Chapter 1 Review Notes

MA107A Linear Algebra A Fall 2021

by Zhang Ce

Something to say:
You are encouraged to review the slides given by your instructor first, since it's more complete and authentic. You are also supposed to work with the exam problems in previous years to get familiar with the problem types. This review note is like a summary and only covers part of knowledge and you can review this material just before the exam to have a quick review and self-test. Hope you all do well in the midterm exam, good luck!

1 Introduction

Summary: Chapter 1.1 is an introduction to the whole chapter, it defines the linear equations, which is the core of linear algebra, and shows us how to transform the linear equation systems into augmented matrix form.

1.1 Linear Equation

A linear equation in n unknowns is an equation of the form

$$a_1x_1 + a_2x_2 + \cdots + a_nx_n = b$$

where $a_1, a_2, \cdots, a_n, b$ are all real numbers.
The equations such as $x^3 + y = 2, \sqrt{x} + cosy = xy$ is not linear.

1.2 Matrix Notation

We can use a coefficient matrix to record all the coefficients in linear system, and we can use an argumented matrix to represent a linear equation system, for example:

$$\begin{cases} x + y = 8 \\ -x + 4y = 1 \end{cases} \iff \begin{bmatrix} 1 & 1 & 8 \\ -1 & 4 & 1 \end{bmatrix}$$

2 The Geometry of Linear Equations

Summary: Chapter 1.2 leads us to understand the linear equations in 2 ways from geometrical perspective: row picture and column picture. You should have an accurate imagination of the row picture and column picture in higher-order linear equation systems after this chapter.

2.1 Row Picture

Given an equation system:

$$\begin{cases} 2x - y = 0 \\ -x + 2y = 3 \end{cases}$$

In geometry, every single row in this system represents a straight line. For the linear equation system above, the row picture is given in Figure 1.

Then, in order to find the solution to this linear system, we only need to find the intersection of the 2 lines, which gives us the solution $x = 1, y = 2$.

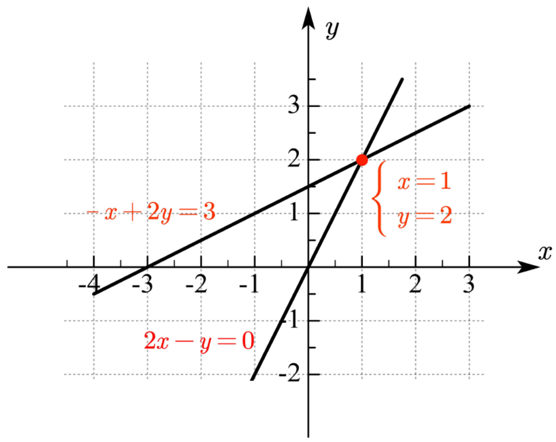


Figure 1: Row picture

2.2 Column Picture

For the linear equation system above, we can transform it into equivalent matrix form:

$$\begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

We can have a further transformation:

$$x \begin{bmatrix} 2 \\ -1 \end{bmatrix} + y \begin{bmatrix} -1 \\ 2 \end{bmatrix} = \begin{bmatrix} 0 \\ 3 \end{bmatrix}$$

Every column (coefficients of a single variable) represents a vector, and the problem is to find how to combine this 2 vectors to get the left side vector b. The column picture is given in Figure 2.

One (column 1) and 2 (column 2) give us the left side vector b, so the solution is $x = 1, y = 2$.

2.3 Higher-Order Imagination

Suppose we have a linear equation system with 4 unknown and 4 equations, the row picture will include 4 three-dimensional space and the problem is to find the intersection, while the column picture have 4 four-dimensional column vectors and the problem is to find a combination to get the right-hand side b.

3 Gaussian Elimination

Summary: Chapter 1.3 introduces a complete algorithm for solving linear system. The general process is to do a series of row operations to get an upper-triangular form, and then we can solve the system by back-substitution.

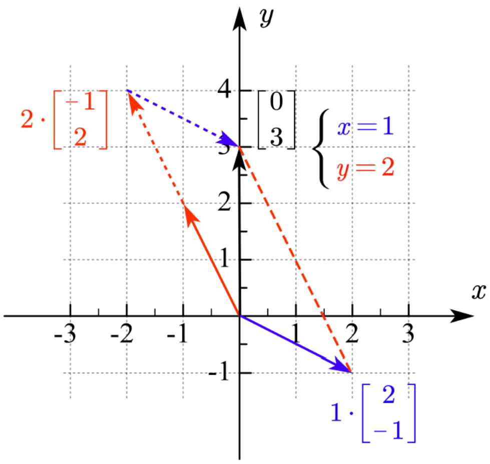


Figure 2: Column picture

3.1 Gaussian Elimination Process

Consider the following system of linear equations:

$$\begin{cases} x + 2y + z = 2 \\ 3x + 8y + z = 12 \\ 4y + z = 2 \end{cases}$$

Augmented matrix:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 8 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix}$$

Now, we eliminate some entries to simplify the equation system. Remember that you can only do row operations.

Eliminate entry on position $(2, 1)$:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 4 & 1 & 2 \end{bmatrix}$$

Eliminate entry on position $(3, 1)$:

The entry is 0, skip this step.

Eliminate entry on position $(3, 2)$:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 0 & 5 & -10 \end{bmatrix}$$

The order is important and you are encouraged to eliminate the entries in the first column, then the second and the following columns.

The Gaussian Elimination process ends here, the system is now easy enough and we can solve it by back-substitution.

