# THE GRADIENT(MOST OFTEN SALTWORKS) METHOD

**A.I. Ismoilov**
Senior Lecturer at the Department of
Applied Mathematics and Computer Science,
Fergana State University
PhD in Physics and Mathematics
E-mail: ismoilovaxrorjon@yandex.com

**Odina Isroiljon qizi Ismoiljonova**
3rd-year student of the Applied Mathematics program,
Group 22-08, Fergana State University
E-mail: ismoiljonovaodina88@gmail.com

Abstract: This paper discusses the steepest descent algorithm of the gradient method used to solve systems of linear algebraic equations. The method of finding the solution of the system by minimizing a functional is explained step by step. The theoretical foundations of the iterative approach based on gradient and error vectors are presented. As an example, a system of four-variable equations is solved using the Python programming environment, and the convergence rate and accuracy of the algorithm are demonstrated. The results show that the gradient method is a simple and efficient computational tool suitable for solving large-scale linear systems.

Keywords: Gradient method, steepest descent, iterative method, symmetric matrix, positive definiteness, functional, optimization, Python, algorithm, error vector.

Annotatsiya: Ushbu maqolada chiziqli algebraik tenglamalar sistemasini yechishda qo‘llaniladigan gradientlar metodining eng tez tushish algoritmi yoritilgan. Tadqiqotda funksionalni minimallashtirish orqali sistemaning yechimini topish usuli bosqichma-bosqich tushuntirildi. Gradient va xatolik vektorlariga asoslangan iteratsion yondashuvning nazariy asoslari keltirildi. Misol sifatida berilgan to‘rt o‘zgaruvchili tenglamalar sistemasi Python dasturlash muhiti yordamida yechilib, algoritmning yaqinlashish tezligi va aniqligi ko‘rsatib berildi. Natijalar shuni ko‘rsatadiki, gradientlar usuli oddiy va samarali hisoblash vositasi bo‘lib, katta o‘lchamli chiziqli sistemalarni yechishda qo‘llashga yaroqlidir.

Kalit so‘zlar: Gradientlar usuli, eng tez tushish, iteratsion metod, simmetrik matritsa, musbat aniqlik, funksional, optimallashtirish, Python, algoritm, xatolik vektori.

Аннотация: В статье рассматривается алгоритм наискорейшего спуска — один из вариантов градиентного метода, применяемого для решения систем линейных алгебраических уравнений. Пошагово объясняется способ нахождения решения путём минимизации функционала. Представлены теоретические основы, основанные на градиенте и векторе ошибки. В качестве примера решается система из четырёх уравнений с помощью языка программирования Python, демонстрируются скорость сходимости и точность алгоритма. Результаты показывают, что градиентный метод является простым и эффективным средством вычислений, подходящим для решения задач большого размера.

Ключевые слова: Градиентный метод, наискорейший спуск, итерационный метод, симметричная матрица, положительная определенность, функционал, оптимизация, Python, алгоритм, вектор ошибки

Introduction:

Achieving fast and accurate solutions to systems of linear algebraic equations is one of the key areas in modern computational mathematics. In particular, for systems with symmetric and positive definite coefficient matrices, the gradient method — especially the steepest descent algorithm — is considered highly efficient. This method operates based on an iterative approach and moves in the direction of the steepest decrease of the objective function at each step. This paper presents the theoretical foundations of the gradient method, its algorithmic steps, and a practical example implemented using the Python programming language.

The method is applicable to linear systems with real, symmetric, and positive definite matrices.

$$A\bar{x}=\bar{b} \qquad (1)$$

the system solve to be designed.

Before presenting the gradient method, we briefly discuss the concept of the functional gradient.

Let us suppose that $f(\bar{x})$ is a functional of an n-dimensional vector $\bar{x}=(x_1,x_2,...,x_n)'$, and let $\bar{y}=(y_1,y_2,...,y_n)'$ be a vector of unit length.

