Physics and Mathematics

UDC 519.612, 517.97

Tsukanova Alisa

Candidate of Physics and Mathematics,

Assistant of the Department of Mathematical Physics and Differential Equations

National Technical University of Ukraine

"Igor Sikorsky Kyiv Polytechnic Institute"

ONE VARIATIONAL METHOD OF SOLVING CLASSICAL PROBLEMS FROM LINEAR ALGEBRA

Summary. *The given paper presents some results of innovative optimization method of gradient descent with elements of classical Gaussian elimination method of solving systems of linear algebraic equations. These results have been obtained after testing our own program in «Visual Basic for Applications».*

Key words: *linear algebra, systems of linear algebraic equations, Gaussian elimination method, Kronecker-Capelli theorem, variational methods, optimization methods, gradient descent method, gradient, anti-gradient, minimizer, «Visual Basic for Applications».*

Introduction. Since the beginning of time, before to start implementation of some ideas that people had planned earlier, they have made some optimal, more or less, decision. For some time, this decision was made without any special analysis, it was based only on pure human experience. But over time it was no longer possible to realize this action without special mathematical methods that carry out global search for necessary optimum. Nowadays, in our crazy time of active computerization, these methods allow us to take, in fact, innovative look at various complicated problems of classical mathematics. The given paper presents

basic results of analysis of well-known classical Gauss method (or Gaussian elimination method of successive elimination of unknowns in solving some set of linear equations) and new optimization method of gradient descent for solving general systems of linear algebraic equations, obtained after testing special program, written in «Visual Basic for Applications».

Main part. An essence of our proposed numerical optimization method of gradient descent is that solving an arbitrary system of linear algebraic equations of the following matrix form

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

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ \vdots & & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix} = A_{m\times n},\ \vec{b} = \begin{pmatrix} b_1 \\ \dots \\ b_m \end{pmatrix} = B_{m\times 1},\ \vec{x} = \begin{pmatrix} x_1 \\ \dots \\ x_n \end{pmatrix} = X_{n\times 1},$$

is reduced to searching of minimizer, i.e. such vector $\vec{x}^* \in \mathbf{R}^n$, that

$$\min_{\vec{x}\in \mathbf{R}^n} f(\vec{x}) = f(\vec{x}^*),$$

for the next residual function

$$f(\vec{x}) = \left(A\vec{x} - \vec{b}, A\vec{x} - \vec{b}\right) = \left\|A\vec{x} - \vec{b}\right\|^2,\ A \in \mathbf{R}^{m\times n},\ \vec{b} \in \mathbf{R}^m,\ m \le n.$$

Formally, this method consists in iterative generation of some sequence of such points $\{\vec{x}_k\}|_{k\ge 0}$ (i.e. some descent trajectory, that is sometimes called a relaxation trajectory, that converges to our real solution $\vec{x}^*$), that

$$f(\vec{x}_{k+1}) \le f(\vec{x}_k),\ k \ge 0,$$

according to the following iterated scheme from two steps.

1. An arbitrary point is selected as an initial approximation $\vec{x}_0$.
2. All the next points $\vec{x}_{k+1}$, $k \ge 0$, are determined by the formula

$$\vec{x}_{k+1} = \vec{x}_k - \lambda_k \vec{g}_k,\ \lambda_k > 0,\ k \ge 0,$$

where $\vec{g}_k = grad\left(f(\vec{x}_k)\right)$, $k \ge 0$. At each step we make some move along the vector of anti-gradient $-grad\left(f(\vec{x}_k)\right)$, $k \ge 0$, in the direction of the fastest decrease of f, and, as a result, we get our necessary solution. Namely, if it turns out that the modulus of our anti-gradient is zero (more precisely, less than

predetermined accuracy), then we are at the minimum point we are looking for. If the criterion for the end of the iteration is not true (the modulus of our anti-gradient is more than predetermined accuracy), then we return to the first step, otherwise we return the exact value of $\vec{x}_{k+1}$, $k \geq 0$. Thus, the main idea of the proposed method is to optimize, that is, to move to minimum in direction of the fastest descent, and this direction is given by the vector of anti-gradient. In details: we choose some starting point in an arbitrary way, calculate the gradient of the considered function in this point, and take a small step in an opposite, anti-gradient, direction. As a result, we arrive to some point where the value of our function is less than the value in the previous point. At this new point we repeat the mentioned procedure: we again calculate the gradient of the function and take a step in an opposite direction. Continuing this process, we move towards some decreasing function. Our relaxation sequence is built, according to the following conclusions: due to special choice of direction of our movement, reducing the value of our function at each step, when moving from the first point $\vec{x}_k$, $k \geq 0$, to the second point $\vec{x}_{k+1}$, $k \geq 0$, we approach its minimum. Since at each step we move along the vector of anti-gradient, in direction of the fastest decrease of f, thus, as a result, we finally arrive to our necessary explicit solution.

