

# ENG1005 Week3 Workshop Problem Set Solutions

Yang Xingyu (33533563)

August 11, 2024

## Applied Class Challenges

### Multiplicative Property of Determinant

**Theorem 1** (Multiplicative Property of Determinant under Matrix Multiplication):  
For any  $n \times n$  square matrices  $\mathbf{A}$  and  $\mathbf{B}$ :

$$|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|,$$

and this implies

$$\text{adj}(\mathbf{AB}) = \text{adj}(\mathbf{B}) \text{adj}(\mathbf{A}).$$

*Proof.* Assuming the following proposition is true, we aim to prove Theorem 1.

**Lemma 1:**  $\forall n \times n$  matrix  $\mathbf{A}$ ,  $\mathbf{A} \text{adj}(\mathbf{A}) = \text{adj}(\mathbf{A})\mathbf{A} = |\mathbf{A}|\mathbf{I}$ .

By rearrangement, we have

$$\mathbf{A} \frac{\text{adj}(\mathbf{A})}{|\mathbf{A}|} = \mathbf{I} \quad (1)$$

Since  $\mathbf{A}$  is non-singular, then there must exist a unique multiplicative inverse of  $\mathbf{A}$ , which we call  $\mathbf{A}^{-1}$ , such that  $\mathbf{AA}^{-1} = \mathbf{A}^{-1}\mathbf{A} = \mathbf{I}$ . Hence, we have

$$\frac{\text{adj}(\mathbf{A})}{|\mathbf{A}|} = \mathbf{A}^{-1} \quad (2)$$

A new lemma is needed here to proceed with the proof.

**Lemma 2:** If a **non-singular** matrix  $\mathbf{A}$  is the product of two other matrices  $\mathbf{M}$  and  $\mathbf{N}$ , such that  $\mathbf{A} = \mathbf{MN}$ , then

$$\mathbf{A}^{-1} = \mathbf{N}^{-1}\mathbf{M}^{-1}.$$

By Lemma 2 and equation (2), we have

$$(\mathbf{MN})^{-1} = \mathbf{N}^{-1}\mathbf{M}^{-1} \quad (3)$$

Again, by equation (2), we have

$$\mathbf{N}^{-1} = \frac{\text{adj}(\mathbf{N})}{|\mathbf{N}|}, \quad \mathbf{M}^{-1} = \frac{\text{adj}(\mathbf{M})}{|\mathbf{M}|}. \quad (4)$$

Combining equations (3) and (4), we get

$$(\mathbf{MN})^{-1} = \mathbf{N}^{-1}\mathbf{M}^{-1} = \frac{\text{adj}(\mathbf{N})}{|\mathbf{N}|} \cdot \frac{\text{adj}(\mathbf{M})}{|\mathbf{M}|}$$

Therefore,

$$(\mathbf{MN})^{-1} = \frac{\text{adj}(\mathbf{MN})}{|\mathbf{MN}|}$$

Given that  $(\mathbf{MN})^{-1} = \mathbf{N}^{-1}\mathbf{M}^{-1}$ , we have:

$$\frac{\text{adj}(\mathbf{MN})}{|\mathbf{MN}|} = \frac{\text{adj}(\mathbf{N})}{|\mathbf{N}|} \cdot \frac{\text{adj}(\mathbf{M})}{|\mathbf{M}|}$$

This implies:

$$\text{adj}(\mathbf{MN}) = \text{adj}(\mathbf{N}) \text{adj}(\mathbf{M})$$

and

$$|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$$

This is because, the structure of the last identity is a fraction on both sides, and the numerator on LHS is a matrix, while a product of matrix on the RHS. The denominator on LHS  $|\mathbf{MN}| \in \mathbb{F}$ , and the denominator on RHS is a product of elements in some  $\mathbb{F}$ . Both the set of all square matrix and any field  $\mathbb{F}$  are vector space, and thus the multiplication is closed on their respective set of elements. Hence, we must have the numerator on the LHS is the product of the numerators on RHS, and same goes for the denominator, due to the closure of multiplication under vector space.

Now we will discuss the case under **non-singular matrices**. When one of or both matrices  $\mathbf{A}, \mathbf{B}$  are both singular, then the product of the two matrices will still be a singular matrix, as the rank of both original matrices are not of full rank. Thus, in whatever case, we have  $0 = 0 \times 0$ . Hence, the part for determinant is proven, and we may move to the adjugate matrix.