3.2 Back-Substitution

After elimination, we get

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 2 & -2 & 6 \\ 0 & 0 & 5 & -10 \end{bmatrix}$$

which is the same with the following linear system

$$\begin{cases} x + 2y + z = 2 \\ 2y - 2z = 6 \\ 5z = -10 \end{cases}$$

Do back-substitution, we can get the solution:

$$\begin{cases} x = 2 \\ y = 1 \\ z = -2 \end{cases}$$

Problem 1. Solve the following system of linear equations by row reduction.

$$\begin{cases} 2x + 3y + z = 8 \\ 4x + 7y + 5z = 20 \\ -2y + 2z = 0 \end{cases}$$

Solution:
Do Gaussian Elimination firstly:

$$\begin{bmatrix} 2 & 3 & 1 & 8 \\ 4 & 7 & 5 & 20 \\ 0 & -2 & 2 & 0 \end{bmatrix} \to \begin{bmatrix} 2 & 3 & 1 & 8 \\ 0 & 1 & 3 & 4 \\ 0 & -2 & 2 & 0 \end{bmatrix} \to \begin{bmatrix} 2 & 3 & 1 & 8 \\ 0 & 1 & 3 & 4 \\ 0 & 0 & 8 & 8 \end{bmatrix}$$

Then we can solve by back-substitution, thesolution is $x = 2, y = 1, z = 1$.

3.3 Singular Cases for Gauss Elimination

3.3.1 Temporal Failure

Consider this linear system:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 3 & 6 & 1 & 12 \\ 0 & 4 & 1 & 2 \end{bmatrix}$$

Eliminate entry on position $(2, 1)$:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 0 & -2 & 6 \\ 0 & 4 & 1 & 2 \end{bmatrix}$$

Eliminate entry on position $(3, 1)$:

The entry is 0, skip this step.

Then the entry on position $(2,2)$ is 0, so we have no chance to eliminate the entry on position $(3,2)$. This is called the temporal failure of Gauss Elimination. Because we can find nonzero entry below, so we can fix it by row exchanges. In this situation, it still has only one solution.

3.3.2 Permanent Failure

Consider this linear system:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 4 & 1 & 12 \\ 3 & 6 & 3 & 2 \end{bmatrix}$$

After Gauss Elimination:

$$\begin{bmatrix} 1 & 2 & 1 & 2 \\ 0 & 4 & 1 & 12 \\ 0 & 0 & 0 & -4 \end{bmatrix}$$

Then the entry on position $(3,3)$ is 0, and there is no nonzero entry below. This is called the permanent failure of Gauss Elimination. In this situation, it has no solution or infinitely many solutions. The last row gives $0 = -4$, making this system have no solution. If it's $0 = 0$, it will give infinite solutions.

Problem 2. If the following linear system has no solution, find a.

$$\begin{bmatrix} 1 & 2 & 1 \\ 2 & 3 & a+2 \\ 1 & a & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 3 \\ 0 \end{bmatrix}$$

Solution:

$$\begin{bmatrix} 1 & 2 & 1 & 1 \\ 2 & 3 & a+2 & 3 \\ 1 & a & -2 & 0 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & -1 & a & 1 \\ 0 & a-2 & -3 & -1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 1 & 1 \\ 0 & -1 & a & 1 \\ 0 & 0 & a^2-2a-3 & a-3 \end{bmatrix}$$

Let $a^2 - 2a - 3 = 0$ and $a - 3 \neq 0$, we can get $a = -1$.

4 Matrix Multiplication

Summary: Chapter 1.4 defines one of the most important matrix arithmetic: matrix multiplication. You are encouraged to learn 4 perspectives of matrix multiplication in this section.

4.1 Matrix Size in Matrix Multiplication

Suppose we have a matrix A with size $m \times n$, and a matrix B with size $n \times p$, then we can find AB with size $m \times p$. For example, consider the following matrix multiplication:

$$\begin{bmatrix}1 & 0 & 0\end{bmatrix}\begin{bmatrix}4 & 2021\\11 & 2\\6 & 4\end{bmatrix}$$

Without computation, the result will be a 1×2 matrix since the left matrix is 1×3 and the right is 3×2.

4.2 Four Methods to Understand Matrix Multiplication

We use an example of matrix multiplication to show the methods:

$$\begin{bmatrix}3 & -1 & 1\\1 & 0 & 1\\0 & 1 & 2\end{bmatrix}\begin{bmatrix}1 & 0 & 1\\1 & -1 & -1\\1 & 2 & 1\end{bmatrix} = \begin{bmatrix}3 & 3 & 5\\2 & 2 & 2\\3 & 3 & 1\end{bmatrix}$$

4.2.1 Method 1: The Regular Way (Row-Col Way)

To calculate one entry in the result, the regular way is to multiply the row in the left matrix and the column in the right matrix. For the example above, if we want to find the $(1, 1)$ entry in the result matrix, we need to multiply the first row of A and the first column of B.

$$\begin{bmatrix}\color{red}{3} & \color{red}{-1} & \color{red}{1}\\1 & 0 & 1\\0 & 1 & 2\end{bmatrix}\begin{bmatrix}\color{red}{1} & 0 & 1\\\color{red}{1} & -1 & -1\\\color{red}{1} & 2 & 1\end{bmatrix} = \begin{bmatrix}\color{red}{3} & 3 & 5\\2 & 2 & 2\\3 & 3 & 1\end{bmatrix}$$

The result comes from $3 \cdot 1 + (-1) \cdot 1 + 1 \cdot 1 = 3$. As you can see, we need to repeat for a total of 9 times, which is really complicated.

4.2.2 Method 2: The Row Way

For this method, we can compute a whole row in the result at once. The process is, take a row from A and see it as a linear combination of rows of B. For example, still for the same multiplication:

$$\begin{bmatrix}\color{red}{3} & \color{red}{-1} & \color{red}{1}\\1 & 0 & 1\\0 & 1 & 2\end{bmatrix}\begin{bmatrix}\color{red}{1} & \color{red}{0} & \color{red}{1}\\\color{red}{1} & \color{red}{-1} & \color{red}{-1}\\\color{red}{1} & \color{red}{2} & \color{red}{1}\end{bmatrix} = \begin{bmatrix}\color{red}{3} & \color{red}{3} & \color{red}{5}\\2 & 2 & 2\\3 & 3 & 1\end{bmatrix}$$

The result comes from $3\begin{bmatrix}1 & 0 & 1\end{bmatrix} - 1\begin{bmatrix}1 & -1 & -1\end{bmatrix} + 1\begin{bmatrix}1 & 2 & 1\end{bmatrix} = \begin{bmatrix}3 & 3 & 5\end{bmatrix}$, which is the first row multiply by matrix B. Then we do the same thing for row 2 and row 3. At this time, we need only to repeat 3 times.