In the mentioned above formula λ_k, $k \geq 0$, determines the distance between two points $\vec{x}_k$ and $\vec{x}_{k+1}$, $k \geq 0$. This distance is called a step of the proposed scheme. The main problem in rather complicated process of choosing this step λ_k, $k \geq 0$, is to ensure that the inequality

$$f(\vec{x}_{k+1}) \leq f(\vec{x}_k), k \geq 0,$$

is true. There are different ways to choose this step multiplier λ_k, $k \geq 0$. Depending on this, different variants of numerical optimization method of gradient descent can be obtained. We have considered the simplest method – the method with adaptive step selection [4]. This method contributes to accelerated approximation of our converging sequence $\{\vec{x}_k\}|_{k\geq 0}$ to necessary solution $\vec{x}^*$ we are looking for.

The program consists of two parts: classical (Gaussian) and innovative (optimization). The result of well-known Gauss method is some conclusion about solvability or unsolvability of the system under consideration, based on application of Kronecker-Capelli theorem from classical algebra to the reduced system [1 – 3]. The result of the second, optimization, part of the program is necessary solution, found with the help of proposed gradient descent method, running time of this method, as well as accuracy of all necessary calculations.

Conclusions. After testing various systems of linear algebraic equations with the help of our program, we have mentioned the next important facts. The Gaussian method algorithm gives results very quickly, in a fraction of seconds, while proposed optimization method produces the most accurate result in a short time not for every system. For some systems the program just loops during execution of this optimization algorithm. This fact experimentally proves that effectiveness of this descent method depends on value of its step: in optimal area significant «yawings» may occur due to need to correct step value of considered descent method. One schematic example of situation, when our method converges very slowly, is demonstrated on the picture below (fig. 1).

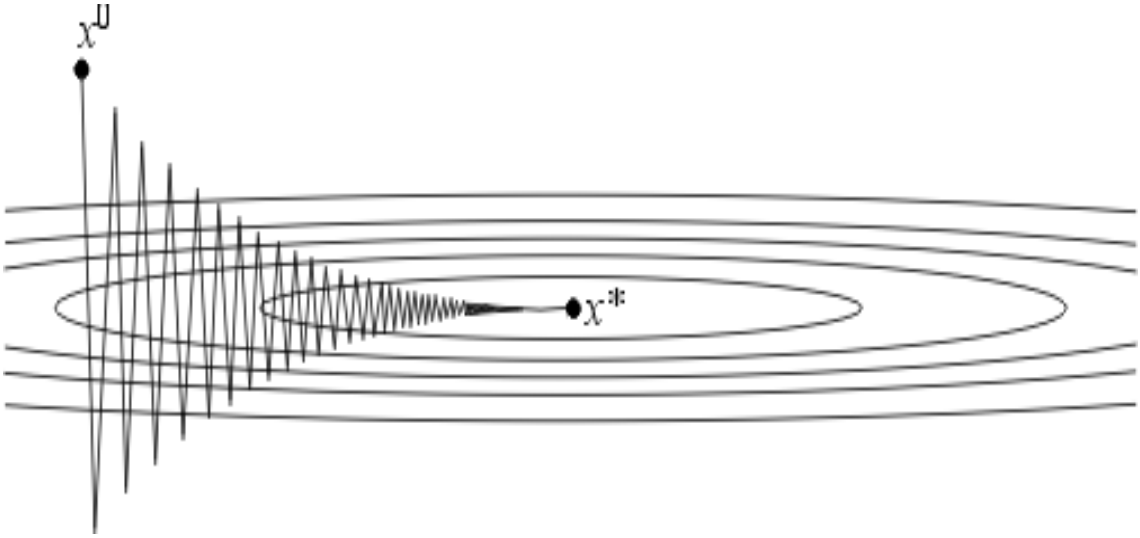

Fig. 1. Schematic example of situation, when our descent method converges very slowly

We suppose the considered software product to be an effective innovative educational and methodological support for studying the course of linear algebra and a cycle of various related disciplines.

References

1. Axler S. Linear Algebra Done Right (Undergraduate Texts in Mathematics). Springer, 2015. 357 p.
2. Shores T.S. Applied Linear Algebra and Matrix Analysis (Undergraduate Texts in Mathematics). Springer, 2018. 491 p.
3. Strang G. Introduction to Linear Algebra. Wellesley, MA: Wellesley Cambridge Press, 2009. 585 p.
4. Бейко І.В., Зінько П.М., Наконечний О. Г. Задачі, методи і алгоритми оптимізації: Навчальний посібник. Рівне: НУВГП, 2011. 624 с.