Considering

$$\begin{cases} (\mathbf{AB}) \text{adj}(\mathbf{AB}) = |\mathbf{AB}|\mathbf{I} \\ \mathbf{A} \text{adj}(\mathbf{A}) = |\mathbf{A}|\mathbf{I} \\ \mathbf{B} \text{adj}(\mathbf{B}) = |\mathbf{B}|\mathbf{I} \end{cases}.$$

As long as more than one of the singular matrix  $\mathbf{A}, \mathbf{B}$ , by what we have previously shown,  $|\mathbf{B}|\mathbf{I} \cdot |\mathbf{A}|\mathbf{I} = |\mathbf{AB}|\mathbf{I} = \mathbf{0}$ . Therefore,

$$(\mathbf{AB}) \text{adj}(\mathbf{AB}) = \mathbf{B} \text{adj}(\mathbf{B}) \cdot \mathbf{A} \text{adj}(\mathbf{A}) = |\mathbf{BA}| \cdot \mathbf{I}.$$

Now consider  $\mathbf{AB} \text{adj}(\mathbf{B}) \text{adj}(\mathbf{A})$ . We have

$$(\mathbf{AB}) \text{adj}(\mathbf{B}) \text{adj}(\mathbf{A}) = \mathbf{A}(\mathbf{B} \text{adj}(\mathbf{B})) \text{adj}(\mathbf{A}) = |\mathbf{B}|\mathbf{I} \cdot \mathbf{A} \text{adj}(\mathbf{A}) = |\mathbf{B}|\mathbf{I} \cdot |\mathbf{A}|\mathbf{I} = |\mathbf{BA}| \cdot \mathbf{I}.$$

Hence,  $\text{adj}(\mathbf{AB}) = \text{adj}(\mathbf{B}) \text{adj}(\mathbf{A})$ . We have also proven the case for one or more singular matrix.

This completes the proof. □

**Remark:** Only when I finished the part discussing case in non-singular matrix and construct  $\mathbf{A}\mathbf{B}\text{adj}(\mathbf{B})\text{adj}(\mathbf{A})$ , I realised that this could be used to rewrite the proof without discussing different cases.

Another approach is using mathematical induction (though I don't really like it). We can prove  $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$  and then show that the other statement holds, which is implied by the previous one.

*Proof.* We use mathematical induction to prove the previous statement.

Let predicate  $P(n)$  to denote that

$$\forall n \times n \text{ square matrix } \mathbf{A}, \mathbf{B}, |\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|.$$

**Base case** ( $n = 1$ )

When  $n = 1$ , we have

$$\mathbf{A} = [a], \quad \mathbf{B} = [b].$$

So  $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}| = ab$ , and obviously  $P(1)$  holds.

**Inductive Hypothesis** ( $n = k$ )

We assume that all  $P(k)$  is true. That is, for any  $k \times k$  matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1k} \\ a_{21} & a_{22} & \dots & a_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k1} & a_{k2} & \dots & a_{kk} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1k} \\ b_{21} & b_{22} & \dots & b_{2k} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k1} & b_{k2} & \dots & b_{kk} \end{bmatrix},$$

We have

$$|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|.$$

**Inductive Step** ( $n = k + 1$ )

Now we need to show that  $P(k + 1)$  is true. That is, for any two  $k + 1 \times k + 1$  matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1,k+1} \\ a_{21} & a_{22} & \dots & a_{2,k+1} \\ \vdots & \vdots & \ddots & \vdots \\ a_{k+1,1} & a_{k+1,2} & \dots & a_{k+1,k+1} \end{bmatrix}, \quad \mathbf{B} = \begin{bmatrix} b_{11} & b_{12} & \dots & b_{1,k+1} \\ b_{21} & b_{22} & \dots & b_{2,k+1} \\ \vdots & \vdots & \ddots & \vdots \\ b_{k+1,1} & b_{k+1,2} & \dots & b_{k+1,k+1} \end{bmatrix},$$

$$|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|.$$

Using Laplace expansion on the first row of the product, we have

$$|\mathbf{AB}| = \sum_{j=1}^{k+1} (-1)^{1+j} (\mathbf{AB})_{1j} \mathbf{M}_{1j}(\mathbf{AB})$$

Notice that, every  $(\mathbf{AB})_{1j}$  is the product of the first row of  $\mathbf{A}$  and the  $j$ -th column of  $\mathbf{B}$ :

$$(\mathbf{AB})_{1j} = \sum_{m=1}^{k+1} A_{1m} B_{mj}. \quad (5)$$

It is also observed that  $\mathbf{M}_{1j}(\mathbf{AB})$  is a determinant of  $k \times k$  matrix, where our inductive hypothesis is applicable.

