1. Pseudosolutions and its applications. Linear regression.

Let's repeat main possible situation for solving linear equations task, which one can be written by the next notation:

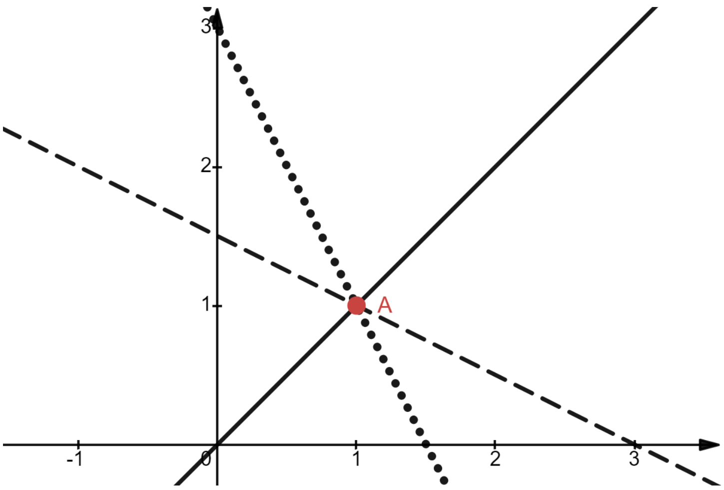
$$A\vec{x} = b,$$

where $A \in M_{m\times n}(\mathbb{C})$, $\vec{b} \in \mathbb{C}^m$, $\vec{x} \in \mathbb{C}^n$.

0. The first case is about square matrix $A \in M_{n\times n}(\mathbb{C})$, $\operatorname{rank} A = n$. In such situation we can easily obtain unique $\overline{x}$ by inverting the matrix of initial coefficients:

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

1. The next easy option is a definite system, when $A \in M_{m\times n}(\mathbb{C}), \operatorname{rank} A = n$. Then unique $\hat{x}$ can be expressed by the following ideas.



Example of definite system.

Consider a system of the form:

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

It is obviously that system have only one correct solution in the point A and it is a solution of a type: $\hat{x} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$. However, we would like to generalize the method of obtaining a solution in such a way that it looks similar to the first (zero) case, namely:

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

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

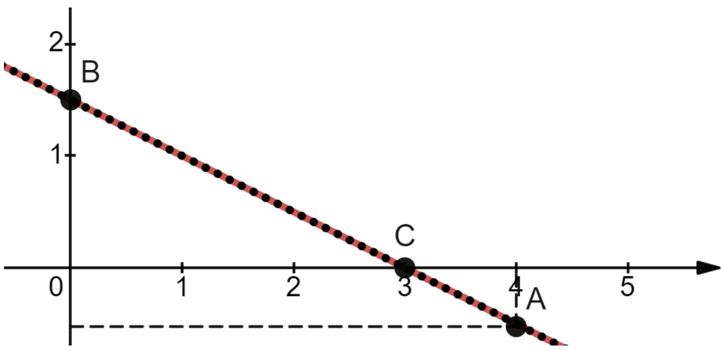
2. Also we can obtain an indefinite solution, that can provide us an infinite amount of solutions.

Consider a system of two equations:

$$\begin{cases} x + 2y = 3, \\ 2x + 4y = 5. \end{cases}$$

It is not so obvious to choose a specific solution here because a whole family of solutions of the following form $\hat{x} = \begin{bmatrix} 3 - 2y \\ y \end{bmatrix}$ is suitable for us.

And now we need to get some understanding about which solution is a kind of optimum. We will discuss it a little bit later, now let's consider one more possible situation.



Example of definite system.

3.

