

## 2. Pseudosolutions and its applications. Linear regression.

### Intuition behind pseudosolutions

Let's recall that the system of linear equations (SLE) can be represented using the following notation

$$A\vec{x} = \vec{b},$$

where

$$A = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix} \in M_{m \times n}(\mathbb{C}), \quad \vec{b} = \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix} \in \mathbb{C}^m, \quad \vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in \mathbb{C}^n.$$

Let's recall possible solution sets of SLE

#### Theorem: Solution sets of SLE

linear system of equations may behave in any one of three possible ways

1. The system is definite - has a single unique solution.
2. The system is indefinite - has infinitely many solutions.
3. The system is inconsistent - has no solution.

**Example 1: Definite system of the square system.**

If  $A \in M_{n \times n}(\mathbb{C})$ ,  $\text{rank } A = n$ , then we can easily obtain unique solution  $\vec{x}$  by inverting the matrix of initial coefficients

$$\vec{x} = A^{-1}\vec{b}.$$

We want to generalize this result for any systems (definite, indefinite and inconsistent of any size). Let's outline intuition of this generalization in the following examples.

### Definite system

**Example 2:** If  $A \in M_{m \times n}(\mathbb{C})$ ,  $\text{rank } A = n$ , then there exist unique solution  $\vec{x}$ . For example, consider a system

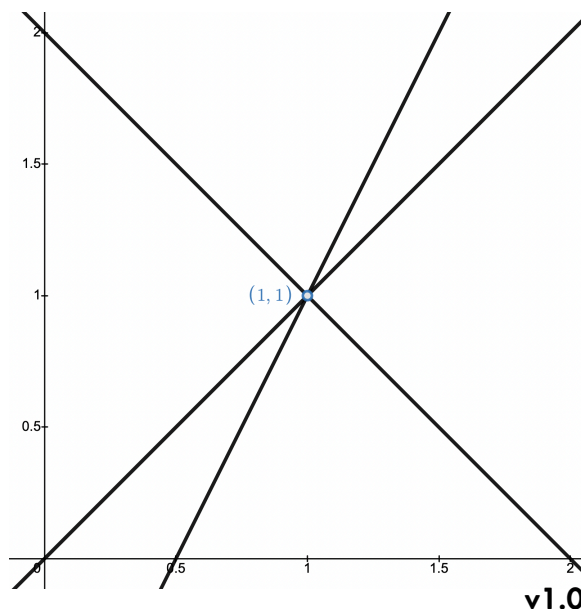
$$\begin{cases} x - y = 0 \\ x + y = 2 \\ 2x - y = 1 \end{cases}.$$

The system has only one solution in the point  $\vec{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ .

However, we would like to generalize the method of obtaining a solution in such a way that it looks similar to the first example (definite system of square system)

$$\vec{x} = ? \cdot \vec{b}.$$

And looking ahead we can obtain such a factor to express solution that way. But now let's get a broader generalization.



Example of definite system.

### Indefinite system

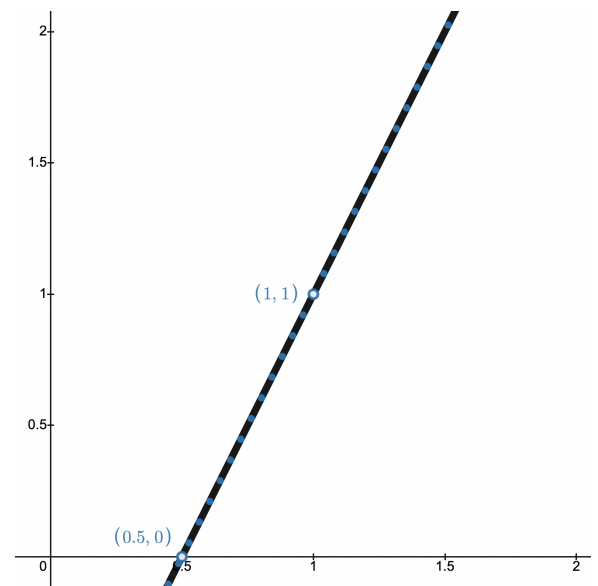
Example 3: Consider a system

$$\begin{cases} 2x - y = 1 \\ 4x - 2y = 2 \end{cases}$$

It is not so obvious to choose a specific solution here because a whole family of solutions of the following form  $\vec{x} = \begin{bmatrix} x \\ 1 - 2x \end{bmatrix}$  is suitable for us. At the same time, we would like express solution of the system in the following form

$$\vec{x} = ? \cdot \vec{b},$$

which assumes one solution, not infinite. For this reason, we need to choose the best solution out of infinite. In other words, we need to define "optimal" solution. Our intuition hints us to a solution point which is the closest to the origin.