So, applying our hypothesis and (5), we have

$$|\mathbf{AB}| = \sum_{j=1}^{k+1} (-1)^{1+j} \left( \sum_{m=1}^{k+1} \mathbf{A}_{1m} \mathbf{B}_{mj} \right) \mathbf{M}_{1j}(\mathbf{A}) \mathbf{M}_{1j}(\mathbf{B}) \quad (6)$$

By the property of double summation, we have

$$|\mathbf{AB}| = \sum_{j=1}^{k+1} \sum_{m=1}^{k+1} (-1)^{1+j} \mathbf{A}_{1m} \mathbf{B}_{mj} \mathbf{M}_{1j}(\mathbf{A}) \mathbf{M}_{1j}(\mathbf{B})$$

We can again, rearrange the summation, because in this case the subscript  $j$  and  $m$  have the same lower and upper boundaries. We can split the double summation in to the product of two  $k + 1$  time summations, and change the subscript variable without changing the result.

$$|\mathbf{AB}| = \sum_{j=1}^{k+1} (-1)^{1+j} \mathbf{A}_{1j} \mathbf{M}_{1j}(\mathbf{A}) \sum_{m=1}^{k+1} (-1)^{1+m} \mathbf{B}_{1m} \mathbf{M}_{1m}(\mathbf{B})$$

Note that

$$\begin{aligned} |\mathbf{A}| &= \sum_{j=1}^{k+1} (-1)^{1+j} \mathbf{A}_{1j} \mathbf{M}_{1j}(\mathbf{A}), \\ |\mathbf{B}| &= \sum_{m=1}^{k+1} (-1)^{1+m} \mathbf{B}_{1m} \mathbf{M}_{1m}(\mathbf{B}). \end{aligned}$$

Hence, we have shown that for any two  $k + 1 \times k + 1$  matrix,  $|\mathbf{AB}| = |\mathbf{A}||\mathbf{B}|$ , and thus  $P(k + 1)$  holds.

To sum up, by induction, we have

$$P(1) \implies P(2) \implies \dots \implies P(n),$$

and the statement is thus proven. □

## Problem 1

**Solution:** By substituting  $(1, 6), (2, 7), (3, 5)$  to

$$y(x) = Ax^2 + Bx + C, \quad (7)$$

we have

$$\begin{cases} A + B + C = 6 \\ 4A + 2B + C = 7 \\ 9A + 3B + C = 5 \end{cases}.$$

In matrix, that is:

$$\begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 6 \\ 7 \\ 5 \end{bmatrix}.$$

## Problem 2

**Solution:** Rewrite the system into coefficient matrix  $\mathbf{A}$ , we have

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{bmatrix}. \quad (8)$$

We use the Laplace expansion to calculate the determinant.

**Definition 1:** For an  $n \times n$  matrix  $\mathbf{B}$ , the determinant  $\det(\mathbf{B})$  can be expanded along any row or column. The formula is:

$$\det(\mathbf{B}) = \sum_{j=1}^n (-1)^{i+j} b_{ij} \det(\mathbf{B}_{ij})$$

where  $b_{ij}$  is the element of matrix  $\mathbf{B}$  at row  $i$  and column  $j$ , and  $\mathbf{B}_{ij}$  is the  $(n-1) \times (n-1)$  submatrix obtained by deleting the  $i$ -th row and  $j$ -th column of  $\mathbf{B}$ .

$\det(\mathbf{A}) =$  We choose to expand along the first row of the matrix  $\mathbf{A}$ :

$$\mathbf{A} = \begin{bmatrix} 1 & 1 & 1 \\ 4 & 2 & 1 \\ 9 & 3 & 1 \end{bmatrix}.$$

Calculating the determinant  $\det(\mathbf{A})$ :

$$\det(\mathbf{A}) = 1 \cdot \begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix} - 1 \cdot \begin{vmatrix} 4 & 1 \\ 9 & 1 \end{vmatrix} + 1 \cdot \begin{vmatrix} 4 & 2 \\ 9 & 3 \end{vmatrix}$$

where,

$$\begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix} = 2 \cdot 1 - 1 \cdot 3 = 2 - 3 = -1$$

$$\begin{vmatrix} 4 & 1 \\ 9 & 1 \end{vmatrix} = 4 \cdot 1 - 1 \cdot 9 = 4 - 9 = -5$$

$$\begin{vmatrix} 4 & 2 \\ 9 & 3 \end{vmatrix} = 4 \cdot 3 - 2 \cdot 9 = 12 - 18 = -6$$

Substituting these into the formula:

$$\det(\mathbf{A}) = 1 \cdot (-1) - 1 \cdot (-5) + 1 \cdot (-6) = -1 + 5 - 6 = -2$$

Therefore, the determinant of matrix  $\mathbf{A}$  is  $\det(\mathbf{A}) = -2$ .

This tells us that the system is solvable and has specific solution, so we can make sure we will find a unique curve to fit the trajectory.