Inconsistent system is a system of a kind:

$$\begin{cases} 2x + y = 3, \\ 2x + y = 6 \end{cases}$$

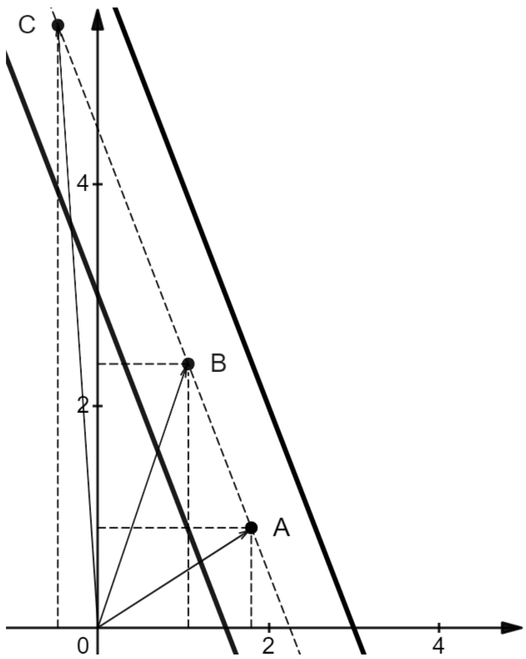
Need to remind, that we want to obtain solution in term of factors:

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

There we have matrix and vector of initial values

$$A = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} \text{ and } \vec{b} = \begin{bmatrix} 3 \\ 6 \end{bmatrix}.$$

Manually we can understand that the best solution will lie somewhere between two parallel lines, perhaps even exactly in the middle. But it is still a whole family of solutions that can be the answer to the request of the product or business problem. We need a general variant to find the best solution. For this we introduce the definition:



Example of inconsistent system.

Definition

Consider a system of a linear equations $A\vec{x} = \vec{b}$ $(A \in M_{m\times n}(\mathbb{C}))$. A vector $\vec{u} \in \mathbb{C}^n$ is called a pseudosolution or a least square solution, if $\forall \vec{x} \in \mathbb{C}^n$ the length of $A\vec{u} - \vec{b}$ is less or equal to the length of $A\vec{x} - \vec{b}$:

$$\left|A\vec{u} - \vec{b}\right| \leq \left|A\vec{x} - \vec{b}\right|.$$

That is: if $f_x = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix} = A\vec{x} - \vec{b}$, then $|f_x|^2 = |f_1|^2 + \ldots + |f_n|^2$ for $\vec{x} = \vec{u}$ is minimal.

Theorem

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, $\vec{u}$ has the minimal length.

Prop

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

<u>Proof:</u> $A\hat{x} - \vec{b} = 0 \Longrightarrow \left|A\hat{x} - \vec{b}\right| = 0 = \min |f_x|^2.$ □

Example 1:

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

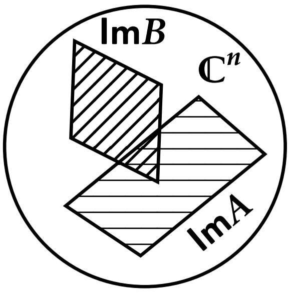
<u>Proof:</u> (Of the theorem) In proof we will use

Theorem: (Pythagoras)

Suppose $\vec{a} \perp \vec{b}$, that is $(\vec{a}, \vec{b}) = 0$. Then for $\vec{c} = \vec{a} + \vec{b}$: $|\vec{c}|^2 = |\vec{a}|^2 + |\vec{b}|^2$. In particular $|\vec{c}| \geq |\vec{a}|$. The equality holds only for $\vec{b} = \vec{0}$.

Lemma

$\mathrm{Im}A \perp \mathrm{Im}B$, where $B = AA^+ - I$.



We should prove: each column A^j of A is orthogonal to the one B^l of B.

$$\text{OR: } \forall l \ \left(B^l, A^j\right) \stackrel{?}{=} 0, \text{ or } B^{l^*} \cdot A^j \stackrel{?}{=} 0,$$

$$\text{or } (B^*)_l \cdot A^j \stackrel{?}{=} 0, \text{ or } B^* A \stackrel{?}{=} 0.$$

$$((AA^+)^* - I^*)A = (AA^+ - I)A = AA^+A - A = 0.$$

□