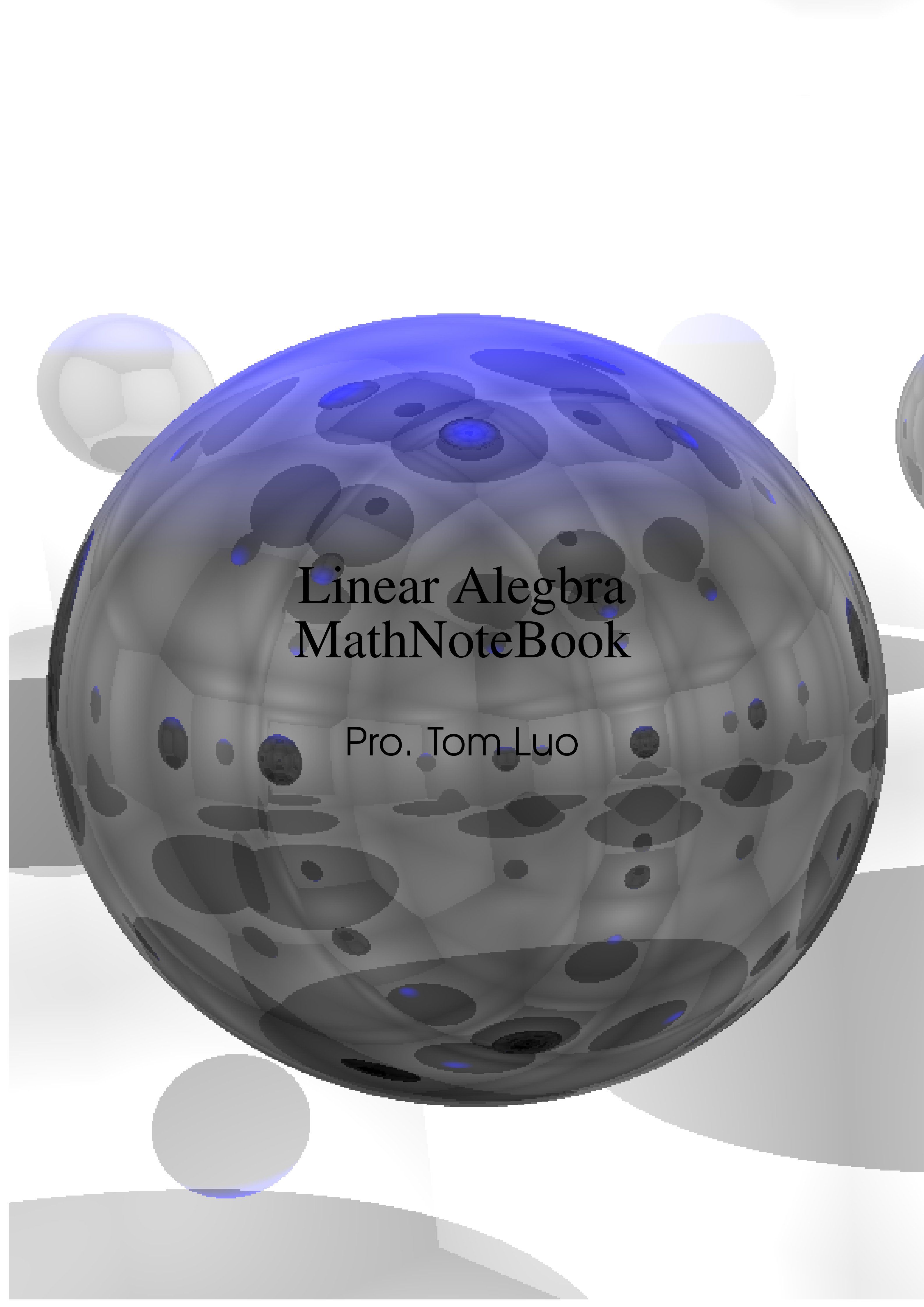