## Problem 3

**Solution:** By lemma 1, we have

$$\mathbf{A}^{-1} = \frac{\text{adj}(\mathbf{A})}{|\mathbf{A}|}.$$

So to get  $\mathbf{A}^{-1}$ , we only need to find  $\text{adj}(\mathbf{A})$ .

We know that  $\text{adj}(\mathbf{A}) = \mathbf{C}^T$ , where  $\mathbf{C}$  is the cofactor matrix of  $\mathbf{A}$ . Each element in  $\mathbf{C}$  is given by

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

where  $M_{ij}$  is the determinant of the  $(n-1) \times (n-1)$  submatrix obtained by deleting the  $i$ -th row and  $j$ -th column of  $\mathbf{A}$ .

After calculation, we have

$$\text{adj}(\mathbf{A}) = \begin{pmatrix} \begin{vmatrix} 2 & 1 \\ 3 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 2 & 1 \end{vmatrix} \\ \begin{vmatrix} 1 & 4 \\ 1 & 9 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 9 & 1 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 1 & 4 \end{vmatrix} \\ \begin{vmatrix} 4 & 2 \\ 9 & 3 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 3 & 9 \end{vmatrix} & \begin{vmatrix} 1 & 1 \\ 4 & 2 \end{vmatrix} \end{pmatrix},$$

and thus, we have

$$\mathbf{A}^{-1} = -\frac{1}{2} \begin{bmatrix} -1 & 2 & -1 \\ 5 & -8 & 3 \\ -6 & 6 & -2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & -1 & \frac{1}{2} \\ -\frac{5}{2} & 4 & -\frac{3}{2} \\ 3 & -3 & 1 \end{bmatrix}.$$

**Remark:** An alternative way here is Gauss-Jordan Elimination. We can apply Gaussian elimination with an identity matrix simultaneously, and when we have the matrix on the left an identity matrix, we will have its inverse matrix on the right.

$$\left[ \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 4 & 2 & 1 & 0 & 1 & 0 \\ 9 & 3 & 1 & 0 & 0 & 1 \end{array} \right]$$

$$R_2 \leftarrow R_2 - 4 \times R_1$$

$$\left[ \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -2 & -3 & -4 & 1 & 0 \\ 9 & 3 & 1 & 0 & 0 & 1 \end{array} \right]$$

$$R_3 \leftarrow R_3 - 9 \times R_1$$

$$\left[ \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -2 & -3 & -4 & 1 & 0 \\ 0 & -6 & -8 & -9 & 0 & 1 \end{array} \right]$$

$$R_2 \leftrightarrow R_3$$

$$\left[ \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -6 & -8 & -9 & 0 & 1 \\ 0 & -2 & -3 & -4 & 1 & 0 \end{array} \right]$$

$$R_3 \leftarrow R_3 - \frac{1}{3}R_2$$

$$\left[ \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -6 & -8 & -9 & 0 & 1 \\ 0 & 0 & -\frac{1}{3} & -1 & 1 & -\frac{1}{3} \end{array} \right]$$

$$R_3 \leftarrow -3 \times R_3$$

$$\left[ \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -6 & -8 & -9 & 0 & 1 \\ 0 & 0 & 1 & 3 & -3 & 1 \end{array} \right]$$

$$R_2 \leftarrow R_2 + 8 \times R_3$$

$$\left[ \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & -6 & 0 & 15 & -24 & 9 \\ 0 & 0 & 1 & 3 & -3 & 1 \end{array} \right]$$

$$R_2 \leftarrow -\frac{1}{6} \times R_2$$

$$\left[ \begin{array}{ccc|ccc} 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & -\frac{5}{2} & 4 & -\frac{3}{2} \\ 0 & 0 & 1 & 3 & -3 & 1 \end{array} \right]$$

$$R_1 \leftarrow R_1 - R_2 - R_3$$

$$\left[ \begin{array}{ccc|ccc} 1 & 0 & 0 & \frac{1}{2} & -1 & \frac{1}{2} \\ 0 & 1 & 0 & -\frac{5}{2} & 4 & -\frac{3}{2} \\ 0 & 0 & 1 & 3 & -3 & 1 \end{array} \right]$$

$$\text{So } \mathbf{A}^{-1} = \begin{bmatrix} \frac{1}{2} & -1 & \frac{1}{2} \\ -\frac{5}{2} & 4 & -\frac{3}{2} \\ 3 & -3 & 1 \end{bmatrix}.$$

The system can be written as

$$\mathbf{A} \cdot \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} 6 \\ 7 \\ 5 \end{bmatrix}$$