4.2.3 Method 3: The Column Way

For this method, similar to the row way, we can compute a whole column in the result at once. The process is, take a column from B and see it as a linear combination of columnss of A. For example, still for the same multiplication:

$$\begin{bmatrix}\color{red}{3} & \color{red}{-1} & \color{red}{1}\\\color{red}{1} & \color{red}{0} & \color{red}{1}\\\color{red}{0} & \color{red}{1} & \color{red}{2}\end{bmatrix}\begin{bmatrix}\color{red}{1} & 0 & 1\\\color{red}{1} & -1 & -1\\\color{red}{1} & 2 & 1\end{bmatrix} = \begin{bmatrix}\color{red}{3} & 3 & 5\\\color{red}{2} & 2 & 2\\\color{red}{3} & 3 & 1\end{bmatrix}$$

The first column of the result is given by $1\begin{bmatrix}3\\1\\0\end{bmatrix}+1\begin{bmatrix}-1\\0\\1\end{bmatrix}+1\begin{bmatrix}1\\1\\2\end{bmatrix}=\begin{bmatrix}3\\2\\3\end{bmatrix}$. We only need to repeat 3 times also.

4.2.4 Method 4: The Col-Row Way

For this method, we do the opposite thing with Method 1. We multiply columns of A and rows of B to calculate the matrix multiplication.

$$\begin{bmatrix}3&-1&1\\1&0&1\\0&1&2\end{bmatrix}\begin{bmatrix}1&0&1\\1&-1&-1\\1&2&1\end{bmatrix}=\begin{bmatrix}3&3&5\\2&2&2\\3&3&1\end{bmatrix}$$

When we multiply a column by a row, we will get a matrix instead of a single entry, so when we take 3 columns of A to multiply 3 rows of B, we can get three 3×3 matrix, then, all we need to do is to add them and find the result.

$$\begin{aligned}&\begin{bmatrix}3\\1\\0\end{bmatrix}\begin{bmatrix}1&0&1\end{bmatrix}+\begin{bmatrix}-1\\0\\1\end{bmatrix}\begin{bmatrix}1&-1&-1\end{bmatrix}+\begin{bmatrix}1\\1\\2\end{bmatrix}\begin{bmatrix}1&2&1\end{bmatrix}\\&=\begin{bmatrix}3&0&3\\1&0&1\\0&0&0\end{bmatrix}+\begin{bmatrix}-1&1&1\\0&0&0\\1&-1&-1\end{bmatrix}+\begin{bmatrix}1&2&1\\1&2&1\\2&4&2\end{bmatrix}=\begin{bmatrix}3&3&5\\2&2&2\\3&3&1\end{bmatrix}\end{aligned}$$

5 Triangular Factors and Row Exchanges

5.1 Elimination Matrices

Now it's time to understand elimination matrices. The goal is to express the elimination process by matrix language.

For example, to eliminate $(2,1)$ position element:

$$\begin{bmatrix}1&2&1\\3&8&1\\0&4&1\end{bmatrix}\rightarrow\begin{bmatrix}1&2&1\\0&2&-2\\0&4&1\end{bmatrix}$$

What we have done: (Row 2) - 3 (Row 1).
Use a elimination matrix E_{21} to represent this process:

$$\begin{bmatrix}1&0&0\\-3&1&0\\0&0&1\end{bmatrix}\begin{bmatrix}1&2&1\\3&8&1\\0&4&1\end{bmatrix}=\begin{bmatrix}1&2&1\\0&2&-2\\0&4&1\end{bmatrix}$$

$$E_{21}A=A'$$

As you might discover, $E_{32}E_{31}E_{21}A=U$.

The process of Gauss Elimination we introduced before:

$$\begin{bmatrix}1 & 2 & 1\\3 & 8 & 1\\0 & 4 & 1\end{bmatrix} \xrightarrow{r2-3r1} \begin{bmatrix}1 & 2 & 1\\0 & 2 & -2\\0 & 4 & 1\end{bmatrix} \rightarrow \begin{bmatrix}1 & 2 & 1\\0 & 2 & -2\\0 & 4 & 1\end{bmatrix} \xrightarrow{r3-2r2} \begin{bmatrix}1 & 2 & 1\\0 & 2 & -2\\0 & 0 & 5\end{bmatrix}$$

Gauss Elimination represented in multipling elimination matrices:

$$\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & -2 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\-3 & 1 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 2 & 1\\3 & 8 & 1\\0 & 4 & 1\end{bmatrix} = \begin{bmatrix}1 & 2 & 1\\0 & 2 & -2\\0 & 0 & 5\end{bmatrix}$$

We define $E = E_{32}E_{31}E_{21}$, then we can get $EA = U$. Finally we can get $A = LU$, $L = E^{-1} = {E_{21}}^{-1}{E_{31}}^{-1}{E_{32}}^{-1}$. Remind that $E_{21}, E_{31}, ..., L$ are all lower triangular matrices.

5.2 A=LU Factorization

A simple comparison of $EA = U$ & $A = LU$ using the example above:

$$E = \begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & -2 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\-3 & 1 & 0\\0 & 0 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\-3 & 1 & 0\\6 & -2 & 1\end{bmatrix}$$

$$L = \begin{bmatrix}1 & 0 & 0\\3 & 1 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 0 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\0 & 1 & 0\\0 & 2 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\3 & 1 & 0\\0 & 2 & 1\end{bmatrix}$$

Computing E is a quite complicated thing, on the contrary, when computing L, we just need to fill the blank and do not need any kinds of matrix multiplication because the entry at different positions will not influence the others.

