Vv417 Lecture 5

Jing Liu

UM-SJTU Joint Institute

September 19, 2019

- With each $n \times n$ matrix $\mathbf{A}$, it is possible to associate a scalar, denoted

  $\det(\mathbf{A})$, whose value will tell us whether the matrix is invertible.

- Before proceeding to the general definition, let us consider the following,
- 1×1 Matrices: $\mathbf{A} = \begin{bmatrix} a_{11} \end{bmatrix}$
- We know $\mathbf{A}$ will have a multiplicative inverse if and only if $a_{11} \neq 0$.
- Thus, suppose we define

$$\det(\mathbf{A}) = a_{11}$$

  then $\mathbf{A}$ will be invertible if and only if $\det(\mathbf{A}) \neq 0$.

- 2×2 Matrices: $\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$
- We know $\mathbf{A}$ will be invertible if and only if it is row equivalent to $\mathbf{I}$.

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \sim \begin{bmatrix} a_{11} & a_{12} \\ 0 & (a_{22} - a_{12}a_{21}/a_{11}) \end{bmatrix}$$

- If $a_{11} \neq 0$, the resulting matrix will be row equivalent to $\mathbf{I}$ if and only if

$$\boxed{a_{22} - a_{12}a_{21}/a_{11}} \neq 0 \implies \boxed{a_{11}a_{22} - a_{12}a_{21}} \neq 0$$

- If $a_{11} = 0$, we can switch the two rows of $\mathbf{A}$. The resulting matrix

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} \sim \begin{bmatrix} a_{21} & a_{22} \\ 0 & a_{12} \end{bmatrix}$$

will be row equivalent to $\mathbf{I}$ if and only if $a_{21}a_{12} \neq 0$. This is equivalent to

$$\boxed{a_{11}a_{22} - a_{12}a_{21}} \neq 0$$

- Thus, if $\mathbf{A}$ is any 2×2 matrix and we define

$$\det(\mathbf{A}) = \boxed{a_{11}a_{22} - a_{12}a_{21}}$$

then $\mathbf{A}$ is invertible if and only if $\det(\mathbf{A}) \neq 0$.

3×3 Matrices

- Similarly we can use row operations to see if the matrix is row equivalent to $\mathbf{I}$

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \sim \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & \dfrac{a_{11}a_{22} - a_{21}a_{12}}{a_{11}} & \dfrac{a_{11}a_{23} - a_{21}a_{13}}{a_{11}} \\ 0 & \dfrac{a_{11}a_{32} - a_{31}a_{12}}{a_{11}} & \dfrac{a_{11}a_{33} - a_{31}a_{13}}{a_{11}} \end{bmatrix}$$

$$= \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ 0 & & \\ 0 & \mathbf{M}_{11} & \end{bmatrix}$$

Q: Why is $\mathbf{A}$ row equivalent to $\mathbf{I}$ if and only if

$$a_{11} \det(\mathbf{M}_{11}) \neq 0$$

- Again, if we define this scalar quantity to be

$$\det(\mathbf{A}) = \boxed{a_{11} \det(\mathbf{M}_{11})}$$

then $\mathbf{A}$ is invertible if and only if $\det(\mathbf{A}) \neq 0$.

Q: How shall we extend the definition of $\det(\mathbf{A})$ for arbitrary $\mathbf{A}$ of size $n \times n$?

- Recall we have

$$\det(\mathbf{A}) = \begin{cases} a_{11} & \text{if } \mathbf{A} \text{ is } 1 \times 1, \\ a_{11}a_{22} - a_{12}a_{21} & \text{if } \mathbf{A} \text{ is } 2 \times 2, \end{cases}$$

- Although the algebra is somewhat messy, the definition of $\det(\mathbf{A})$ for 3×3,

$$\det(\mathbf{A}) = a_{11} \det(\mathbf{M}_{11})$$

can be put in terms of elements of $\mathbf{A}$ explicitly as

$$a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}$$

Q: Can you see the pattern in the indices of a_{ij}?

Q: Can you see how the signs, $\pm$, are related to the indices of a_{ij}?

Definition

Any arrangement of a set $\mathcal{S} = \{1, 2, \ldots, n\}$ in a specific order, for example,

$$\sigma_{\text{no}} = (1, 2, \ldots, n) \qquad \text{or} \qquad \sigma = (k_1, k_2, \ldots, k_i, \ldots, k_j, \ldots, k_n)$$

is called a permutation of $\mathcal{S}$, where σ_{no} above is defined to be in the nature order.

