The Cauchy–Binet formula

ho boon suan

January 2021

The Cauchy–Binet formula is a generalization of the identity $\det(AB) = \det(A)\det(B)$ to non-square matrices. More specifically, if A and B are $m \times n$ and $n \times m$ matrices respectively, then

$$\det(AB) = \sum_{S \subseteq \binom{[n]}{m}} \det(A_{[m]\times S}) \det(B_{S\times[m]}),$$

where $[n] := \{1, 2, \ldots, n\}$, $\binom{[n]}{m}$ denotes the set of m element subsets of $[n]$, and $A_{R\times S} := (a_{ij})_{i\in R, j\in S}$ is the submatrix of A with rows indexed by R and columns indexed by S. (In particular, $A = A_{[m]\times[n]}$.) For example, if $m = 2$ and $n = 3$, writing $|A| := \det(A)$ for convenience, we get the identity

$$\begin{aligned} &\det\left[\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix}\begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix}\right] \\ &= \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix}\begin{vmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{vmatrix} + \begin{vmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{vmatrix}\begin{vmatrix} b_{11} & b_{12} \\ b_{31} & b_{32} \end{vmatrix} + \begin{vmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{vmatrix}\begin{vmatrix} b_{21} & b_{22} \\ b_{31} & b_{32} \end{vmatrix}. \end{aligned}$$

If $m = n$, the formula is precisely $\det(AB) = \det(A)\det(B)$. If $m > n$, then $\binom{[n]}{m} = \varnothing$ and so $\det(AB) = 0$, reflecting the fact the $m \times m$ matrix AB cannot have full rank as $\operatorname{rank}(AB) \le \operatorname{rank}(A) \le n < m$.

We present two proofs of the formula. The first proof relies on the exterior algebra, and the second proof makes use of characteristic polynomials.

Via exterior powers

Let $m < n$. The $m \times n$ matrix A can be interpreted as a linear map $L_A \colon \mathbf{k}^n \to \mathbf{k}^m$, where $\mathbf{k}$ is a field. We shall investigate what maps the $m \times m$ matrices $A_{[m]\times S}$ and $B_{S\times[m]}$ represent. Denote by $e_1, \ldots, e_n$ the standard basis for $\mathbf{k}^n$ and fix $S = \{s_1, \ldots, s_m\}$ with $1 \le s_1 < \cdots < s_m \le n$. We define an m-dimensional subspace of $\mathbf{k}^n$ by

$$V_S := \operatorname{span}\{e_{s_1}, \ldots, e_{s_m}\} \subseteq \mathbf{k}^n.$$

A natural way to obtain a map between m-dimensional spaces from L_A is by first applying some inclusion $\mathbf{k}^m \hookrightarrow \mathbf{k}^n$ before applying L_A. Similarly, since L_B is a map $\mathbf{k}^m \to \mathbf{k}^n$, it is natural to apply a projection $\mathbf{k}^n \twoheadrightarrow \mathbf{k}^m$ after applying L_B to obtain a map between m-dimensional spaces. We are thus led to consider the maps

$$V_S \overset{\iota_S}{\hookrightarrow} \mathbf{k}^n \overset{L_A}{\to} \mathbf{k}^m$$

where ι_S denotes the natural inclusion, and

$$\mathbf{k}^m \overset{L_B}{\to} \mathbf{k}^n \overset{\pi_S}{\twoheadrightarrow} V_S$$

where π_S denotes the natural projection onto V_S. Identifying $\mathbf{k}^m \cong V_S$ by $e_i \mapsto e_{s_i}$, we find that $L_A \circ \iota_S$ and $\pi_S \circ L_B$ are represented by $A_{[m]\times S}$ and $B_{S\times[m]}$ respectively. (This fact is perhaps best appreciated with a concrete example as given in the margin, noting that multiplying a matrix on the right gives linear combinations of columns while multiplying on the left gives linear combinations of rows.) Passing to the m-th exterior power for L_B, we get

$$(\Lambda^m(\pi_S L_B))(e_1 \wedge \cdots \wedge e_m) = \det(B_{S\times[m]}) e_{s_1} \wedge \cdots \wedge e_{s_m}.$$