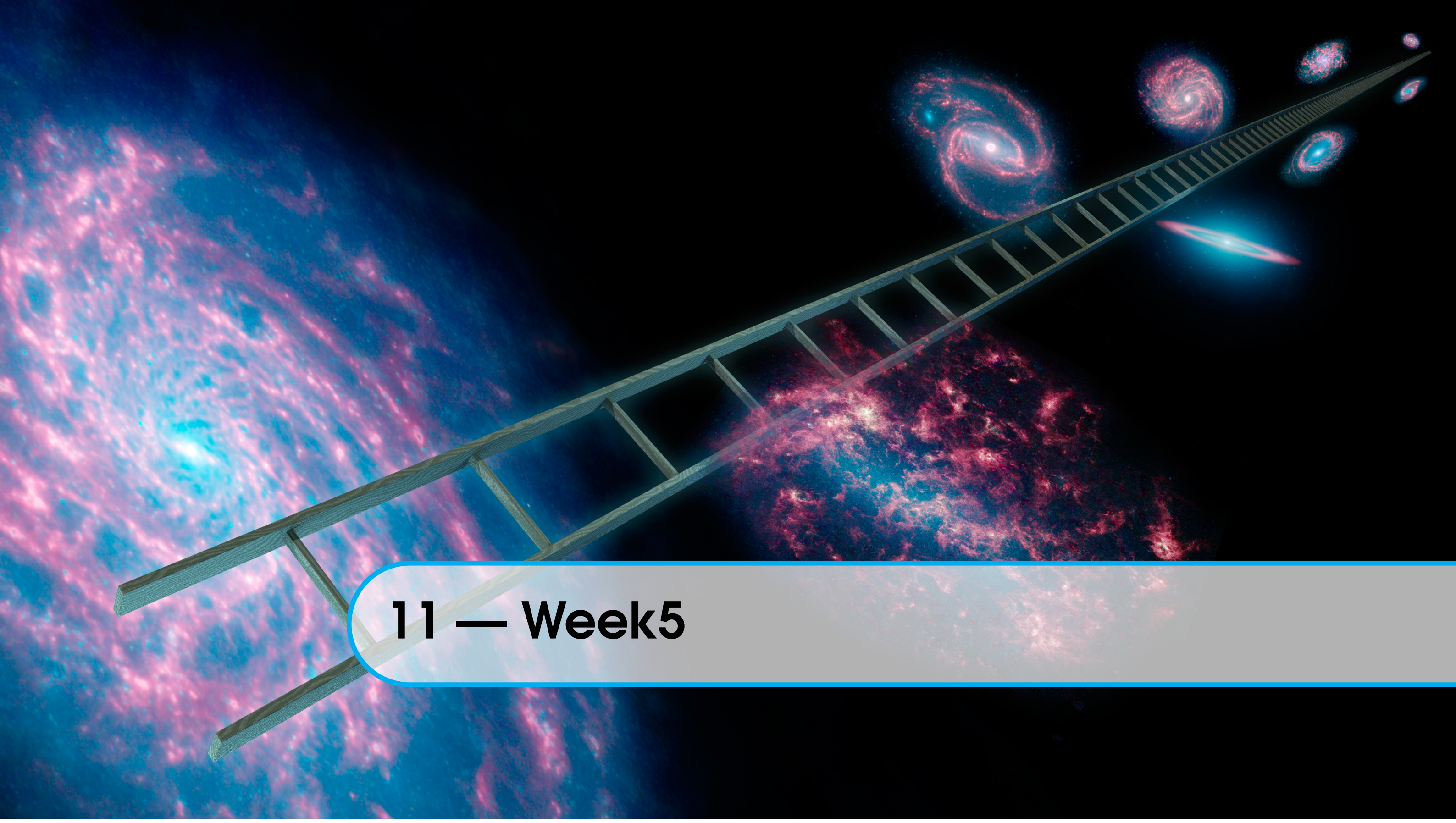