A pair of elements (k_i, k_j) in σ is said to be out of the nature order if

$$k_i > k_j \qquad \text{where} \quad i < j$$

A permutation σ is said to be

- even if there is an even number of pairs of (k_i, k_j) out of the nature order
- odd if there is an odd number of pairs of (k_i, k_j) out of the nature order

The Levi-Civita symbol, $\varepsilon_\sigma = \varepsilon_{k_1 \cdots k_n}$, is defined by

$$\varepsilon_\sigma = \begin{cases} +1 & \text{if } \sigma \text{ is an even permutation,} \\ -1 & \text{if } \sigma \text{ is an odd permutation,} \\ 0 & \text{otherwise.} \end{cases}$$

Definition

The determinant of an $n \times n$ matrix $\mathbf{A} = [a_{ij}]$, denoted

$$\det(\mathbf{A}),$$

is a scalar associated with the matrix $\mathbf{A}$ that is defined as follows:

$$\det(\mathbf{A}) = \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_n=1}^{n} \varepsilon_{k_1 \cdots k_n} a_{1k_1} a_{2k_2} a_{3k_3} \cdots a_{nk_n}$$

- Notice the determinant for an $n \times n$ matrix is an n-fold summation.
- However, only $n!$ numbers of ε_σ out of n^n possible ε_σ are nonzero.
- So the determinant is also a summation over the $n!$ distinct permutations of

$$\{1, 2, \ldots, n\}$$

Q: Given $\mathbf{A}$ is triangular, why is $\det(\mathbf{A})$ equal to the product of the diagonals?

Interchanging two rows of $\mathbf{A}$

Suppose $\mathbf{A}$ is an $n \times n$ matrix and

$$\mathbf{E}_{i,j} = \begin{bmatrix} 1 & & & & & & \\ & \ddots & & & & & \\ & & 0 & & 1 & & \\ & & & \ddots & & & \\ & & 1 & & 0 & & \\ & & & & & \ddots & \\ & & & & & & 1 \end{bmatrix}$$

is the elementary matrix corresponding to interchanging row i with row j, then

$$\det(\mathbf{E}_{i,j}\mathbf{A}) = -\det(\mathbf{A})$$

Furthermore,

$$\det(\mathbf{E}_{i,j}\mathbf{A}) = \det(\mathbf{E}_{i,j})\det(\mathbf{A})$$

Proof

- Let $\mathbf{A}^* = \mathbf{E}_{i,j}\mathbf{A}$, then the elements $a^*_{rs} = [\mathbf{A}^*]_{rs}$ are related to those of $\mathbf{A}$

$$a^*_{rs} = \begin{cases} a_{rs} & \text{if } r \neq i, j, \\ a_{js} & \text{if } r = i, \\ a_{is} & \text{if } r = j. \end{cases}$$

- According to the definition of the determinant, we have

$$\begin{aligned} \det(\mathbf{A}^*) &= \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_n=1}^{n} \varepsilon_{k_1 \cdots k_i \cdots k_j \cdots k_n} a^*_{1k_1} a^*_{2k_2} \cdots a^*_{ik_i} \cdots a^*_{jk_j} \cdots a^*_{nk_n} \\ &= \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_n=1}^{n} \varepsilon_{k_1 \cdots k_i \cdots k_j \cdots k_n} a_{1k_1} a_{2k_2} \cdots a_{jk_i} \cdots a_{ik_j} \cdots a_{nk_n} \end{aligned}$$

- If $i - j = 1$, that is, row i and row j are adjacent, then it is clear that

$$\varepsilon_{k_1 \cdots k_j k_i \cdots k_n} = -\varepsilon_{k_1 \cdots k_i k_j \cdots k_n}$$

Proof

- If $i - j > 1$,

$$(k_1, k_2, \ldots, k_j, \ldots, \underbrace{k_i}_{\longleftarrow}, \ldots, k_n)$$

  interchanging row i to row j can be done by first successively interchanging adjacent rows to row i in the direction of row j to obtain

$$(k_1, k_2, \ldots, k_i, k_j, \ldots, k_{i-1}, k_{i+1}, \ldots, k_n)$$

  which requires $i - j$ adjacent row-switching.

- Next the old row j,

$$(k_1, k_2, \ldots, k_i, \underbrace{k_j}_{\longrightarrow}, \ldots, k_{i-1}, k_{i+1}, \ldots, k_n)$$

  which is now in row $j + 1$, is pushed into row i to complete the interchanging

- Notice this requires additional $(i - j) - 1$ adjacent row-switching.