Since $\Lambda^m(\pi_S L_B) = \Lambda^m \pi_S \circ \Lambda^m L_B$, it follows that

$$(\Lambda^m L_B)(e_1 \wedge \cdots \wedge e_m) = \sum_{\substack{S=\{s_1,\dots,s_m\} \\ 1\le s_1<\cdots<s_m\le n}} \det(B_{S\times[m]}) e_{s_1} \wedge \cdots \wedge e_{s_m}.$$

Since the m-th exterior power for L_A gives

$$(\Lambda^m L_A)(e_{s_1} \wedge \cdots \wedge e_{s_m}) = \det(A_{[m]\times S}) e_1 \wedge \cdots \wedge e_m$$

where we have once again identified $\mathbf{k}^m \cong V_S$ as above, we compute

$$\begin{aligned}
&(\Lambda^m L_{AB})(e_1 \wedge \cdots \wedge e_m) \\
&\quad = (\Lambda^m L_A) \sum_{\substack{S=\{s_1,\dots,s_m\} \\ 1\le s_1<\cdots<s_m\le n}} \det(B_{S\times[m]}) e_{s_1} \wedge \cdots \wedge e_{s_m} \\
&\quad = \sum_{\substack{S=\{s_1,\dots,s_m\} \\ 1\le s_1<\cdots<s_m\le n}} \det(B_{S\times[m]}) (\Lambda^m L_A)(e_{s_1} \wedge \cdots \wedge e_{s_m}) \\
&\quad = \left(\sum_{S\subseteq\binom{[n]}{m}} \det(A_{[m]\times S}) \det(B_{S\times[m]}) \right) e_1 \wedge \cdots \wedge e_m.
\end{aligned}$$

An example with $m = 2$, $n = 3$, and $S = \{1,3\} \subseteq [3]$. We have naturally identified $\mathbf{k}^m \cong V_S$ by $e_i \mapsto e_{s_i}$.

$L_A \circ \iota_S = L_{A_{[2]\times S}}$:

$$\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} a_{11} & a_{13} \\ a_{21} & a_{23} \end{pmatrix}$$

$\pi_S \circ L_B = L_{B_{S\times[2]}}$:

$$\begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \\ b_{31} & b_{32} \end{pmatrix} = \begin{pmatrix} b_{11} & b_{12} \\ b_{31} & b_{32} \end{pmatrix}$$

Via the characteristic polynomial

Given an $n \times n$ matrix X, we work with the polynomial $\det(zI_n + X)$ in z whose coefficients are those of the characteristic polynomial, without the signs for convenience. We first show that the coefficient of z^{n-m} in this polynomial is equal to the sum of $m \times m$ principal minors of X, where $1 \le m \le n$. We compute

$$\begin{aligned}
\det(zI_n + X) &= \sum_{\sigma\in\mathfrak{S}_n} (\operatorname{sgn}\sigma) \prod_{1\le m\le n} (z\delta_{m,\sigma(m)} + X_{m,\sigma(m)}) \\
&= \sum_{\sigma\in\mathfrak{S}_n} (\operatorname{sgn}\sigma) \sum_{S\subseteq[n]} \prod_{i\in S} X_{i,\sigma(i)} \prod_{j\in[n]-S} z\delta_{j,\sigma(j)} \\
&= \sum_{S\subseteq[n]} \sum_{\sigma\in\mathfrak{S}_n} (\operatorname{sgn}\sigma) \prod_{i\in S} X_{i,\sigma(i)} \prod_{j\in[n]-S} z\delta_{j,\sigma(j)} \\
&= \sum_{S\subseteq[n]} z^{n-|S|} \sum_{\sigma\in\mathfrak{S}_S} (\operatorname{sgn}\sigma) \prod_{i\in S} X_{i,\sigma(i)} \\
&= \sum_{S\subseteq[n]} z^{n-|S|} \det(X_{S\times S}) \\
&= \sum_{0\le m\le n} z^{n-m} \sum_{S\in\binom{[n]}{m}} \det(X_{S\times S}).
\end{aligned}$$

The *Kronecker delta* $\delta_{i,j}$ is equal to 1 if $i = j$ and is 0 otherwise.