Linear Alegbra MathNoteBook

Pro. Tom Luo



11 — Week5

11.1 Tuesday

11.1.1 Formulas for Determinant

We want to use the **3 basic properties** to derive the formula for determinant:

1. **The determinant of the n by n identity matrix is 1.**

$$\begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} = 1 \qquad \text{and} \qquad \begin{vmatrix} 1 & & \\ & \ddots & \\ & & 1 \end{vmatrix} = 1.$$

2. **The determianant changes sign when two rows are exchanged.** (sign reversal)

$$\text{Check: } \begin{vmatrix} c & d \\ a & b \end{vmatrix} = -\begin{vmatrix} a & b \\ c & d \end{vmatrix} \qquad \text{(both sides equal } bc - ad\text{).}$$

3. **The determinant is a linear function of each row separately.** (all other rows stay fixed).

$$\textbf{multiply row 1 by any number } t \qquad \begin{vmatrix} ta & tb \\ c & d \end{vmatrix} = t\begin{vmatrix} a & b \\ c & d \end{vmatrix}$$

$$\textbf{Add row 1 of } A \textbf{ to row 1 of } B\textbf{:} \qquad \begin{vmatrix} a_1 + a_2 & b_1 + b_2 \\ c & d \end{vmatrix} = \begin{vmatrix} a_1 & b_1 \\ c & d \end{vmatrix} + \begin{vmatrix} a_2 & b_2 \\ c & d \end{vmatrix}$$

Although we derive the formula for $\det \boldsymbol{A}$ is $\det \boldsymbol{A} = \pm \prod_i \text{pivots}_i$ (product of pivots), it is not **explicit**. We begin some example to show how to derive the explicit formula for determinant.

■ **Example 11.1** To derive the formula for determinant, let's start with $n = 2$.
Given $\boldsymbol{A} = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, our goal is to get $ad - bc$.
We can break each row into two simpler rows:

$$\begin{vmatrix} a & b \end{vmatrix} = \begin{vmatrix} a & 0 \end{vmatrix} + \begin{vmatrix} 0 & b \end{vmatrix} \qquad \text{and} \qquad \begin{vmatrix} c & d \end{vmatrix} = \begin{vmatrix} c & 0 \end{vmatrix} + \begin{vmatrix} 0 & d \end{vmatrix}$$

Now apply property 3, first in row 1(with row 2 fixed) and then in row 2(with row 1 fixed):

$$\begin{aligned}\begin{vmatrix} a & b \\ c & d \end{vmatrix} &= \begin{vmatrix} a & 0 \\ c & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & d \end{vmatrix} \\ &= \begin{vmatrix} a & 0 \\ c & 0 \end{vmatrix} + \begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} + \begin{vmatrix} 0 & b \\ 0 & d \end{vmatrix}\end{aligned}$$

The last line has $2^2 = 4$ determinants. The first and fourth are zero since their rows are **dep.** (one row is a multiple of the other row.) We left two terms to compute:

$$\begin{vmatrix} a & 0 \\ 0 & d \end{vmatrix} + \begin{vmatrix} 0 & b \\ c & 0 \end{vmatrix} = ad \begin{vmatrix} 1 & 0 \\ 0 & 1 \end{vmatrix} + bc \begin{vmatrix} 0 & 1 \\ 1 & 0 \end{vmatrix} = ad - bc$$

The permutation matrices $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ have determinant $+1$ or -1. ■

■ **Example 11.2** Now we try $n = 3$. Each row splits into 3 simpler rows such as $\begin{bmatrix} a_{11} & 0 & 0 \end{bmatrix}$. Hence $\det \boldsymbol{A}$ will split into $3^3 = 27$ simple determinants. For simple determinant, if one column has two nonzero entries, (For example, $\begin{bmatrix} a_{11} & 0 & 0 \\ a_{21} & 0 & 0 \\ 0 & 0 & a_{33} \end{bmatrix}$), then its determinant will be zero.

Hence we only need to foucus on the matrix that **the nonzero terms come from defferent columns**:

$$\begin{aligned}\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} &= \begin{vmatrix} a_{11} & & \\ & a_{22} & \\ & & a_{33} \end{vmatrix} + \begin{vmatrix} & a_{12} & \\ & & a_{23} \\ a_{31} & & \end{vmatrix} + \begin{vmatrix} & & a_{13} \\ a_{21} & & \\ & a_{32} & \end{vmatrix} \\ &+ \begin{vmatrix} a_{11} & & \\ & & a_{23} \\ & a_{32} & \end{vmatrix} + \begin{vmatrix} & a_{12} & \\ a_{21} & & \\ & & a_{33} \end{vmatrix} + \begin{vmatrix} & & a_{13} \\ & a_{22} & \\ a_{31} & & \end{vmatrix}\end{aligned}$$