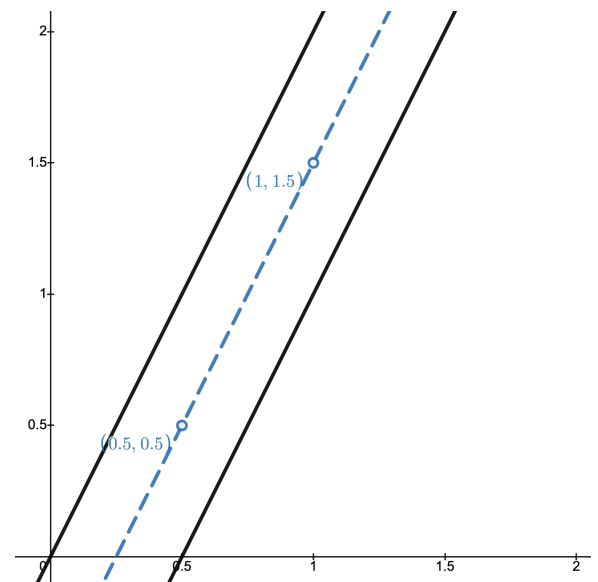
Example of indefinite system.

### Inconsistent system

Example 4: Consider a system

$$\begin{cases} 2x - y = 0, \\ 2x - y = 1 \end{cases}$$

The system does not have solution. Geometrically the system describes two parallel lines. Our intuition hints us to choose a point on a line right in middle of the parallel lines. However, it is still a whole family of solutions that can be the answer to the request of the product or business problem.



Example of inconsistent system.

## Pseudosolution

**Definition: Pseudosolution**

Consider a system of a linear equations  $A\vec{x} = \vec{b}$ . A vector  $\vec{u} \in \mathbb{C}^n$  is called a pseudosolution (least square solution), if

$$|A\vec{x} - \vec{b}| \geq |A\vec{u} - \vec{b}|, \quad \forall \vec{x} \in \mathbb{C}^n.$$

That is, vector  $f(\vec{x}) = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix} = A\vec{x} - \vec{b}$  reaches it's minimum length at  $\vec{x} = \vec{u}$ . That is,  $\vec{x} = \vec{u}$  is the solution of the following problem

$$|f(x)|^2 = |f_1|^2 + \dots + |f_n|^2 \rightarrow \min_{\vec{x}}.$$

**Theorem:  $\vec{u} = A^+ \vec{b}$** 

The vector  $\vec{u} = A^+ \vec{b}$  is a pseudosolution of the system of linear equations  $A\vec{x} = \vec{b}$ . Moreover, among all pseudosolutions, the vector  $\vec{u}$  is the unique pseudosolution with minimal length.

**Prop** If  $\hat{x}$  is a solution, then it is a pseudosolution.

**Proof:**  $A\hat{x} - \vec{b} = 0 \implies |A\hat{x} - \vec{b}| = 0 = \min |f_x|^2.$  □

**Prop**

Type of a system	Solution
definite	$\hat{x} = A^+ \vec{b}$ is the solution
indefinite	$\hat{x} = A^+ \vec{b}$ is the solution of minimal length
inconsistent	$\hat{x} = A^+ \vec{b}$ is the pseudosolution of minimal length

**Lemma**

$$\text{Im}(AA^+ - I) \perp \text{Im}(A).$$

**Proof:** Let us denote  $AA^+ - I$  by  $M$ . Let us denote vector columns that generate  $\text{Im}(M)$  and  $\text{Im}(A)$  by  $M^i$  and  $A^j$  respectively. Then the following holds

$$\begin{aligned} & \text{Im}(M) \perp \text{Im}(A) & (1) \\ \Leftrightarrow & M^i \perp A^j, \quad \forall i, j \\ \Leftrightarrow & (M^i, A^j) = 0, \quad \forall i, j \\ \Leftrightarrow & (M^i)^* \cdot A^j = 0, \quad \forall i, j \\ \Leftrightarrow & M^* A = 0, & (2) \end{aligned}$$

which means that by proving (2), we get (1). Let us prove (2).

$$M^* A = (AA^+ - I)^* A = ((AA^+)^* - I^*) A = (AA^+)^* A - I^* A \stackrel{\text{axiom III}}{=} AA^+ A - I \stackrel{\text{axiom I}}{=} A - A = 0.$$

□

**Theorem: Pythagoras**

Suppose  $\vec{a} \perp \vec{b}$ , then for  $\vec{c} = \vec{a} + \vec{b}$ , we have  $|\vec{c}|^2 = |\vec{a}|^2 + |\vec{b}|^2$ . In particular  $|\vec{c}| \geq |\vec{b}|$ . The equality holds only for  $\vec{a} = \vec{0}$ .