For the process of Gauss Elimination before:

$$\begin{bmatrix}1 & 2 & 1\\3 & 8 & 1\\0 & 4 & 1\end{bmatrix} \xrightarrow{l_{21}=3} \begin{bmatrix}1 & 2 & 1\\0 & 2 & -2\\0 & 4 & 1\end{bmatrix} \rightarrow \begin{bmatrix}1 & 2 & 1\\0 & 2 & -2\\0 & 4 & 1\end{bmatrix} \xrightarrow{l_{32}=2} \begin{bmatrix}1 & 2 & 1\\0 & 2 & -2\\0 & 0 & 5\end{bmatrix}$$

So the LU factorization of the matrix above is:

$$A = \begin{bmatrix}1 & 2 & 1\\3 & 8 & 1\\0 & 4 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\3 & 1 & 0\\0 & 2 & 1\end{bmatrix}\begin{bmatrix}1 & 2 & 1\\0 & 2 & -2\\0 & 0 & 5\end{bmatrix} = LU$$

If we want to find the LDU factorization, we need to extract the pivots in the diagonal matrix and divide each row by the corresponding pivot.

$$A = \begin{bmatrix}1 & 2 & 1\\3 & 8 & 1\\0 & 4 & 1\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\3 & 1 & 0\\0 & 2 & 1\end{bmatrix}\begin{bmatrix}1 & 2 & 1\\0 & 2 & -2\\0 & 0 & 5\end{bmatrix} = \begin{bmatrix}1 & 0 & 0\\3 & 1 & 0\\0 & 2 & 1\end{bmatrix}\begin{bmatrix}1 & 0 & 0\\0 & 2 & 0\\0 & 0 & 5\end{bmatrix}\begin{bmatrix}1 & 2 & 1\\0 & 1 & -1\\0 & 0 & 1\end{bmatrix} = LDU$$

Problem 3. Do the LU factorization for matrix A:

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 4 \\ 1 & 4 & 8 \end{bmatrix}$$

Solution:

$$\begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 4 \\ 1 & 4 & 8 \end{bmatrix} \xrightarrow{l_{21}=1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 3 \\ 1 & 4 & 8 \end{bmatrix} \xrightarrow{l_{31}=1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 3 \\ 0 & 3 & 7 \end{bmatrix} \xrightarrow{l_{32}=1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 3 \\ 0 & 0 & 4 \end{bmatrix}$$

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 4 & 4 \\ 1 & 4 & 8 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 3 & 3 \\ 0 & 0 & 4 \end{bmatrix} = LU$$

5.3 Permutation Matrices

For permutation matrices, they have only a single one in a row and a column. For a particular permutation matrix P:

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

- For "1" at position $(1, 1)$: (Row 1) $\rightarrow$ (Row 1).
- For "1" at position $(2, 3)$: (Row 3) $\rightarrow$ (Row 2).
- For "1" at position $(3, 2)$: (Row 2) $\rightarrow$ (Row 3).

For those matrices that make Gaussian Elimination temporarily fail, they can not have $A = LU$ factorization, but they can have $PA = LU$ factorization since they can be fixed by row exchanges.

6 Inverses

Summary: Chapter 1.6 defines inverse of a matrix and also teaches us a method to find the inverse of matrix: Gauss-Jordan method. You also need to learn how to determine a matrix is invertible or not.

6.1 Definition of Inverses

An $n \times n$ matrix A is said to be invertible if there is an $n \times n$ matrix B such that

$$AB = BA = I$$

In this case, B is called an inverse of A.

6.2 Existence of Inverses

Suppose there is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = 0$, then A cannot have an inverse.
For this matrix A, find a nonzero solution for $A\mathbf{x} = 0$:

$$\begin{bmatrix} 1 & 2 \\ 3 & 6 \end{bmatrix} \begin{bmatrix} 2 \\ -1 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

Suppose there is an inverse to make $A^{-1}A = I$, multiply A^{-1} on both sides:

$$A\mathbf{x} = 0 \to A^{-1}A\mathbf{x} = A^{-1}0 \to \mathbf{x} = 0$$

Well, but there is a nonzero solution (2,-1) for matrix A! So matrix A cannot have an inverse.

6.3 Calculation of Inverses

The Gauss-Jordan method goes as the following: We put the matrix A at the left and write an identity matrix I at the right, and we do row operations to eliminate the left matrix to I, the right matrix turns to A^{-1}.
For example:

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 7 \end{bmatrix}$$

The Gauss-Jordan method goes as:

$$\begin{bmatrix} 1 & 2 & 1 & 0 \\ 3 & 7 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 1 & 0 \\ 0 & 1 & -3 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 7 & -2 \\ 0 & 1 & -3 & 1 \end{bmatrix}$$

So the inverse of A is $A^{-1} = \begin{bmatrix} 7 & -2 \\ -3 & 1 \end{bmatrix}$.

Problem 4. Calculate the inverse of matrix A:

$$A = \begin{bmatrix} 1 & 2 & 0 \\ 2 & 6 & 4 \\ 0 & 4 & 11 \end{bmatrix}$$

Solution:

$$\begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 2 & 6 & 4 & 0 & 1 & 0 \\ 0 & 4 & 11 & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 2 & 4 & -2 & 1 & 0 \\ 0 & 4 & 11 & 0 & 0 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 2 & 4 & -2 & 1 & 0 \\ 0 & 0 & 3 & 4 & -2 & 1 \end{bmatrix}$$

$$\to \begin{bmatrix} 1 & 2 & 0 & 1 & 0 & 0 \\ 0 & 6 & 0 & -22 & 11 & -4 \\ 0 & 0 & 3 & 4 & -2 & 1 \end{bmatrix} \to \begin{bmatrix} 3 & 0 & 0 & 25 & -11 & 4 \\ 0 & 6 & 0 & -22 & 11 & -4 \\ 0 & 0 & 3 & 4 & -2 & 1 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & \frac{25}{3} & -\frac{11}{3} & \frac{4}{3} \\ 0 & 1 & 0 & -\frac{11}{3} & \frac{11}{6} & -\frac{2}{3} \\ 0 & 0 & 1 & \frac{4}{3} & -\frac{2}{3} & \frac{1}{3} \end{bmatrix}$$

