2. 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:

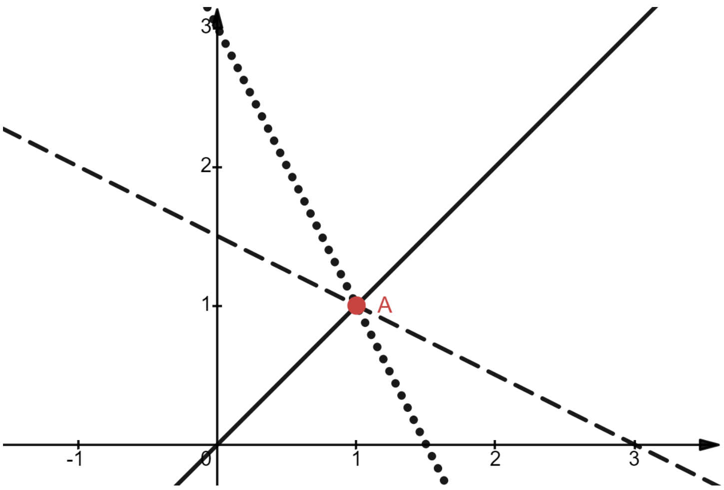
$$A\vec{x} = b,$$

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.

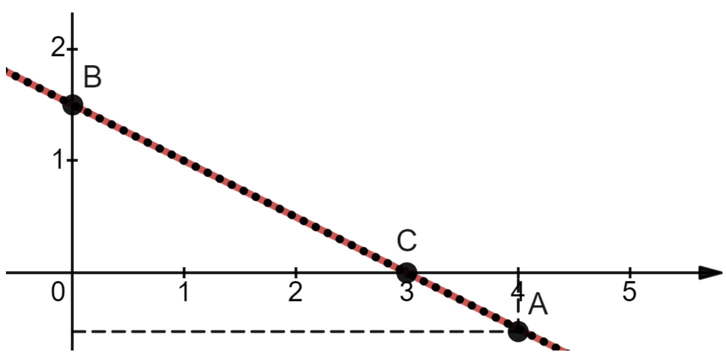
2. Also we can obtain an indefinite solution, that can provide us an infinite amount of solutions.

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:

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

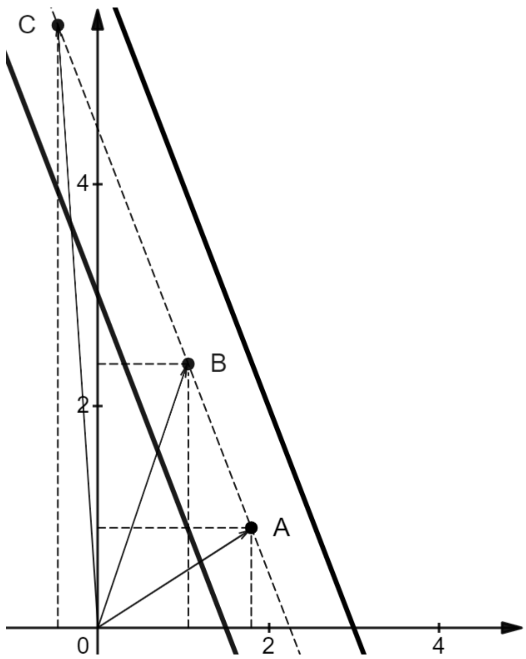
Need to remind, that we want to obtain solution in term of factors:

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

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 $A\vec{x} - \vec{b}$:

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

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.

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

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

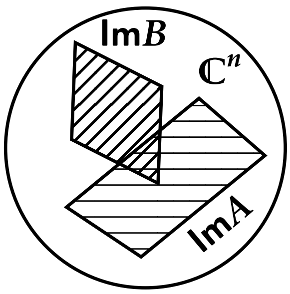
Proof: (Of the theorem) In proof we will use

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 $B = AA^+ - I$.



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

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

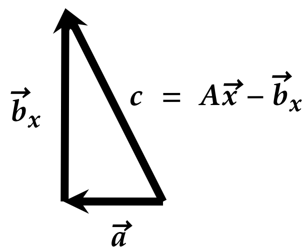
$$\text{or } (B^*)_l \cdot A^j \stackrel{?}{=} 0, \text{ or } B^* A \stackrel{?}{=} 0.$$

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

□

We need to prove that $\vec{u}$ is a pseudosolution. Let $\vec{x} \in \mathbb{C}^n$. We need to show:

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



Here $\vec{c} = A\vec{x} - \vec{b} = A\vec{x} - A\vec{u} + A\vec{u} - \vec{b} = A(\vec{x} - \vec{u}) + AA^+\vec{b} - \vec{b} =$

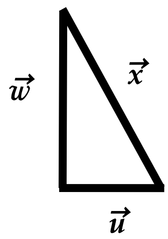
$$= A(\vec{x} - \vec{u}) + (AA^+ - I)\vec{b} = \underbrace{A(\vec{x} - \vec{u})}_{\vec{b}_x} + \underbrace{B\vec{b}}_{\vec{a}}.$$

By Lemma, $\vec{b}_x \perp \vec{a}$. By Pythagoras theorem, $|\vec{c}| = \min \iff \vec{b}_x = \vec{0}$. For example, it is so for $\vec{x} = \vec{u}$. So $\vec{u}$ is a pseudosolution.

We have shown, what $\vec{x}$ is a pseudosolution $\iff \vec{b}_x = 0$ or $A(\vec{x} - \vec{u}) = 0$, or $A\vec{x} = A\vec{u}$, or $A\vec{x} = AA^+\vec{b}$.

Suppose $\vec{x}$ is another pseudosolution. We need to prove that $|\vec{u}| \leq |\vec{x}|$. Let $\vec{w} = \vec{x} - \vec{u}$.

If we prove that $\vec{w} \perp \vec{u}$ then $|\vec{x}| \geq |\vec{u}|$. We have



$$\left(\vec{u}, \vec{w}\right) = \vec{u}^* \vec{w} = (A^+\vec{b})^* \vec{w} = b^* A^{+^*} \vec{w}, \text{ where } A\vec{w} = 0$$

Here $(A^+)* \stackrel{II}{=} (A^+AA^+)^* = A^{+^*}(A^+A)^* \stackrel{IV}{=} A^{+^*}A^+A$. So

$$(\vec{u}, \vec{w}) = b^* A^{+^*} A^+ \underbrace{A\vec{w}}_{0} = 0.$$

□

Let's return to our inconsistent system and find pseudosolution:

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

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 $\Longleftrightarrow$ $\vec{x}$ is a solution of the "normal"system of a kind:

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

Prop

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

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

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

Now let's find all pseudosolutions for example with inconsistent system. 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,

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

Linear regression problem