Since  $\mathbf{A}^{-1}$  has the same dimension as  $\mathbf{A}$ , we can left multiply  $\mathbf{A}^{-1}$  on both sides, so we have

$$\mathbf{A}^{-1} \cdot \mathbf{A} \cdot \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \mathbf{A}^{-1} \cdot \begin{bmatrix} 6 \\ 7 \\ 5 \end{bmatrix} \implies \begin{bmatrix} A \\ B \\ C \end{bmatrix} = \mathbf{A}^{-1} \cdot \begin{bmatrix} 6 \\ 7 \\ 5 \end{bmatrix}$$

So,

$$\begin{bmatrix} A \\ B \\ C \end{bmatrix} = \begin{bmatrix} \frac{1}{2} \times 6 - 7 + \frac{1}{2} \times 5 \\ -\frac{5}{2} \times 6 + 4 \times 7 - \frac{3}{2} \times 5 \\ 3 \times 6 - 3 \times 7 + 5 \end{bmatrix} = \begin{bmatrix} -\frac{3}{2} \\ \frac{11}{2} \\ 2 \end{bmatrix}.$$

Hence, the solution is  $\begin{cases} A = -\frac{3}{2} \\ B = \frac{11}{2} \\ C = 2 \end{cases}$

## Problem 4

Substitute  $\begin{cases} A = -\frac{3}{2} \\ B = \frac{11}{2} \\ C = 2 \end{cases}$  to (7), we have

$$y(x) = -\frac{3}{2}x^2 + \frac{11}{2}x + 2$$

The missile lands when  $y(x) = 0$ . To get the position when  $y(x) = 0$ , we only need to solve the quadratic equation:

$$-\frac{3}{2}x^2 + \frac{11}{2}x + 2 = 0 \equiv -3x^2 + 11x + 4 = 0$$

This can be factorised to

$$(x - 4)(-3x - 1) = 0$$

Hence,  $x = 4$  or  $x = -\frac{1}{3}$ , the latter does not make sense here, since the the missile is launched at where  $x = 0$ . So, it lands when it is 4 units away from the origin.

## Problem 5

**Solution:** Yes, we can find a unique trajectory function for the object.

This is because it requires only four unknowns to identify the function, and we are provided four points on the function.

If we substitute these points into the expression respectively, we will get a system of linear equation, whose coefficient matrices satisfies all row vectors are independent from each other. So, the augmented matrix  $(A|b)$  satisfies  $\text{rank}(A) = \text{rank}(b) = n$ , where  $n$  is the number of unknowns.

Therefore, by Rouché-Capelli Theorem, we can find a specific solution for the system of linear equations, so that a unique cubic function can be identified.

## problem 6

**Solution:** By substituting the points into the relation, we have

$$\mathbf{D} = \begin{cases} A + B + C + D = 6 \\ 8A + 4B + 2C + D = 7 \\ 27A + 9B + 3C + D = 5 \\ 64A + 16B + 4C + D = 4 \end{cases} \quad (9)$$



In matrix, we have

$$\begin{bmatrix} 1 & 1 & 1 & 1 \\ 8 & 4 & 2 & 1 \\ 27 & 9 & 3 & 1 \\ 64 & 16 & 4 & 1 \end{bmatrix} \begin{bmatrix} A \\ B \\ C \\ D \end{bmatrix} = \begin{bmatrix} 6 \\ 7 \\ 5 \\ 4 \end{bmatrix}.$$

## Problem 7

This matrix has a very interesting property, which could help us simplify the calculation of  $|\mathbf{D}|$ . We can exchange column 1 and column 4, and then exchange column 2 and column 3. This will not change the determinant, since we exchanged it twice,  $-(-|\mathbf{D}|) = |\mathbf{D}|$ .

The exchange makes

$$\mathbf{D} = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 1 & 2 & 4 & 8 \\ 1 & 3 & 9 & 27 \\ 1 & 4 & 16 & 64 \end{bmatrix}.$$

This allows us to identify the matrix as a Vandermonde matrix, whose determinant can be obtained by  $\prod_{0 \leq i < j \leq n} (x_j - x_i)$ , and  $n = 4$  in this case.

So we have

$$\det(\mathbf{D}) = \prod_{0 \leq i < j \leq 4} (x_j - x_i) = (4 - 3) \times (4 - 2) \times (4 - 1) \times (3 - 2) \times (3 - 1) \times (2 - 1) = 12.$$

This means we can find a unique function through the four data points, since it's a non-singular matrix,  $\det(\mathbf{D}) \neq 0$ .