**Proof:** We leave as an exercise for the reader. □

**Proof:** (of Theorem  $\vec{u} = A^+ \vec{b}$ )

First, let us

prove that  $\vec{u} = A^+ \vec{b}$  is a pseudosolution. Let us denote

$$\vec{A}_x := A(\vec{x} - \vec{u}) = A\vec{x} - A\vec{u} \in \mathbb{C}^n, \quad \vec{B} := (AA^+ - I)\vec{b} = A\vec{u} - \vec{b} \in \mathbb{C}^n, \quad \vec{C} := \vec{A}_x + \vec{B} = A\vec{x} - \vec{b}.$$

Using Lemma, we get  $\vec{A}_x \perp \vec{B}$ . Using Pythagoras theorem, we get

$$\begin{aligned} |C| &\geq |B| \\ \Leftrightarrow \\ |A\vec{x} - \vec{b}| &\geq |A\vec{u} - \vec{b}|, \quad \forall \vec{x} \in \mathbb{C}^n, \end{aligned}$$

which means that  $\vec{u}$  is a pseudosolution.

Secondly, let us prove that  $\vec{u} = A^+ \vec{b}$  is the unique pseudosolution with minimal length. Let us denote  $w := \vec{x} - \vec{u}$ . Suppose  $\vec{x}$  is another pseudosolution, which means that

$$A\vec{x} = A\vec{u} \Leftrightarrow A(\vec{x} - \vec{u}) = 0 \Leftrightarrow Aw = 0. \quad (3)$$

Let us consider the following

$$\begin{aligned} (\vec{u}, \vec{w}) &= \vec{u}^* \vec{w} = (A^+ \vec{b})^* \vec{w} = \vec{b}^* A^+ \vec{w} \\ &\stackrel{II}{=} \vec{b}^* (A^+ AA^+)^* \vec{w} = \vec{b}^* A^+ (A^+ A)^* \vec{w} \\ &\stackrel{IV}{=} \vec{b}^* A^+ A^+ A \vec{w} \\ &= \vec{b}^* A^+ A^+ \underbrace{A \vec{w}}_{\vec{0}} \stackrel{(3)}{=} 0, \end{aligned}$$

which means that  $u \perp w$ . Using Pythagoras theorem, we get  $|\vec{x}| \geq |\vec{u}|$ , which means that  $\vec{u}$  is pseudosolution with minimal length. By Pythagoras theorem  $|\vec{x}| = |\vec{u}| \Leftrightarrow \vec{w} = 0 \Leftrightarrow \vec{x} = \vec{u}$ , which means that  $\vec{u}$  is unique. □

**Example 5:** Find pseudosolution of the following inconsistent system

$$A = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}, \quad \vec{b} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}$$

Then pseudosolution can be found by the formula

$$\hat{x} = A^+ \vec{b}.$$

Pseudoinverse matrix to  $A$  can be obtained by

$$A^+ = \left( \begin{bmatrix} 1 \\ 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \end{bmatrix} \right)^+ = \begin{bmatrix} 2 & 1 \end{bmatrix}^+ \cdot \begin{bmatrix} 1 \\ 1 \end{bmatrix}^+ = \frac{1}{2} \cdot \frac{1}{5} \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}$$

Then we can get a pseudosolution

$$\hat{x} = \frac{1}{10} \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 6 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 18 & 9 \end{bmatrix}$$

Prop

$\vec{x}$  is a pseudosolution  $\Leftrightarrow \vec{x}$  is a solution of the "normal" system of a kind

$$A^* A \vec{x} = A^* \vec{b}.$$

Prop

All pseudosolutions of  $A\vec{x} = \vec{b}$  are given by the formula

$$\vec{x} = A^+ \vec{b} - (A^+ A - I) \vec{y},$$

where  $\vec{y} \in \mathbb{C}^n$  – an arbitrary vector.

**Example 6:** Find all pseudosolutions of the system from previous example.  
We have already obtained one pseudosolution

$$\hat{x} = \frac{1}{10} \begin{bmatrix} 18 \\ 9 \end{bmatrix}.$$

Now we need to obtain  $A^+ A - I$

$$A^+ A - I = \frac{1}{10} \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \cdot \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \frac{1}{10} \begin{bmatrix} 8 & 4 \\ 4 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = \begin{bmatrix} -0.2 & 0.4 \\ 0.4 & -0.8 \end{bmatrix}.$$

Finally, all pseudosolutions are given by the formula

$$\vec{x} = \frac{1}{10} \begin{bmatrix} 18 \\ 9 \end{bmatrix} - \begin{bmatrix} -0.2 & 0.4 \\ 0.4 & -0.8 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} = \begin{bmatrix} 1.8 + 0.2y_1 - 0.4y_2 \\ 0.9 - 0.4y_1 + 0.8y_2 \end{bmatrix}.$$