2.1 向量空间与子空间

一. 向量空间 Vector Space

1. 本质：集合

2. 运算：

(1) 加法 Addition

$\forall v, w \in V, v+w \in V$（封闭性）

(2) 数乘 Scalar Multiplication

$\forall v \in V, \forall \lambda \in \mathbb{R}(\mathbb{C}, \cdots), \lambda v \in V$（封闭性）

3. 八大性质

加法公理：加法交换律、加法结合律、加法单位元、加法逆元；

乘法公理：乘法结合律、乘法单位元；

分配律（2条）

牢记！

4. 例

$\mathbb{R}^n$：全体n维实列向量的集合

$\mathbb{R}^{m,n}$：全体m行n列实矩阵的集合

…

二. 子空间 Subspace

1. 定义：对于 $U \subset V$（V是向量空间，下同），若在使用与V一致的运算下，U也是向量空间，则称U是V的一个子空间.

2. 判定准则：

1° $U \neq \varnothing$（$0 \in U$）

2° $\forall u, v \in U, u+v \in U$（加法封闭）

3° $\forall u \in U, \forall \lambda \in \mathbb{R}, \lambda u \in U$（数乘封闭）

第1°条保证了空集不是子空间

3. 常见的子空间:

$\mathbb{R}^2$ 的全部子空间: $\begin{cases} \{0\} \\ \mathbb{R}^2 \text{本身} \\ \text{过原点的直线} \end{cases}$

$\mathbb{R}^3$ 的全部子空间: $\begin{cases} \{0\} \\ \mathbb{R}^3 \text{本身} \\ \text{过原点的直线、过原点的平面} \end{cases}$

(2) $U = \{(x_1, x_2, x_3, x_4, x_5) : x_1 + 2x_2 + 3x_3 + 4x_4 + 5x_5 = 0\} \subset \mathbb{R}^5$

是 $\mathbb{R}^5$ 的子空间.

(3) 若 $U_1, U_2 \subset V$ 是子空间 则 $U_1 \cap U_2$ 也是 V 的子空间.

(4) 记 V 为全体次数小于等于 n 的实系数多项式的集合,
则以下均为 V 的子空间:

1° $U_1 = \{f \in V : f(2) = 0\}$

2° $U_2 = \{f \in V : f'(-1) = 0\}$

3° $U_3 = \{f \in V : \int_{-1}^{1} f\,dx = 0\}$

(5) $U \subset \mathbb{R}^{n,n}$, U 为全体对称矩阵的集合.

(请自行验证 (2) ~ (5)!)

4. 矩阵的列空间: 各列向量张成的空间 Column Space

若 $A = [\alpha_1 \cdots \alpha_n]$, 则

$$C(A) = \{c_1\alpha_1 + \cdots + c_n\alpha_n : c_1, \cdots, c_n \in \mathbb{R}\} \subset \mathbb{R}^m$$

$C(A)$ 也可表示为 $\{Ac : c \in \mathbb{R}^n\}$. (为什么?)

矩阵的零空间: 使 $Ax = 0$ 的全体 $x \in \mathbb{R}^n$ 组成的空间. Nullspace

$$N(A) = \{x \in \mathbb{R}^n : Ax = 0\} \subset \mathbb{R}^n.$$ (为什么零空间是子空间?)

2.2. $Ax=0$. $Ax=b$, 基础解系

一. $Ax=0$ 的解法： Homogeneous Equations

1. 化为 $Ux=0$, 简化方程组.
2. 化为 $Rx=0$, 用零空间矩阵.（具体如何执行？）

解 $Ax=0$ 的本质：

求 $N(A)$

二. 秩 Rank

1. 定义：A 的主元个数.
2. 性质：对于 $m\times n$ 矩阵，秩为 r，必有

$r\le m$; $r\le n$;

在 $A\to U\to R$ 过程中，秩不发生改变.

三. $Ax=b$ 的解法： Non-homogeneous Equations

Step 1. 求出 $Ax=0$ 的全部解；（$X_{nullspace}$

Step 2. 求 $Ax=b$ 的一个解（用 $Ux=c$ / $Rx=d$ 解决）

（$X_{particular}$

Step 3. $X_{complete}=X_{particular}+X_{nullspace}$.

（为什么 Step 3 成立？）

注：任何 $Ax=b$ 的一个解均可充当特解；

对于 $Ax=b\neq 0$，若有解，其全部解系集不构成子空间；

$Ax=b$ 未必有解；$Ax=b$ 有解当且仅当 $b\in C(A)$，当且仅当将 $[A\ b]$ 化简至 $[U\ c]$ 中，U 的每个全零行与 c 中的 0 对应.

III. 满秩的基础解系：

$A: m\times n,\ \operatorname{rank}A = r.$

1. $m > n = r$:

$[A] \to R = \begin{bmatrix} I_r \\ 0 \end{bmatrix}$

	无解	恰1解	无穷多解
$Ax=0$		✓	
$Ax=b$	✓	✓	

2. $n > m = r$:

$[A] \to R = [I_r\ F]$

	无解	恰1解	无穷多解
$Ax=0$			✓
$Ax=b$			✓

3. $m = n = r$:

$[A] \to R = I$

	无解	恰1解	无穷多解
$Ax=0$		✓	
$Ax=b$		✓	

（为什么？）

2.3 线性无关、线性相关、张成、基、维数

一. 线性无关、线性相关、张成

1. 线性无关 Linear independence

定义：对于 $v_1,\dots,v_n \in V$，若对于 $c_1,\dots,c_n \in \mathbb{R}$，

$$c_1v_1+\dots+c_nv_n = 0 \Rightarrow c_1 = \dots = c_n = 0,$$

则称 $v_1,\dots,v_n$ 线性无关.