There are $3! = 6$ ways to permutate the three columns, so there leaves six determinants. The six permutations of $(1,2,3)$ is given by:

$$\textbf{Column numbers } = (1,2,3),(2,3,1),(3,1,2),(1,3,2),(2,1,3),(3,2,1).$$

The last three are *odd permutations* (One exchange from identity permutation $(1,2,3)$.) The first three are *even permutations*. (zero or two exchange from identity permutation $(1,2,3)$.) When the column number is (α,β,ω), we get the entries $a_{1\alpha}, a_{2\beta}, a_{3\omega}$. The permutation (α,β,ω) comes with a plus or minus sign. If you don't understand, look at example below:

$$\begin{aligned}\det \boldsymbol{A} = a_{11}a_{22}a_{33} &\begin{vmatrix} 1 & & \\ & 1 & \\ & & 1 \end{vmatrix} + a_{12}a_{23}a_{31} \begin{vmatrix} & 1 & \\ & & 1 \\ 1 & & \end{vmatrix} + a_{13}a_{21}a_{32} \begin{vmatrix} & & 1 \\ 1 & & \\ & 1 & \end{vmatrix} \\ + a_{11}a_{23}a_{32} &\begin{vmatrix} 1 & & \\ & & 1 \\ & 1 & \end{vmatrix} + a_{12}a_{21}a_{33} \begin{vmatrix} & 1 & \\ 1 & & \\ & & 1 \end{vmatrix} + a_{13}a_{22}a_{31} \begin{vmatrix} & & 1 \\ & 1 & \\ 1 & & \end{vmatrix}\end{aligned}$$

The first three (even) permutation matrices have $\det \boldsymbol{P} = +1$, the last three (odd) permutation matrices have $\det \boldsymbol{P} = -1$. Hence we have:

$$\begin{aligned}\det \boldsymbol{A} &= a_{11}a_{22}a_{33} + a_{12}a_{23}a_{31} + a_{13}a_{21}a_{32} - a_{11}a_{23}a_{32} - a_{12}a_{21}a_{33} - a_{13}a_{22}a_{31} \\ &= a_{11}(a_{22} - a_{33}) + a_{12}(a_{23}a_{31} - a_{21}a_{33}) + a_{13}(a_{21}a_{32} - a_{22}a_{31})\end{aligned}$$

■

n by n formula

Now we can see n by n formula. There are $n!$ permutations of columns, so we have $n!$ terms for determinant.
Assuming $(\alpha, \beta, \dots, \omega)$ is the permutation of $(1, 2, \dots, n)$. The coorsponding terms is $a_{1\alpha}a_{2\beta}\dots a_{n\omega} \det \boldsymbol{P}$, where $\boldsymbol{P}$ is the permutation matrix with column number $\alpha, \beta, \dots, \omega$.
The complete determinant of $\boldsymbol{A}$ is the sum of these $n!$ simple determinant. $a_{1\alpha}a_{2\beta}\dots a_{n\omega}$ is obtained by choosing **one entry from every row and every column:**

Definition 11.1 — Big formula for determinant.

$$\begin{aligned}\det \boldsymbol{A} &= \text{sum of all } n! \text{ column permutations} \\ &= \sum (\det \boldsymbol{P}) a_{1\alpha}a_{2\beta}\dots a_{n\omega} = \textbf{BIG FORMULA}\end{aligned}$$

where $\boldsymbol{P}$ is permutation matrix with column number $(\alpha, \beta, \dots, \omega)$. And $\{\alpha, \beta, \dots, \omega\}$ is a permutation of $\{1, 2, \dots, n\}$. ■

R

Complexity Analysis

However, if we want to use big formula to compute matrix, we need to do $n!(n-1)$ multiplications. If we use formula $\det \boldsymbol{A} = \pm \prod pivots$, we only need to do $O(n^3)$ multiplications. Hence the letter one is quite more efficient.

Verify property

We can use the big formula to verify property 1 to property 3:

- $\det \boldsymbol{I} = 1$:
  Only when $(\alpha, \beta, \dots, \omega)$=$(1, 2, \dots, n)$, there is no zero entries for $a_{1\alpha}a_{2\beta}\dots a_{n\omega}$. Hence $\det \boldsymbol{A} = a_{11}a_{22}\dots a_{nn} = 1$.
- **sign reversal**:
  If two rows are interchanged, then all determinant of permutation matrix will change its sign, hence the value for determinant $\boldsymbol{A}$ is opposite.
- **The determinant is a linear function of each row separately**.
  If we separate out the fator $a_{11}, a_{12}, \dots, a_{1\alpha}$ **that comes from the first row**, this property is easy to check. For 3 by 3 matrix, separate the usual 6 terms of the determinant into 3 pairs:

  $$\det \boldsymbol{A} = a_{11}(a_{22}a_{33} - a_{23}a_{32}) + a_{12}(a_{23}a_{31} - a_{21}a_{33}) + a_{13}(a_{21}a_{32} - a_{22}a_{31}).$$

  Those three quantities in parentheses are called **cofactors**. They are 2×2 determinant coming from matrices in row 2 and 3. The first row contributes the factors a_{11}, a_{12}, a_{13}. The lower rows contribute the cofactors $(a_{22}a_{33} - a_{23}a_{32}), (a_{23}a_{31} - a_{21}a_{33}), (a_{21}a_{32} - a_{22}a_{31})$. Certainly $\det \boldsymbol{A}$ depends **linearly** on a_{11}, a_{12}, a_{13}, which is property 3.