**Remark:** Here we attach the proof of the determinant (Kalman, 1984; Rushanan, 1989).

**Definition 2** (Vandermonde Matrix): A Vandermonde matrix is an  $(m + 1) \times (n + 1)$  matrix  $\mathbf{V}$  such that:

$$\mathbf{V} = \begin{bmatrix} 1 & x_0 & x_0^2 & \dots & x_0^n \\ 1 & x_1 & x_1^2 & \dots & x_1^n \\ 1 & x_2 & x_2^2 & \dots & x_2^n \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_m & x_m^2 & \dots & x_m^n \end{bmatrix}.$$

The Vandermonde matrix is widely used in the interpolation of polynomials, and in this case of finding coefficients, we must have a square matrix. The following rule is very useful in finding the determinant of a Vandermonde matrix.

**Corollary 1:** When  $m = n$ , the Vandermonde matrix is square, and its determinant has a special form. The determinant of a square Vandermonde matrix  $\mathbf{V}$  of size  $(n + 1) \times (n + 1)$  is given by:

$$\det(\mathbf{V}) = \prod_{0 \leq i < j \leq n} (x_j - x_i).$$

*Proof.* By iterating through each column of the matrix, from right to left, until we reach the first column (we will not do anything to it). We will subtract the value on the previous column (on the left) multiplied by  $x_0$ , from the current elements. This will make the first row (except the first entry) 0, and the rest of the rows (each) will have common factor  $(x_i - x_0)$ ,  $1 \leq i \leq n$ .

$$\mathbf{V} = \begin{bmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ 1 & x_1 - x_0 & x_1(x_1 - x_0) & x_1^2(x_1 - x_0) & \cdots & x_1^{n-1}(x_1 - x_0) \\ 1 & x_2 - x_0 & x_2(x_2 - x_0) & x_2^2(x_2 - x_0) & \cdots & x_2^{n-1}(x_2 - x_0) \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n - x_0 & x_n(x_n - x_0) & x_n^2(x_n - x_0) & \cdots & x_n^{n-1}(x_n - x_0) \end{bmatrix}$$

If we measure  $\det(\mathbf{V})$  using Laplace expansion, we will find that  $\det(\mathbf{V}) = \det(\mathbf{M})$ , where

$$\mathbf{M} = \begin{bmatrix} x_1 - x_0 & x_1(x_1 - x_0) & x_1^2(x_1 - x_0) & \cdots & x_1^{n-1}(x_1 - x_0) \\ x_2 - x_0 & x_2(x_2 - x_0) & x_2^2(x_2 - x_0) & \cdots & x_2^{n-1}(x_2 - x_0) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ x_n - x_0 & x_n(x_n - x_0) & x_n^2(x_n - x_0) & \cdots & x_n^{n-1}(x_n - x_0) \end{bmatrix}.$$

Since each row has common factors, we can just take it outside as a scalar when we deal with its determinant. So we have,

$$\begin{aligned} \det(\mathbf{V}) &= (x_1 - x_0)(x_2 - x_0) \cdots (x_n - x_0) \begin{vmatrix} 1 & x_1 & x_1^2 & \cdots & x_1^{n-1} \\ 1 & x_2 & x_2^2 & \cdots & x_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_n & x_n^2 & \cdots & x_n^{n-1} \end{vmatrix} \\ &= \prod_{1 \leq i \leq n} (x_i - x_0) \det(\mathbf{V}') \end{aligned}$$

It is observed that  $\mathbf{V}'$  is also a Vandermonde matrix, so we can apply the same algorithm again.

The execution of the algorithm comes to an end when  $\mathbf{V}'$  is a  $1 \times 1$  square matrix with 1 as the only entry. Then, we will get all possible permutations of differences between elements in column two  $x_i, x_j$ , where  $j > i$ .

Hence, we can deduce that

$$\det(\mathbf{V}) = \prod_{0 \leq i < j \leq n} (x_j - x_i).$$

□

---

**Note**

The proof maybe not very exhaustive and rigorous, but is enough to show how its determinant works.

## Problem 8

**Solution:** Use the following wolfram code,

```
1 m = {{1, 1, 1, 1}, {8, 4, 2, 1}, {27, 9, 3, 1}, {64, 16, 4,
    1}};
2 b = {6, 7, 5, 4};
3 x = LinearSolve[m, b]
```

we get the solution for (9):

$$A = \frac{2}{3}, \quad B = -\frac{11}{2}, \quad C = \frac{77}{6}, \quad D = -2.$$

## Problem 9

**Solution:** By substituting the solutions into the polynomial, we have

$$y(x) = \frac{2}{3}x^3 - \frac{11}{2}x^2 + \frac{77}{6}x - 2.$$

We draw  $y(x)$  using tikz and pgfplots below.