Proof

- The total number of adjacent interchanges involved in the interchanging is

$$N = 2(i-j)-1$$

which is always an odd integer, thus

$$\varepsilon_{k_1\cdots k_j\cdots k_i\cdots k_n} = (-1)^N \varepsilon_{k_1\cdots k_i\cdots k_j\cdots k_n} = -\varepsilon_{k_1\cdots k_i\cdots k_j\cdots k_n}$$

- Hence

$$\begin{aligned}\det(\mathbf{A}^*) &= \sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\cdots\sum_{k_n=1}^{n}\varepsilon_{k_1\cdots k_i\cdots k_j\cdots k_n} a^*_{1k_1}a^*_{2k_2}\cdots a^*_{ik_i}\cdots a^*_{jk_j}\cdots a^*_{nk_n}\\ &= \sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\cdots\sum_{k_n=1}^{n}\varepsilon_{k_1\cdots k_i\cdots k_j\cdots k_n} a_{1k_1}a_{2k_2}\cdots a_{jk_i}\cdots a_{ik_j}\cdots a_{nk_n}\\ &= -\sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\cdots\sum_{k_n=1}^{n}\varepsilon_{k_1\cdots k_j\cdots k_i\cdots k_n} a_{1k_1}a_{2k_2}\cdots a_{nk_n} = -\det(\mathbf{A})\end{aligned}$$

Proof

- For the second part of the theorem, since identity matrices are triangular,

$$\det(\mathbf{I}) = 1$$

together with the fact that $\mathbf{E}_{i,j}$ is one interchange of rows away from $\mathbf{I}$

$$\det(\mathbf{E}_{i,j}) = -1$$

- Therefore,

$$\begin{aligned}\det(\mathbf{E}_{i,j}\mathbf{A}) = \det(\mathbf{A}^*) &= -\det(\mathbf{A}) \\ &= \det(\mathbf{E}_{i,j})\det(\mathbf{A})\end{aligned}$$

Q: What is the relationship between

$$\det(\mathbf{A}) \qquad \text{and} \qquad \det(\mathbf{A}\mathbf{E}_{i,j})$$

Theorem

Suppose $\mathbf{A}$ is an $n \times n$ matrix, then

$$\det\left(\mathbf{A}^{\mathrm{T}}\right) = \det\left(\mathbf{A}\right)$$

- What does the above theorem imply when we interchange two columns?

Proof

- Using the definition, we have

$$\det\left(\mathbf{A}^{\mathrm{T}}\right) = \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_n=1}^{n} \varepsilon_{k_1 \cdots k_n} a_{k_1 1} a_{k_2 2} a_{k_3 3} \cdots a_{k_n n}$$

- Since $\sigma = (k_1, k_2, k_3, \ldots, k_n)$ is a permutation of

$$\{1, 2, 3, \ldots, n\}$$

we can interchange terms in σ so that σ becomes σ_{no}.

Proof

- In terms of elements of $\mathbf{A}$, it is always possible to rearrange the terms so that

$$a_{k_1 1} a_{k_2 2} a_{k_3 3} \cdots a_{k_n n} = a_{1 k_1^*} a_{2 k_2^*} a_{3 k_3^*} \cdots a_{n k_n^*}$$

  where $\sigma^* = (k_1^*, k_2^*, k_3^*, \ldots, k_n^*)$ is some other permutation of $\{1, 2, 3, \ldots, n\}$
- Note the number of interchanges, N, needed in taking $\sigma_{\text{no}} = (1, 2, 3, \ldots, n)$ to σ^* is the same as the number of interchanges in taking σ to σ_{no}, so

$$\varepsilon_\sigma = (-1)^N \varepsilon_{\sigma_{\text{no}}} = \varepsilon_{\sigma^*}$$

- Therefore,

$$\begin{aligned} \det\left(\mathbf{A}^{\mathrm{T}}\right) &= \sum_{k_1=1}^{n} \sum_{k_2=1}^{n} \cdots \sum_{k_n=1}^{n} \varepsilon_{k_1 \cdots k_n} a_{k_1 1} a_{k_2 2} a_{k_3 3} \cdots a_{k_n n} \\ &= \sum_{k_1^*=1}^{n} \sum_{k_2^*=1}^{n} \cdots \sum_{k_n^*=1}^{n} \varepsilon_{k_1^* \cdots k_n^*} a_{1 k_1^*} a_{2 k_2^*} a_{3 k_3^*} \cdots a_{n k_n^*} = \det\left(\mathbf{A}\right) \end{aligned}$$

- As you know, we don't usually used the definition directly to find $\mathbf{A}$.