Here the sign stays the same when we pass from $\mathfrak{S}_n$ to the subgroup $\mathfrak{S}_S$. This can be seen by thinking of $\operatorname{sgn}\sigma$ as counting the number of transpositions of σ, modulo 2.

Before proving the Cauchy–Binet formula, we will need the identity

$$\det(zI_n + BA) = z^{n-m}\det(zI_m + AB),$$

where $m \leq n$, and A and B are $m \times n$ and $n \times m$ matrices respectively. We first show the result for when $z = 1$ and $m = n$. In this case, the identity reads $\det(I_m + BA) = \det(I_m + AB)$. We may consider the identity $\det((I_m + BA)B) = \det(B(I_m + AB))$ as a polynomial identity in the domain $\mathbf{Z}[a_{ij}, b_{ij}]$, where we may cancel $\det B$ from both sides to obtain the result. We may then apply the result over any field via the universal property of polynomial rings, sending each indeterminate a_{ij} to the field element $a_{ij} \in \mathbf{k}$. We may then extend the result to when $m < n$ by padding the rectangular matrices with zeroes to form square matrices. In detail, we get

If one prefers to avoid such abstract nonsense proofs, one may simply note that $B(I_m + AB)B^{-1} = I_m + BA$, and thus the result holds for invertible B, which are dense in the space of $m \times m$ matrices.

$$\begin{pmatrix} B & 0_{n\times(n-m)} \end{pmatrix} \begin{pmatrix} A \\ 0_{(n-m)\times n} \end{pmatrix} = BA$$

and

$$\begin{pmatrix} A \\ 0_{(n-m)\times n} \end{pmatrix} \begin{pmatrix} B & 0_{n\times(n-m)} \end{pmatrix} = \begin{pmatrix} AB & 0_{m\times(n-m)} \\ 0_{(n-m)\times m} & 0_{(n-m)\times(n-m)} \end{pmatrix},$$

and the result follows since

$$\det\begin{pmatrix} I_m + AB & 0 \\ 0 & I_{n-m} \end{pmatrix} = \det(I_m + AB)\det(I_{n-m}),$$

which can be seen by using the Leibniz permutation expansion of the determinant. Finally, for $z \neq 1$, we employ a scaling argument. The case for $z = 0$ is left as an exercise; consider $z \neq 0$. We set $A' := z^{-1}A$, and compute

$$\begin{aligned} \det(zI_n + BA) &= \det(zI_n + zBA') \\ &= z^n \det(I_n + BA') \\ &= z^n \det(I_m + A'B) \\ &= z^{n-m}\det(zI_m + zA'B) \\ &= z^{n-m}\det(zI_m + AB). \end{aligned}$$

The Cauchy–Binet formula is now within our reach. Comparing the coefficients of z^{n-m} in $\det(zI_n + BA) = z^{n-m}\det(zI_m + AB)$, we find that the sum of principal $m \times m$ minors of BA is equal to $\det(AB)$; that is,

$$\det(AB) = \sum_{S \in \binom{[n]}{m}} \det((BA)_{S\times S}).$$

If $S = \{s_1, \ldots, s_m\}$, then

$$\begin{aligned} ((BA)_{S\times S})_{ij} &= (BA)_{s_i,s_j} \\ &= \sum_{1\leq k\leq m} B_{s_i,k}A_{k,s_j} \\ &= \sum_{1\leq k\leq m} (B_{S\times[m]})_{i,k}(A_{[m]\times S})_{k,j} \\ &= (B_{S\times[m]}A_{[m]\times S})_{ij}, \end{aligned}$$

proving the result.