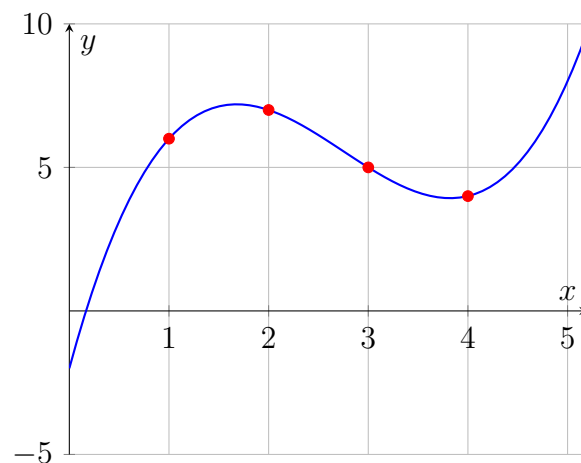


Figure 1:  $y(x) = \frac{2}{3}x^3 - \frac{11}{2}x^2 + \frac{77}{6}x - 2$

## Problem 10

**Solution:** The missile will never land. If we assume that this missile will must hit its target at some point  $(x, 0)$ , then this will not happen. Because the value of  $y(x)$  keeps increasing after it reaches its second stationary point  $(4, 4)$ .

## Problem 11

**Solution:** The three points are not colinear. If we suppose they are colinear, then any two direction vector between the any of two points chosen will be the same. However,  $(2, 7) - (1, 6) = (1, 1)$ , while  $(3, 5) - (2, 7) = (1, -2)$ . Hence do not lie on a straight line.

## Problem 12

By substituting the predicted values into the relation, we have

**Solution:**

$$\begin{cases} S + T = y(1) \\ 2S + T = y(2) \\ 3S + T = y(3) \end{cases}.$$

## Problem 13

**Solution:** The system can be written as:

$$\begin{bmatrix} y(1) \\ y(2) \\ y(3) \end{bmatrix} = S \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + T \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

To make it clearer, we can add  $\mathbf{0}$  to the RHS.

$$\begin{bmatrix} y(1) \\ y(2) \\ y(3) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} + S \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + T \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$$

We first focus on the RHS. The expression on the right can be interpreted as creating a new plane from the origin in  $\mathbb{R}^3$  using two column vectors scaled by some parameter  $S$  and  $T$  respectively. From the size of the column vector, we know they are vectors in  $\mathbb{R}^3$ .

From the LHS, we know that the result of the linear transformation on the right will create a new point in  $\mathbb{R}^3$ , combining this to what we derived earlier, we can confirm that

$$\exists \alpha \subseteq \mathbb{R}^3, \forall S, T \in \mathbb{R}, \begin{bmatrix} S + T \\ 2S + T \\ 3S + T \end{bmatrix} \in \alpha.$$

## Problem 14

**Solution:** To find the minimum distance, between  $(6, 7, 5)$  and the plane, we simply need to find the projection of some direction vector between a point on the plane and the actual coordinate on the normal to the plane.

We already know that  $\begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$  and  $\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}$  are two direction vectors on the solution plane, so we can find a normal vector to the plane (call  $\alpha$ ):

$$\vec{n} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \times \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 2 \\ -1 \end{bmatrix}$$

From earlier, we know that  $\begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} \in \alpha$ , so we have a direction vector from the plane to actual coordinate, we call  $\vec{v}$ :

$$\vec{v} = \begin{bmatrix} 6 \\ 7 \\ 5 \end{bmatrix} - \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix} = \begin{bmatrix} 6 \\ 7 \\ 5 \end{bmatrix}$$

Now we can calculate the minimum distance  $d$  as projection of  $\vec{v}$  on  $\vec{n}$ .

$$d = \frac{|\vec{v} \cdot \vec{n}|}{|\vec{n}|} = \frac{|6 \times (-1) + 7 \times 2 + 5 \times (-1)|}{\sqrt{6}} = \frac{|-6 + 14 - 5|}{\sqrt{6}} = \frac{\sqrt{6}}{2}$$

## Problem 15

**Solution:** We can find the optimised line by least squares method.

$$S = \frac{N \sum xy - \sum x \sum y}{N \sum x^2 - (\sum x)^2}, \quad T = \frac{\sum y - S \sum x}{N}$$

We have

$$S = \frac{3 \times 35 - 6 \times 18}{3 \times 14 - 6^2} = \frac{105 - 108}{42 - 36} = \frac{-3}{6} = -\frac{1}{2}$$

and

$$T = \frac{18 - (-\frac{1}{2}) \times 6}{3} = \frac{18 + 3}{3} = \frac{21}{3} = 7$$

So the optimised line is  $f(x) = -\frac{1}{2}x + 7$ .

**Remark:** We may show how it works (only for the linear function, because I do not know much about vector calculus, hope I'll learn that soon in ENG2005).