2. 线性相关 Linear dependence

定义：对于 $v_1,\dots,v_n \in V$，若不是线性无关的，则称它们是线性相关的.

也就是说，存在不全为0标量 $c_1,\dots,c_n$ 使

$$c_1v_1+\cdots+c_nv_n=0.$$

3. 张成　　Span.

对于 $v_1,\dots,v_n\in V$，其全体线性组合构成的空间称为 $v_1,\dots,v_n$ 的张成空间：

$$\mathrm{span}(v_1,\dots,v_n)=\{c_1v_1+\cdots+c_nv_n : c_1,\dots,c_n\in\mathbb{R}\}.$$

若 $\mathrm{span}(v_1,\dots,v_n)=V$，则称 $v_1,\dots,v_n$ 张成 V.

4. 结论：

(1) $v_1,\dots,v_n\in V$ 线性无关当且仅当对于 $\forall v\in\mathrm{span}(v_1,\dots,v_n)$，$v$ 对应的 $v_1,\dots,v_n$ 的线性表示是唯一的.

(请自行证明！)

(2) V 中一组向量的张成空间是包含这组向量的最小子空间.

(集合的"大小"用什么比较？)

(3) 若 $v_1,\dots,v_n$ 线性相关，则 $\exists v_j$ 使 v_j 可以被其余的向量线性表示.

("$\exists$"不能改成"$\forall$"，为什么？)

(4) 对于向量空间 V，V 的每个线性无关组的长度 $\leq$ 张成组的长度.

("长度"即向量组中向量的个数. (3)(4) 不需证明).

(试着用 (4) 证明：

V 中任何长度小于 $\dim V$ 的向量组不可能张成 V；

V 中任何长度大于 $\dim V$ 的向量组不可能线性无关.)

(5) 矩阵各列线性无关当且仅当 $N(A)=\{0\}$.

(请自行证明！)

二. 基与维数　　Basis and dimension

1. 基：定义：线性无关的张成组是V的一组基。

2. 维数：定义：基的长度称为V的维数。

3. 结论：

(1) $v_1, \dots, v_n \in V$是V的一组基当且仅当$\forall v \in V$，$\exists$唯一的$c_1, \dots, c_n$使

$$v = c_1 v_1 + \cdots + c_n v_n.$$

(为什么？)

(2) 每个V中的张成组均可化简为V的一个基。

(从中移除n ($n \geq 0$)个向量)

(3) 每个V中的线性无关组均可扩充为V的一个基

(从中添加n ($n \geq 0$)个向量)

(4) V中任何两个基的长度相同。

(这保证了"维数"定义良好。)

(如何证明？)

(5) 任何长度为$\dim V$的线性无关组均是V的一组基；

任何长度为$\dim V$的张成组均是V的一组基。

(例如，$\begin{bmatrix}1\\2\end{bmatrix}$，$\begin{bmatrix}1\\3\end{bmatrix}$是$\mathbb{R}^2$的一组基，因为它们线性无关，且长度$= \dim \mathbb{R}^2 = 2$。)

(6) 若U是V的子空间，且$\dim U = \dim V$，则$U = V$。

(为什么？)

4. 常见向量空间的维数

(1) $\dim \mathbb{R}^n = n$；

(2) $\dim \mathbb{R}^{m,n} = m \cdot n$；

(3) $\dim C(A) = \operatorname{rank} A$。(各主元列为$C(A)$的一组基)。

2.4 The Four Fundamental Subspaces

Fundamental Theorem of Linear Algebra, Part I

1. $C(A)$ = column space of A; dimension r.
2. $N(A)$ = nullspace of A; dimension $n-r$.
3. $C(A^{\mathrm{T}})$ = row space of A; dimension r.
4. $N(A^{\mathrm{T}})$ = left nullspace of A; dimension $m-r$.

3D Fundamental Theorem of Linear Algebra, Part II

The nullspace is the *orthogonal complement* of the row space in $\mathbf{R}^n$.

The left nullspace is the *orthogonal complement* of the column space in $\mathbf{R}^m$.