11.1.2 Determinant by Cofactors

We could write the determinant in this form:

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = \begin{vmatrix} a_{11} & & \\ & a_{22} & a_{23} \\ & a_{32} & a_{33} \end{vmatrix} + \begin{vmatrix} & a_{12} & \\ a_{21} & & a_{23} \\ a_{31} & & a_{33} \end{vmatrix} + \begin{vmatrix} & & a_{13} \\ a_{21} & a_{22} & \\ a_{31} & a_{32} & \end{vmatrix}.$$

If we define $\boldsymbol{A}_{1j}$ to be **the submatrix obtained by removing row 1 and column j**, We could compute $\det \boldsymbol{A}$ in this way:

The cofactors along row 1 are $C_{1j} = (-1)^{1+j} \det \boldsymbol{A}_{1j} \quad j = 1, 2, \dots, n.$

The cofactor expansion is $\det \boldsymbol{A} = a_{11}C_{11} + a_{12}C_{12} + \cdots + a_{1n}C_{1n}.$

More generally, we can cross row i to get the determinant:

Definition 11.2 — Determinant. The determinant is the **dot product** of any row i of $\boldsymbol{A}$ with its cofactors using other rows:

Cofactor Formula $\quad \det \boldsymbol{A} = a_{i1}C_{i1} + a_{i2}C_{i2} + \cdots + a_{in}C_{in}.$

Each cofactor C_{ij} is defined as:

Cofactor $\quad C_{ij} = (-1)^{i+j} \det \boldsymbol{A}_{ij}$

where A_{ij} is the submatrix obtained by removing row 1 and column 1.

Moreover, we can construct $\det \boldsymbol{A}$ from its properties. Since we have $\det \boldsymbol{A} = \det \boldsymbol{A}^{\mathrm{T}}$, we can expand the determinant in cofactors *down a column* instead of across a row. Down column j the entries are a_{1j} to a_{nj}, the cofactors are C_{1j} to C_{nj}. The determinant is given by:

Cofactors down column j: $\quad \det \boldsymbol{A} = a_{1j}C_{1j} + a_{2j}C_{2j} + \cdots + a_{nj}C_{nj}.$

11.1.3 Determinant Applications

Inverse

It's easy to check that the inverse of 2 by 2 matrix $\boldsymbol{A}$ is

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}^{-1} = \frac{1}{ad-bc} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} = \frac{1}{\det \boldsymbol{A}} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}.$$

We could use determinant to compute inverse! Before that let's define **cofactor matrix**:

Definition 11.3 — cofactor matrix. The cofactor matrix of $n \times n$ matrix $\boldsymbol{A}$ is given by:

$$\boldsymbol{C} = \left[C_{ij}\right]_{1 \le i,j \le n}$$

where C_{ij} is the cofactor for a_{ij}.

Then we try to derive the inverse of matrix $\boldsymbol{A}$: For $n \times n$ matrix $\boldsymbol{A}$, the product of $\boldsymbol{A}$ and the **transpose** of *cofactor matrix* is given by:

$$\boldsymbol{A}\boldsymbol{C}^{\mathrm{T}} = \begin{bmatrix} a_{11} & \dots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \dots & a_{nn} \end{bmatrix} \begin{bmatrix} C_{11} & \dots & C_{n1} \\ \vdots & \ddots & \vdots \\ C_{1n} & \dots & C_{nn} \end{bmatrix} = \begin{bmatrix} \det \boldsymbol{A} & & \\ & \det \boldsymbol{A} & \\ & & \det \boldsymbol{A} \end{bmatrix} \tag{11.1}$$

Explain:

- Row 1 of $\boldsymbol{A}$ times the column 1 of $\boldsymbol{C}^{\mathrm{T}}$ yields the first $\det\boldsymbol{A}$ on the right:

$$a_{11}C_{11}+a_{12}C_{12}+\cdots+a_{1n}C_{1n}=\det\boldsymbol{A}$$

Similarly, row j of $\boldsymbol{A}$ times column j of $\boldsymbol{C}^{\mathrm{T}}$ yields the determinant.

