1. Difference between fundamental and applied linear algebra. Problems with the real data. Pseudoinverse matrices. Skeletonization.

Let's consider some of the standard linear algebra problems, for example, solving the systems of linear equations. It can be written:

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

where $A = M_{m \times n}(F)$ – the matrix of coefficients, $\vec{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \in F^n$ – unknown vector and $\vec{b} \in F^m$ – known

vector. Solving such systems is our goal. In the best situation we can write out the solution:

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

Or we have another method in a more general situation, when the matrix of the coefficients can be rectangular or degenerate, not inverted. We can provide Gaussian elimination:

After that we can easily express one variable in terms of another one step by step.

But in the real world with the linear models, the initial data can be inaccurate due to the observational errors in some physical cases or human reliability in, for example, social or business situations. It can lead to some problems. For example, in Gaussian elimination, you need to choose a pivot variable, and if it contains some errors, then other computations will increase them. In situations with high order errors, such methods cannot be applied. But even if you have the exact formula and enough resources for calculating the inverted matrix, you will realize that the inverse is obtained with some errors. Another problem is rounding. It happens because of the precise nature of Gaussian elimination algorithm or other algorithms and machine precision. It is hard and slow to calculate the precision of the solution.

The next problem is about speed or complexity of calculations. For example, the complexity for Gaussian elimination is $\mathcal{O}(n^3)$. It is not suitable for dealing with, for example, video or signals in real time.

Suppose you have created a linear model that more or less describes the production or business process. For example, you get the vector of unknowns $\vec{x}$, which means that you should sell x_1 copies of product 1 and x_2 copies of product 2, etc. How much must you produce or sell? It is a set of possible scenarios needed to solve. Then you can face an indefinite system because your initial data are not exact.

Indefinite (indeterminate) system

For example, it means system with less equations than variables or linearly dependent equations. It can be some system of a kind:

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

So we have two lines of the same kind. The problem is how to pick one value of the vector $\vec{x}$.

Mathematically, we can pick every point from this line, but, for example, negative values (point A), obviously looks strange or even inappropriate for our case. Or, for example, you can pick either point B, or C. Again, mathematically there are solutions of our systems, but it can be no optimum solution. We will understand it on the next lecture. Another problem is about inconsistent system if there is no set of values for the unknown vector $\vec{x}$ that satisfies all of the equations. Or in mathematical terms, it happens when rank of initial coefficients matrix is less than rank of augmented matrix (a matrix of a kind $[A|b]$). Such systems can be either determined (defined) or underdetermined.
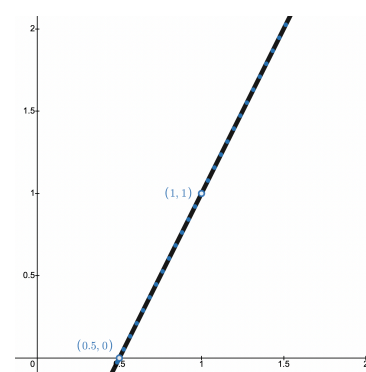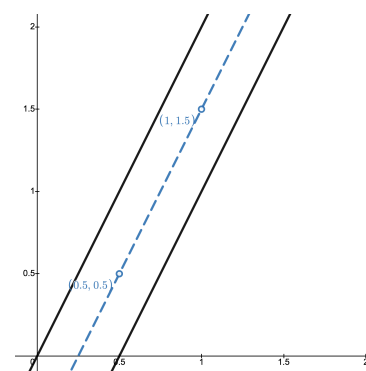

Example of indefinite system.

Inconsistent system

For example, you have approximately calculated several coefficients of the initial matrix and after the approximate calculations you obtained the system of equations, for example, a kind of:

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

It looks like 2 parallel lines. Of course, rarely some $\vec{x}$ exists, but we cannot solve this system because the coefficients are not exact (they contain some random errors, rounding after calculations, etc.). But answers can still be the same. How much do we need to sell? We can choose, for example, solutions A or B or even C, but what would be more optimal or correct. We will give an answer to this question right off the bat in the next lecture. Now let's define some new concepts that will help us with it.

Example of inconsistent system.

Pseudoinverse Matrices

Definition: Pseudoinverse

Let $A \in M_{m \times n}(\mathbb{C})$. Then C is called pseudoinverse matrix to A, or Moore-Penrose (pseudo) inverse, if it is satisfies Penrose axioms:

    I. $ACA = A$;
   II. $CAC = C$;
  III. $(AC)^* = AC$;
  IV. $(CA)^* = CA$.

Notation: $C = A^+$.

Note

If $A \in M_{m \times n}(\mathbb{C})$, then $C \in M_{n \times m}(\mathbb{C})$.

Note

if $A \in M_{n \times n}$, $\det A \neq 0$, then A^{-1} is a pseudoinverse.

Prop

If a pseudoinverse matrix C to A exists, it is unique.

<u>**Proof:**</u>   Suppose B is some another pseudoinverse matrix to A. Then:

$$AB \overset{I}{=} (AC)(AB) \overset{III}{\Rightarrow} (AC)^*(AB)^* \Rightarrow C^*(A^*B^*A^*) = C^*(ABA)^* = (AC)^* = AC.$$

Similarly, $BA = CA$. Now, $B \overset{II}{=} BAB = BAC = CAC = C.$ □

Example 1: $A = O_{m \times n} = \begin{bmatrix} 0 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 0 \end{bmatrix}$, then $A^+ = O_{n \times m}$.

Example 2: $\begin{bmatrix} 5 & 0 \\ 0 & 0 \\ 0 & 0 \end{bmatrix}^+ = \begin{bmatrix} \frac{1}{5} & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix}$.

Example 3: $\begin{bmatrix} 1 \\ 2 \end{bmatrix}^+ = \begin{bmatrix} \frac{1}{5} & \frac{2}{5} \end{bmatrix}$.

Example 4: Let $A = \vec{a} \in \mathbb{C}^n$, then

$$\vec{a}^+ = \frac{1}{a^*a}\vec{a}^* = \frac{1}{|\vec{a}|^2}\vec{a}^* = \frac{1}{|a_1|^2 + \ldots + |a_n|^2}\vec{a}^*.$$

Prop

Suppose that $A \in M_{m \times n}(\mathbb{C})$ has full column rank, that is, $\operatorname{rank} A = n$. Then $A^+ = (\underbrace{A^*A}_{n \times n})^{-1}A^*$.

Prop

Suppose that $A \in M_{m \times n}(\mathbb{C})$ has full row rank, that is, $\operatorname{rank} A = m$. Then $A^+ = A^*(\underbrace{AA^*}_{m \times m})^{-1}$.

Exercise: Check I-IV axioms for these A^+.   **!!!**

Definition: Skeletonization

A full rank decomposition (or skeletonization) of a matrix $A \in M_{m \times n}(\mathbb{C})$ with $r = \operatorname{rank} A$ is a decomposition:

$$A = F \cdot G, \qquad \begin{matrix} F \in M_{m \times r}(\mathbb{C}), \\ G \in M_{r \times n}(\mathbb{C}). \end{matrix}$$

Then $\operatorname{rank} F = \operatorname{rank} G = r$. F, G are called matrices of full rank.

Theorem

For each matrix $A \in M_{m \times n}(\mathbb{C})$, its pseudoinverse matrix A^+ exists. If $A = F \cdot G$ is a full rank decomposition, then

$$A^+ = G^+F^+ = G^*(GG^*)^{-1}(F^*F)^{-1}F^*.$$