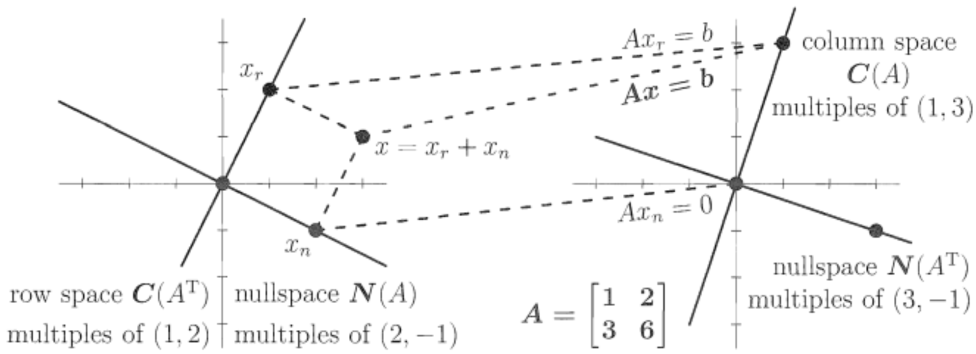


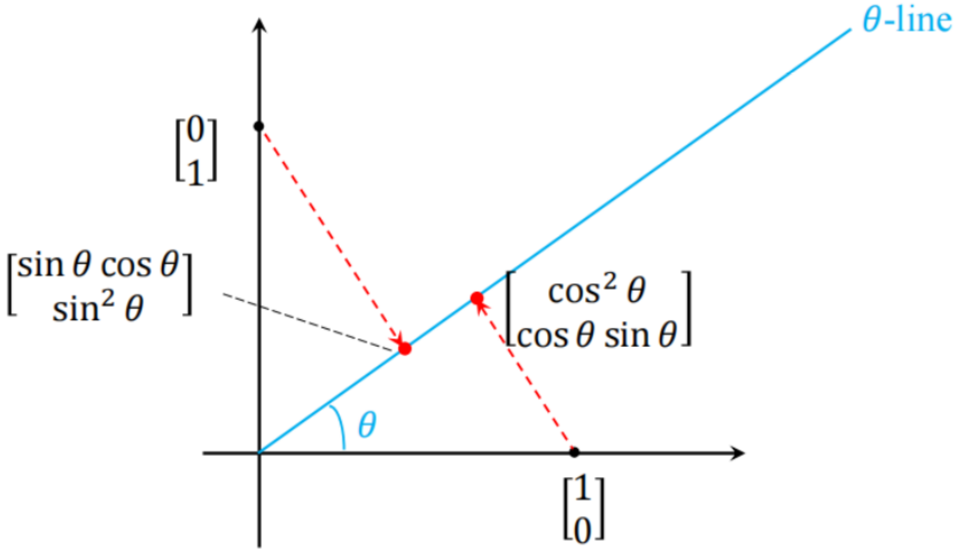
Figure 2.5: The four fundamental subspaces (lines) for the singular matrix A.

Every matrix of rank 1 has the simple form $A = uv^{\mathrm{T}}$ = column times row.

R $[I]$	$r=m=n$	inv.	1 solution	Both	左逆=右逆
$\begin{bmatrix} I \\ 0 \end{bmatrix}$	$r=n<m$	full col rank	1/0 solution	Unique	左逆 $(A^TA)^{-1}A^T$
$[I\ F]$	$r=m<n$	full row rank	∞	Exists	右逆 $A^T(AA^T)^{-1}$
$\begin{bmatrix} I & F \\ 0 & 0 \end{bmatrix}$	$r<m, r<n$		$0/\infty$	None	

2.6 Linear Transformations

- Projection onto the θ-line(the line at the angle θ from x-axis): p

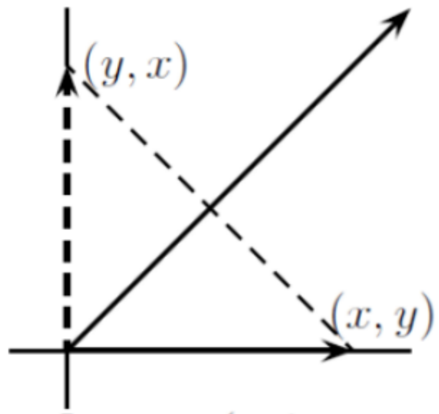


$$p^2 = p$$

- The reflection with respect to the line $x = y$

Let $A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, then

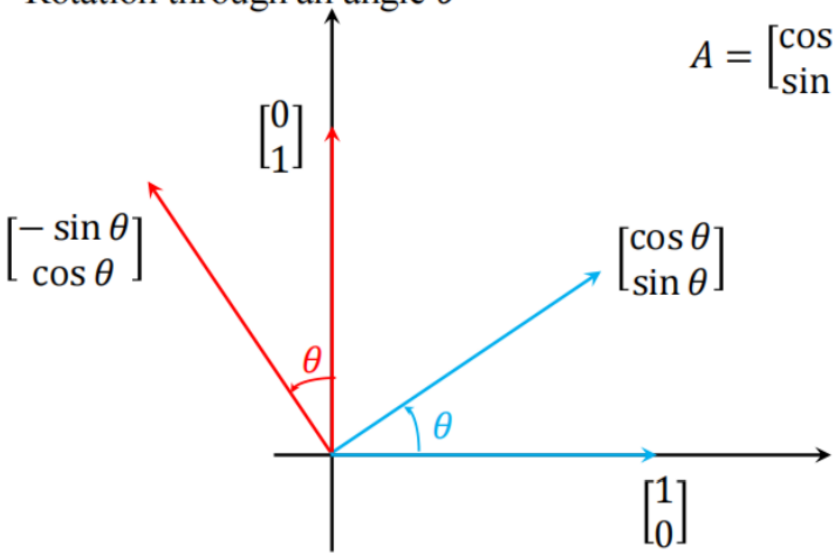
$$A\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix}$$



reflection (45° mirror)

- Rotation through an angle θ

$$A = \begin{bmatrix} \cos\theta & -\sin\theta \\ \sin\theta & \cos\theta \end{bmatrix}$$



$$A^n = \begin{bmatrix} \cos(n\theta) & -\sin(n\theta) \\ \sin(n\theta) & \cos(n\theta) \end{bmatrix}$$

线性映射的矩阵

设V, W分别为n, m维的向量空间

$\vec{v_1},\cdots,\vec{v_n}$是V一组基, $\vec{w_1},\cdots,\vec{w_m}$是W一组基

T是从V到W的一个线性映射, 且

$$T(\vec{v_1}) = A_{1,1}\vec{w_1} + \cdots + A_{m,1}\vec{w_m}$$
$$T(\vec{v_2}) = A_{1,2}\vec{w_1} + \cdots + A_{m,2}\vec{w_m}$$
$$\cdots$$
$$T(\vec{w_n}) = A_{1,n}\vec{w_1} + \cdots + A_{m,n}\vec{w_m}$$

写成矩阵形式

$$\left[T(\vec{v_1}) \ \cdots \ T(\vec{v_n})\right] = \left[\vec{w_1} \ \cdots \ \vec{w_m}\right]\begin{bmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots \\ A_{m,1} & \cdots & A_{m,n} \end{bmatrix}$$

$\forall \vec{v} \in V.\ \exists a_i \ st.\ \vec{v} = \sum_{i=1}^{n} a_i \vec{v_i}$