The idea here is basically to minimise the difference between the predicted value by the model (the function here).

If we compare the difference between our prediction and real value for the same  $x$  one by one, we can define something called the loss (function) of the model, like

$$loss(S, T) = \sum_{i=1}^n (y_i - f(x)) = \sum_{i=1}^n [y_i - (Sx_i + T)]$$

where

- $n$  is the number of samples.
- $x_i, y_i$  are the  $i$ th  $x$  value and actual value respectively.
- $S, T$  are the parameters introduced in the problem.

But in real practice, we tend to square the difference, as sum up everything directly can sometimes cause confusion, for example, when we have some negative and positive terms that offset each other. Also we are doing this to make sure all errors can be measured by a positive value. Additionally, we use the mean of the squared sum to measure the disparity to make sure each sample has equal contribution to the total

loss (Zhou, 2021). Therefore, we define the most common mean square loss function, in this case, as

$$\text{loss}(S, T) = \frac{1}{n} \sum_{i=1}^n [y_i - (Sx_i + T)]^2.$$

The principal goal of linear regression is basically minimise the value of the function. But we can just rule out the constant coefficient before the summation, as it does not affect the optimisation, since  $\frac{1}{n} > 0$ . So actually, we only need to minimise

$$g(S, T) = \sum_{i=1}^n [y_i - (Sx_i + T)]^2.$$

This makes it clear its geometrical meaning. We see that this function add up the squared Euclidean distances between each predicted data and actual data.

We can get  $\min(g(S, T))$  by taking partial derivatives on  $S$  and  $T$  respectively, since this function is a binary quadratic function, meaning that we have an extreme value when the value of derivative functions equals zero.

First, we take the partial derivative of  $g(S, T)$  with respect to  $T$ :

$$\frac{\partial g(S, T)}{\partial T} = \frac{\partial}{\partial T} \sum_{i=1}^n [y_i - Sx_i - T]^2$$

Expanding using the chain rule:

$$\frac{\partial g(S, T)}{\partial T} = -2 \sum_{i=1}^n [y_i - Sx_i - T]$$

Setting the derivative to zero gives:

$$\sum_{i=1}^n [y_i - Sx_i - T] = 0$$

which can be rearranged to:

$$\sum_{i=1}^n y_i = S \sum_{i=1}^n x_i + nT$$

Thus, the intercept  $T$  is:

$$T = \frac{\sum_{i=1}^n y_i - S \sum_{i=1}^n x_i}{n}$$

Next, we take the partial derivative of  $g(S, T)$  with respect to  $S$ :

$$\frac{\partial g(S, T)}{\partial S} = \frac{\partial}{\partial S} \sum_{i=1}^n [y_i - Sx_i - T]^2$$

Again, expanding using the chain rule:

$$\frac{\partial g(S, T)}{\partial S} = -2 \sum_{i=1}^n x_i [y_i - Sx_i - T]$$

Setting the derivative to zero gives:

$$\sum_{i=1}^n x_i y_i = S \sum_{i=1}^n x_i^2 + T \sum_{i=1}^n x_i$$

Substituting the expression for  $T$  from the previous step, and simplifying, we obtain the formula for the slope  $S$ :

$$S = \frac{n \sum_{i=1}^n x_i y_i - \sum_{i=1}^n x_i \sum_{i=1}^n y_i}{n \sum_{i=1}^n x_i^2 - (\sum_{i=1}^n x_i)^2}$$

So that's how we derive the formula for this basic case.

## References

- Kalman, D. (1984). The Generalized Vandermonde Matrix. *Mathematics Magazine*, 57(1), 15–21. <https://doi.org/10.1080/0025570X.1984.11977069>
- Rushanan, J. J. (1989). On the Vandermonde Matrix. *The American Mathematical Monthly*, 96(10), 921–924. <https://doi.org/10.1080/00029890.1989.11972306>
- Zhou, Z.-H. (2021). *Machine learning*. Springer. [https://monash.primo.exlibrisgroup.com/discovery/fulldisplay?docid=alma9940279982801751&context=L&vid=61MONASH\\_AU:MUMUI&lang=en&search\\_scope=MalaysiaAll&adaptor=Local%20Search%20Engine&tab=MalaysiaLocalCDIs&query=any,contains,Machine%20Learning%20\(Zhou\)&offset=0](https://monash.primo.exlibrisgroup.com/discovery/fulldisplay?docid=alma9940279982801751&context=L&vid=61MONASH_AU:MUMUI&lang=en&search_scope=MalaysiaAll&adaptor=Local%20Search%20Engine&tab=MalaysiaLocalCDIs&query=any,contains,Machine%20Learning%20(Zhou)&offset=0)