- *How to explain the zeros off the main diagonal in equation (11.1)?* Rows of $\boldsymbol{A}$ are multiplying $\boldsymbol{C}^{\mathrm{T}}$ from **different** columns. Why is the answer zero? For example, the (2,1)th entry of the result is given by

Row 2 of $\boldsymbol{A}$
Row 1 of $\boldsymbol{C}$
$$a_{21}C_{11}+a_{22}C_{12}+\cdots+a_{2n}C_{1n}=0.$$

Answer: If the second row of $\boldsymbol{A}$ is copied into its first row, we define this new matrix as $\boldsymbol{A}^*$. Thus the determinant of $\boldsymbol{A}^*$ is given by:

$$\begin{vmatrix} a_{21} & a_{22} & \dots & a_{2n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = \begin{vmatrix} a_{21} & & & \\ & a_{22} & \dots & a_{2n} \\ & a_{32} & \dots & a_{3n} \\ & \vdots & \ddots & \vdots \\ & a_{n2} & \dots & a_{nn} \end{vmatrix} + \begin{vmatrix} & a_{22} & & \\ a_{21} & & \dots & a_{2n} \\ a_{31} & & \dots & a_{3n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & & \dots & a_{nn} \end{vmatrix} + \cdots + \begin{vmatrix} & & & a_{2n} \\ a_{21} & a_{22} & a_{2(n-1)} & \\ a_{31} & a_{32} & a_{3(n-1)} & \\ \vdots & \vdots & \vdots & \\ a_{n1} & a_{n2} & a_{n(n-1)} & \end{vmatrix}$$

Equivalently, we have

$$\det\boldsymbol{A}^* = \begin{vmatrix} a_{21} & a_{22} & \dots & a_{2n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ a_{31} & a_{32} & \dots & a_{3n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix} = a_{21}C_{11}+a_{22}C_{12}+\cdots+a_{2n}C_{1n}$$

Since $\boldsymbol{A}^*$ has two equal rows, the determinant must be zero.
Hence $a_{21}C_{11}+a_{22}C_{12}+\cdots+a_{2n}C_{1n}=0$. Similarly, **all entries off the main diagonal are zero.**

Thus the equation (11.1) is correct:

$$\boldsymbol{A}\boldsymbol{C}^{\mathrm{T}} = \begin{bmatrix} \det\boldsymbol{A} & & \\ & \det\boldsymbol{A} & \\ & & \det\boldsymbol{A} \end{bmatrix} = \det(\boldsymbol{A})\boldsymbol{I} \implies \boldsymbol{A}^{-1} = \frac{1}{\det\boldsymbol{A}}\boldsymbol{C}^{\mathrm{T}}.$$

Hence we could compute the inverse by computing many determinant of submatrix:

Definition 11.4 — Inverse. **The $(\boldsymbol{i},\boldsymbol{j})$th entry of $\boldsymbol{A}^{-1}$ is the cofactor C_{ji}(not C_{ji}) divided by $\det\boldsymbol{A}$:**

Formula for $\boldsymbol{A}^{-1}$ $\quad (\boldsymbol{A}^{-1})_{ij} = \dfrac{C_{ji}}{\det\boldsymbol{A}}$ and $\boldsymbol{A}^{-1} = \dfrac{\boldsymbol{C}^{\mathrm{T}}}{\det\boldsymbol{A}}$

Cramer's Rule

Cramer's Rule solves $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$.

Assume $\boldsymbol{A}$ is a $n \times n$ matrix that is **nonsingular**. Then we can use determinant to solve this system:

Let's start with $n = 3$. We could multiply $\boldsymbol{A}$ with a new matrix $\boldsymbol{C}_1$ to get $\boldsymbol{B}_1$:

$$\textbf{Key idea:} \qquad \boldsymbol{A}\boldsymbol{C}_1 = \begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} \begin{bmatrix} x_1 & 0 & 0 \\ x_2 & 1 & 0 \\ x_3 & 0 & 1 \end{bmatrix} = \begin{bmatrix} b_1 & a_{12} & a_{13} \\ b_2 & a_{22} & a_{23} \\ b_3 & a_{32} & a_{33} \end{bmatrix} = \boldsymbol{B}_1$$