Definition

Let $\mathbf{A}$ be an $n \times n$ matrix. The cofactor of the element a_{ij} is defined by

$$C_{ij} = (-1)^{i+j} M_{ij}$$

where M_{ij} is the minor of the element a_{ij}, which is the determinant of the matrix obtained by deleting the ith row and the jth column of $\mathbf{A}$.

Cofactor Expansion

Suppose $\mathbf{A}$ is an $n \times n$ matrix, where $n \geq 2$, and C_{ij} denotes the cofactors, then

$$\det(\mathbf{A}) = \begin{cases} a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in} & \text{for} \quad i = 1, \ldots n \\ a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj} & \text{for} \quad j = 1, \ldots n \end{cases}$$

- It states $\det(\mathbf{A})$ is a weighted sum of cofactors with a_{ij} as the coefficients.

Exercise

Find $\det(\mathbf{A})$, where $\mathbf{A} = \begin{bmatrix} 0 & 2 & 3 & 0 \\ 0 & 4 & 5 & 0 \\ 0 & 1 & 0 & 3 \\ 2 & 0 & 1 & 3 \end{bmatrix}$.

Solution

- Using the cofactor expansion on the row/column that has the most zeros,

$$\begin{aligned}\det(\mathbf{A}) &= (-1)^{4+1} \cdot 2 \det\left(\begin{bmatrix} 2 & 3 & 0 \\ 4 & 5 & 0 \\ 1 & 0 & 3 \end{bmatrix}\right) = -2 \cdot (-1)^{3+3} \cdot 3 \det\left(\begin{bmatrix} 2 & 3 \\ 4 & 5 \end{bmatrix}\right) \\ &= -6 \cdot (10 - 12) = 12\end{aligned}$$

- To most of you, Laplace's cofactor expansion is not new, it is often used as the definition in an elementary course, but now we are ready to prove it.

Proof

- Recall the definition

$$\det(\mathbf{A}) = \sum_{k_1=1}^{n}\sum_{k_2=1}^{n}\cdots\sum_{k_n=1}^{n}\varepsilon_{k_1\cdots k_n}a_{1k_1}a_{2k_2}a_{3k_3}\cdots a_{nk_n}$$

$$= \underbrace{a_{11}\sum_{k_2=1}^{n}\cdots\sum_{k_n=1}^{n}\varepsilon_{1k_2\cdots k_n}a_{2k_2}a_{3k_3}\cdots a_{nk_n}}_{\alpha}$$

$$+\sum_{k_1=2}^{n}\sum_{k_2=1}^{n}\cdots\sum_{k_n=1}^{n}\varepsilon_{k_1\cdots k_n}a_{1k_1}a_{2k_2}a_{3k_3}\cdots a_{nk_n}$$

- Notice if any one of those k_i in α is equal to 1,

$$\sigma = (1, k_2, \ldots, k_i, \ldots, k_n)$$

then σ is no longer a permutation of $\{1, 2, \ldots, n\}$, in which case

$$\varepsilon_\sigma = \varepsilon_{1k_2\cdots k_n} = 0$$

Proof

- If $\sigma = (1, k_2, \ldots, k_i, \ldots, k_j, \ldots, k_n)$ is a permutation of $\{1, 2, \ldots, n\}$, then

$$1 < k_i \qquad \text{for all} \qquad i = 2, \ldots, n$$

thus there are the same number of paris of

$$k_i > k_j \qquad \text{where} \qquad i < j$$

in σ as there are in

$$(k_2, k_3, \ldots, k_i, \ldots, k_j, \ldots k_n)$$

which is a permutation of $\{2, 3, \ldots, n\}$.

- Furthermore, note shifting indices according to the following definition

$$k_j^* = k_{j+1} - 1 \qquad \text{for} \quad j = 1, \ldots, (n-1)$$

has no effect on the relative ordering of the permutation.

Proof

- Since the relative ordering is preserved, the number of paris of

$$k_i > k_j \qquad \text{where} \qquad i < j$$

in $\sigma = (1, k_2, \ldots, k_i, \ldots, k_j, \ldots, k_n)$ is also equal to the number of pairs of

$$k_i^* > k_j^* \qquad \text{where} \qquad i < j$$

in $\sigma^* = (k_1^*, k_2^*, \ldots, k_i^*, \ldots, k_j^*, \ldots, k_{n-1}^*)$ which is a permutation of

$$\sigma_{\text{no}} = (1, 2, \ldots, n-1)$$

- Hence

$$\varepsilon_{1 k_2 \cdots k_i \cdots k_n} = \begin{cases} 0 & \text{if there exists } k_i = 1 \text{ for any } i = 2, \ldots n, \\ \varepsilon_{k_1^* \cdots k_{n-1}^*} & \text{otherwise.} \end{cases}$$

