleted row, from every other vertex in the tree T. Construct a matrix N, whose rows are indexed by S and columns are indexed by $[n-1]$ and is given by the following

$$N_{ej} = \begin{cases} 0 & \text{if the unique path from } j \text{ to } n \text{ in } T \text{ does not pass through } e \\ 1 & \text{if the unique path from } j \text{ to } n \text{ in } T \text{ passes through } e \text{ along the orientation } \mathcal{O} \\ -1 & \text{if the unique path from } j \text{ to } n \text{ in } T \text{ passes through } e \text{ against the orientation } \mathcal{O} \end{cases}$$

Now, for each pair of vertices $i, j \in [n-1]$, we have

$$(M_0(G,\mathcal{O})_{[n-1],S} N)_{ij} = \sum_{\substack{e \in S \\ i \text{ is a source of } e}} N_{ej} - \sum_{\substack{e \in S \\ i \text{ is a target of } e}} N_{ej}.$$

The right hand side is now going to be ± 1 if and only if $i = j$. By suitably multiplying the columns of N by ± 1, an integral matrix inverse of of $M_0(G,\mathcal{O})_{[n-1],S}$ can be obtained. Hence, we prove the claim that $\det(M_0(G,\mathcal{O})_{[n-1],S})^2 = 1$ and we are done with the proof. □