Taking determinants both sides, then we have

$$\det(\boldsymbol{A}\boldsymbol{C}_1) = \det(\boldsymbol{A})\det(\boldsymbol{C}_1) = \det(\boldsymbol{A})(x_1) = \det \boldsymbol{B}_1 \implies x_1 = \frac{\det \boldsymbol{B}_1}{\det \boldsymbol{A}_1}.$$

The matrix $\boldsymbol{C}_1$ is obtained by putting $\boldsymbol{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}$ into the *first* column of the *identity matrix*.

Similarly, we could get all x_j in this way. $(i = 1, \dots, n)$.

Definition 11.5 — Cramer's Rule. If $\det \boldsymbol{A}$ is not zero, $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$ could be solved by determinants:

$$x_1 = \frac{\det \boldsymbol{B}_1}{\det \boldsymbol{A}} \qquad x_2 = \frac{\det \boldsymbol{B}_2}{\det \boldsymbol{A}} \qquad \dots\dots \qquad x_n = \frac{\det \boldsymbol{B}_n}{\det \boldsymbol{A}}$$

The matrix $\boldsymbol{B}_j$ has the jth column of $\boldsymbol{A}$ replaced by the vector $\boldsymbol{b}$. In other words,

$$\boldsymbol{B}_j = \begin{bmatrix} a_{11} & \dots & b_1 & \dots & a_{1n} \\ a_{21} & \dots & b_2 & \dots & a_{2n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \dots & b_n & \dots & a_{nn} \end{bmatrix} \qquad j = 1, \dots, n.$$

11.1.4 Orthogonality and Projection

Definition 11.6 — Orthogonal vectors.
Two vectors $\boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$ are orthogonal when their inner product is zero:

$$<\boldsymbol{x}, \boldsymbol{y}> = \sum_{i=1}^{n} x_i y_i = 0.$$

R Note that the inner product of two vectors satisfies the *commutative rule*. In other words, $<\boldsymbol{x}, \boldsymbol{y}> = <\boldsymbol{y}, \boldsymbol{x}>$ for vectors $\boldsymbol{x}$ and $\boldsymbol{y}$. Generally, if the result of inner product is a scalar, then inner product satisfies commutative rule.

An important case is the inner product of a vector with *itself*. The inner product $<\boldsymbol{x}, \boldsymbol{x}>$ gives the *length of* $\boldsymbol{v}$ *squared*.

Definition 11.7 — length/norm.
The **length(norm)** $||\boldsymbol{x}||$ of a vector $\boldsymbol{x} \in \mathbb{R}$ is the square root of $< \boldsymbol{x}, \boldsymbol{x} >$:

$$\textbf{length} = ||\boldsymbol{x}|| = \sqrt{< \boldsymbol{x}, \boldsymbol{x} >} = \sqrt{x_1^2 + \cdots + x_n^2}.$$

Function space

We can talk about inner product between functions under function space. For example, if we define $V = \{f(t) | \int_0^1 f^2(t)dt < \infty\}$, then we can define inner product and norm under V:

Definition 11.8 — Inner product; norm. The **inner product** of $f(x)$ and $g(x)$, and the **norm** are defined as:

$$< f, g >= \int_0^1 f(x)g(x)dx \qquad \text{and} \qquad ||f||^2 = \sqrt{\int_0^1 f^2(x)dx}$$

Moreover, when $< f, g >= 0$, we say two functions are **orthogonal** and denote it as $f \perp g$.

Cauchy-Schwarz Inequality

In $\mathbb{R}^2$, suppose $\boldsymbol{x} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, \boldsymbol{y} = \begin{pmatrix} y_1 \\ y_2 \end{pmatrix}$, then we set:

$$x_1 = ||\boldsymbol{x}|| \cos\theta \quad x_2 = ||\boldsymbol{x}|| \sin\theta \qquad y_1 = ||\boldsymbol{y}|| \cos\varphi \quad y_2 = ||\boldsymbol{y}|| \sin\varphi$$