Proof

- If follows then that

$$\begin{aligned}\alpha &= a_{11}\sum_{k_2=1}^{n}\cdots\sum_{k_n=1}^{n}\varepsilon_{1k_2\cdots k_n}a_{2k_2}a_{3k_3}\cdots a_{nk_n}\\ &= a_{11}\sum_{k_2=2}^{n}\cdots\sum_{k_n=2}^{n}\varepsilon_{1k_2\cdots k_n}a_{2k_2}a_{3k_3}\cdots a_{nk_n}\\ &= a_{11}\sum_{k_1^*=1}^{n-1}\cdots\sum_{k_{n-1}^*=1}^{n-1}\varepsilon_{k_1^*\cdots k_{n-1}^*}a_{2(k_1^*+1)}a_{3(k_1^*+1)}\cdots a_{n(k_{n-1}^*+1)}\\ &= a_{11}\sum_{k_1^*=1}^{n-1}\cdots\sum_{k_{n-1}^*=1}^{n-1}\varepsilon_{k_1^*\cdots k_{n-1}^*}a_{1k_1^*}^*a_{2k_2^*}^*\cdots a_{(n-1)k_{n-1}^*}^* = a_{11}\det(\mathbf{A}^*)\end{aligned}$$

where $\mathbf{A}^* = [a_{ij}^*]$ is the submatrix without the 1st row and column of $\mathbf{A}$.

- Since $C_{11} = (-1)^{1+1}M_{11} = M_{11} = \det(\mathbf{A}^*)$, we have $\alpha = a_{11}C_{11}$.

Proof

- So we have shown the term involving a_{11} in

$$\det(\mathbf{A}) = a_{11}C_{11} + \sum_{k_1=2}^{n}\sum_{k_2=1}^{n}\cdots\sum_{k_n=1}^{n}\varepsilon_{k_1\cdots k_n}a_{1k_1}a_{2k_2}a_{3k_3}\cdots a_{nk_n}$$

  is the cofactor of a_{11}. Note the sum can be split according to coefficients a_{ij}

$$\det(\mathbf{A}) = \begin{cases} a_{i1}\hat{C}_{i1} + a_{i2}\hat{C}_{i2} + \cdots + a_{in}\hat{C}_{in} & \text{for} \quad i = 1, \ldots n \\ a_{1j}\hat{C}_{1j} + a_{2j}\hat{C}_{2j} + \cdots + a_{nj}\hat{C}_{nj} & \text{for} \quad j = 1, \ldots n \end{cases}$$

  where $\hat{C}_{ij}$ is a sum involves no elements from row i or column j.

- We just need to show for every coefficient a_{ij},

$$\hat{C}_{ij} = C_{ij} \qquad \text{where } C_{ij} \text{ is the cofactor of } a_{ij}.$$

- Notice we can put a_{ij} into the 1st-row-1st-column position by interchanging adjacent rows $i-1$ times and interchanging adjacent columns $j-1$ times.

Proof

- Let the result matrix after the interchanges to be $\mathbf{A}'$. Consequently

$$\det(\mathbf{A}) = (-1)^{i+j-2}\det(\mathbf{A}') = (-1)^{i+j}\det(\mathbf{A}')$$

- Since the relative positions of elements in the rest of rows and columns in

$$\mathbf{A}$$

remain intact after the interchanges, that is, the submatrix

$$\mathbf{M}_{ij}$$

obtained by deleting row i and column j of $\mathbf{A}$ is identical to the submatrix

$$\mathbf{M}'_{11}$$

obtained by deleting the 1st row and the 1st column of $\mathbf{A}'$. Hence

$$C'_{11} = M'_{11} = M_{ij}$$

Proof

- By the construction of $\mathbf{A}'$, we have

$$a_{ij} = a'_{11}$$

- Using our results we have discussed so far in this proof, we have

$$\det(\mathbf{A}) = (-1)^{i+j}\left(a'_{11}C'_{11} + \beta\right)$$
$$= a_{ij}C_{ij} + (-1)^{i+j}\beta$$

  where β denotes terms that does not involve $a_{ij} = a'_{11}$, which shows indeed

$$\hat{C}_{ij} = C_{ij}$$

- This establishes that the cofactor expansion is true along a particular row, the expansion along a particular column follows immediately, since

$$\det(\mathbf{A}) = \det\left(\mathbf{A}^{\mathrm{T}}\right) \quad \square$$