即 $\vec{v} = \left[\vec{v_1} \cdots \vec{v_n}\right]\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \Rightarrow \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} = \left[\vec{v_1} \cdots \vec{v_n}\right]^{-1}\vec{v}$

$$T(\vec{v}) = T\left(\sum_{i=1}^{n} a_i \vec{v_i}\right) = \sum_{i=1}^{n} a_i T(\vec{v_i})$$

即

$$T(\vec{v}) = \left[T(\vec{v_1}) \cdots T(\vec{v_n})\right]\begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix}$$
$$= \underbrace{\left[\vec{w_1} \cdots \vec{w_m}\right]\begin{bmatrix} A_{1,1} & \cdots & A_{1,n} \\ \vdots & \ddots & \vdots \\ A_{n,1} & \cdots & A_{n,n} \end{bmatrix}\left[\vec{v_1} \cdots \vec{v_n}\right]^{-1}\vec{v}}_{\text{该线性映射完整过程}}$$

3.1 orthogonal vectors and subspaces

① inner product

$u = (a_1, a_2, \cdots, a_n)^T \quad v = (b_1, b_2, \cdots, b_n)^T$

$u^T v = a_1 b_1 + a_2 b_2 + \cdots + a_n b_n$

$u^T v = v^T u \qquad (u+v)^T w = u^T w + v^T w$

$(cu)^T v = c u^T v = u^T (cv)$

② length

$\|u\| = \sqrt{a_1^2 + a_2^2 + \cdots + a_n^2} = \sqrt{\sum_{i=1}^{n} a_i^2}$

③ orthogonal vectors:

if $u^T v = 0$, then u, v are orthogonal (perpendicular)

④ mutually orthogonal vectors:

if nonzero vectors $v_1, v_2, \cdots, v_k$ are mutually orthogonal, then those vectors are linearly independent.

⑤ orthonormal basis:

a basis $\{v_1, v_2, \cdots, v_n\}$ is called orthonormal if $\|v_i\| = 1$, $v_i^T v_j = 0 \ (i \neq j)$

⑥ orthogonal subspaces:

let U, W be two subspace of a vector space $V = R^n$. If every vector v in U is orthogonal to every vector w in W, then $U \perp W$

⑦ fundamental theorem of orthogonality

$N(A) \perp C(A^T) \qquad C(A) \perp N(A^T)$

⑧ orthogonal complement.

Given a subspace V of R^n, the space of all vectors orthogonal to V is called the orthogonal complement of V.

$N(A) = (C(A^T))^{\perp} \qquad N(A^T) = (C(A))^{\perp}$

dim (column space) + dim (left nullspace) = number of rows

dim (row space) + dim (nullspace) = number of columns

④ lemma : a system $Ax=b$ has solutions if and only if
$y^Tb=0$, whenever $y^TA=0$

⑥ each matrix A transforms its row space $C(A^T)$ onto its column space $C(A)$

3.2

① proj_a :

the projection of the vectors of V onto the line in the direction of a.

$(v-P_v a)\perp a \Rightarrow P_v = \frac{a^Tv}{a^Ta}$ (a number)

$\text{proj}_a(v) = \frac{a^Tv}{a^Ta}a = \frac{aa^T}{a^Ta}v$

② Cauchy-Schwarz inequality

$u=(a_1,a_2,\cdots,a_n)$ $v=(b_1,b_2,\cdots,b_n)$

$|a_1b_1+a_2b_2+\cdots+a_nb_n| \le \sqrt{a_1^2+a_2^2+\cdots+a_n^2}\sqrt{b_1^2+b_2^2+\cdots+b_n^2}$

$|u^Tv| \le \|u\|\,\|v\|$

③ projection matrix

$P=\frac{aa^T}{a^Ta}$ $P^T=P$ $P^2=P$

④ cosines

$\cos\theta = \frac{u^Tv}{\|u\|\|v\|}$ $\left|\frac{u^Tv}{\|u\|\|v\|}\right| \le 1$

3.3

① when $Ax=b$ is inconsistent. find $\hat{x}$ such that $\|A\hat{x}-b\|$ is as small as possible

$\hat{x}$: least square solutions.

searching for $\hat{x}$ is the same as locating the point $p=A\hat{x}$ that is closer to be than any other point in the $C(A)$

$(b-A\hat{x})\perp C(A)$ so $b-A\hat{x}$ lies in the left nullspace of A.

$A^T(b-A\hat{x})=0$

so: normal equations $A^TA\hat{x}=A^Tb$

best estimate $\hat{x}$ $\hat{x}=(A^TA)^{-1}A^Tb$

projection : $p=A(A^TA)^{-1}A^Tb$

$\boxed{\text{the premise is that } A^TA \text{ is invertible}}$

A^TA is invertible exactly when the columns of A are linearly independent.

$\boxed{\text{Please use Gaussian elimination to solve } A^TA\hat{x}=A^Tb}$

if $b \in C(A)$ $\quad p=b$

if b is perpendicular to every column $\quad A^Tb=0 \quad p=0$

if A is invertible, $p=A(A^TA)^{-1}A^Tb = AA^{-1}(A^T)^{-1}A^Tb=b$.

② when A is vector a

$$\hat{x}=\frac{a^Tb}{a^Ta}$$

③ the matrices A^TA and A have the same nullspace. (please prove it)

In particular, if A has full column rank, then A^TA is invertible.

④ projection matrices:

if A^TA is invertible, $P=A(A^TA)^{-1}A^T \quad P^2=P,\ P^T=P$

⑤ $e=b-Pb \perp C(A)$

⑥ any symmetric matrix with $P^2=P$ represents a projection.

⑦ least-squares fitting of data

$$A=\begin{bmatrix} 1 & t_1 \\ \vdots & \vdots \\ 1 & t_m \end{bmatrix} \quad x=\begin{bmatrix} C \\ D \end{bmatrix} \quad b=\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

$$A^TA\hat{x}=A^Tb$$