The inner product of $\boldsymbol{x}$ and $\boldsymbol{y}$ is given by:

$$\begin{aligned} < \boldsymbol{x}, \boldsymbol{y} >= \boldsymbol{x}^{\mathrm{T}}\boldsymbol{y} &= x_1x_2 + y_1y_2 \\ &= ||\boldsymbol{x}|| ||\boldsymbol{y}|| (\cos\theta\cos\varphi + \sin\theta\sin\varphi \\ &= ||\boldsymbol{x}|| ||\boldsymbol{y}|| \cos(\theta - \varphi) \end{aligned}$$

Since $|\cos(\theta - \varphi)|$ never exceeds 1, the cosine formula gives great inequality:

Theorem 11.1 — Cauchy Schwarz Inequality.
$| < \boldsymbol{x}, \boldsymbol{y} > | \leq ||\boldsymbol{x}|| ||\boldsymbol{y}||$ for two vectors $\boldsymbol{x}$ and $\boldsymbol{y}$.

Proof. Firstly, we want to find t^* such that $\min ||\boldsymbol{x} - t\boldsymbol{y}||^2 = ||\boldsymbol{x} - t^*\boldsymbol{y}||^2$.

$$\begin{aligned} ||\boldsymbol{x} - t\boldsymbol{y}||^2 &=< \boldsymbol{x} - t\boldsymbol{y}, \boldsymbol{x} - t\boldsymbol{y} >=< \boldsymbol{x}, \boldsymbol{x} > + < -t\boldsymbol{y}, \boldsymbol{x} > + < \boldsymbol{x}, -t\boldsymbol{y} > + < -t\boldsymbol{y}, -t\boldsymbol{y} > \\ &= ||\boldsymbol{x}||^2 - t < \boldsymbol{y}, \boldsymbol{x} > -t < \boldsymbol{x}, \boldsymbol{y} > +t^2||\boldsymbol{y}||^2 \\ &= ||\boldsymbol{x}||^2 - 2t < \boldsymbol{x}, \boldsymbol{y} > +t^2||\boldsymbol{y}||^2 \end{aligned}$$

Hence the minimizer $t*$ must satisfy

$$\Delta = 0 \implies t^* = \frac{< \boldsymbol{x}, \boldsymbol{y} >}{||\boldsymbol{y}||^2}$$

Hence we have

$$
\begin{aligned}
\|\boldsymbol{x}-t\boldsymbol{y}\|_{\min}^2 = \|\boldsymbol{x}-t^*\boldsymbol{y}\|^2 = \|\boldsymbol{x}\|^2 - \frac{<\boldsymbol{x},\boldsymbol{y}>^2}{\|\boldsymbol{y}\|^2} \\
= \frac{\|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2 - <\boldsymbol{x},\boldsymbol{y}>^2}{\|\boldsymbol{y}\|^2} \geq 0 \\
\implies \|\boldsymbol{x}\|^2\|\boldsymbol{y}\|^2 \geq <\boldsymbol{x},\boldsymbol{y}>^2
\end{aligned}
$$

Or equivalently,

$$|<\boldsymbol{x},\boldsymbol{y}>| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|.$$

■

If we consider functions f,g as vectors, then **Cauchy-Schwarz** inequality also holds:

$$\left[\int_0^1 f(t)g(t)dt\right] \leq \int_0^1 f^2 dt \int_0^1 g^2 dt$$

Since $|<\boldsymbol{x},\boldsymbol{y}>| \leq \|\boldsymbol{x}\|\|\boldsymbol{y}\|$, we have

$$-1 \leq \frac{<\boldsymbol{x},\boldsymbol{y}>}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} \leq 1$$

If we define $\frac{<\boldsymbol{x},\boldsymbol{y}>}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|} = \cos\theta$, then $<\boldsymbol{x},\boldsymbol{y}> = \|\boldsymbol{x}\|\|\boldsymbol{y}\|\cos\theta$.

R This equality holds for **Hilbert space**, which will be discussed later.

Orthogonal for space

Also, we can discuss orthogonality for space:

Definition 11.9 — Orthogonal subspaces. Two subspaces $\boldsymbol{U}$ and $\boldsymbol{V}$ of a vector space are **orthogonal** if every vector $\boldsymbol{u}$ in $\boldsymbol{U}$ is *perpendicular* to every vector $\boldsymbol{v}$ in $\boldsymbol{V}$:

Orthogonal subspaces $\quad \boldsymbol{u}^{\mathrm{T}}\boldsymbol{v} = 0$ for all $\boldsymbol{u}$ in $\boldsymbol{U}$ and all $\boldsymbol{v}$ in $\boldsymbol{V}$.

■