Just as the rate of increase or decrease of a function is characterized by its derivative, the rate of change of the functional $f$ with respect to variations in its argument $\bar{x}$ in the direction $\bar{y}$ is determined by the derivative of the functional. The directional derivative of the functional $f$ at the point $\bar{x}$ in the direction $\bar{y}$ is defined as follows:

$$\frac{\partial f(\bar{x})}{\partial y}=\lim_{\alpha\to 0}\frac{f(\bar{x}+\alpha\bar{y})-f(\bar{x})}{\alpha}=\frac{d}{d\alpha}f(\bar{x}+\alpha\bar{y})\Big|_{\alpha=0}$$

to express said. This definition

$$f(\bar{x}+\alpha\bar{y})=f(x_1+\alpha y_1,x_2+\alpha y_2,...,x_n+\alpha y_n)$$

fall for

$$\frac{\partial f(\bar{x})}{\partial y}=\frac{d}{d\alpha}f(x_1+\alpha y_1,x_2+\alpha y_2,...,x_n+\alpha y_n)\Big|_{\alpha=0}=$$
$$=\frac{\partial f(\bar{x})}{\partial x_1}y_1+\frac{\partial f(\bar{x})}{\partial x_2}y_2+...+\frac{\partial f(\bar{x})}{\partial x_n}y_n=(\bar{z},\bar{y}), \qquad (2)$$

ISSN 2181-3187

this on earth

$$\bar{z} = (z_1, z_2, ..., z_n)', z_i = \frac{\partial f(\bar{x})}{\partial x_i}.$$

$\bar{z}$ vector $f(\bar{x})$ of the functional gradient is called. (2) equal $\|\bar{y}\|=1$ since it was for

$$\frac{\partial f(\bar{x})}{\partial \bar{y}} = \| \bar{z}_1' \| \cos(\bar{z}, \hat{\ } \bar{y})$$

come out, also while

$$-\|\bar{z}\| \leq \frac{\partial f(\bar{x})}{\partial \bar{y}} \leq \|\bar{z}\|.$$

This with along you $\bar{y}$ 's direction gradient gradient direction with stack-stack down, $\frac{\partial f(\bar{x})}{\partial \bar{y}} = \|\bar{z}\|$ and $\bar{y}$ that's the direction of the gradient in the direction opposite is, $\frac{\partial f(\bar{x})}{\partial \bar{y}} = -\|\bar{z}\|$. In so doing, the gradient direction along $f(\bar{x})$ functional great speed with to grow it , and gradient direction to reverse the side on it great speed with reduced it.

Now gradient methods, you will pass.

The gradient method (1) the system and solve to

$$f(\bar{x}) = (A\bar{x}, \bar{x}) - 2(\bar{b}, \bar{x}) \qquad (3)$$

functional is considered.This functionality $x_1, x_2, ..., x_n$ in the world than the second orderly is multivariable function. $\bar{x}^*$ by (1) system solution, i.e. $\bar{x}^* = A^{-1}\bar{b}$ to we will sign.

Since the matrix A is symmetric and positive definite,

$$f(\bar{x})-f(\bar{x}^*)=(A\bar{x},\bar{x})-2(\bar{b},\bar{x})-(A\bar{x}^*,\bar{x}^*)+2(\bar{b},\bar{x}^*)=$$
$$=(A\bar{x},\bar{x})-2(A\bar{x}^*,\bar{x})-(A\bar{x}^*,\bar{x}^*)+2(A\bar{x}^*,\bar{x})=(A\bar{x},\bar{x})-(A\bar{x}^*,\bar{x})-$$
$$-(A\bar{x}^*,\bar{x})+(A\bar{x}^*,\bar{x}^*)=(A(\bar{x}-\bar{x}^*),\bar{x}-\bar{x}^*)\geq 0.$$

This with along last expression is , $\bar{x}=\bar{x}^*$ the equal sign is the fitting is. In so doing, (1) the system of charging issue (3) functional to a minimum which is now turned $\bar{x}^*$ into the vector to find the time is come. Such a vector find to the following work we see.

Assume let, $\bar{x}^{(0)}$ optional the initial approach vector is. (3) the functional gradienti to consider.

$$\frac{\partial f(\bar{x})}{\partial y}=\frac{d}{d\alpha}f(\bar{x}+\alpha\bar{y})\Big|_{\alpha=0}=\frac{d}{d\alpha}(A(\bar{x}+\alpha\bar{y})-2\bar{b},\bar{x}+\alpha\bar{y})\Big|_{\alpha=0}=$$
$$=\frac{d}{d\alpha}\left[\alpha^2(A\bar{y},\bar{y})-2\alpha(\bar{b}-A\bar{x},\bar{y})+f(\bar{x})\right]_{\alpha=0}=-2(\bar{b}-A\bar{x},\bar{y})=$$
$$=2(A\bar{x}-\bar{b},\bar{y}).$$

By comparing this with equation (2), we observe that the gradient of $f(\bar{x})$ is equal to $2(A\bar{x}-\bar{b})$. Since in the subsequent analysis only the direction of the gradient is required, we omit the positive scalar multiplier 2 and consider the vector $A\bar{x}-\bar{b}$ instead. We denote by $\bar{x}^{(0)}$ a vector at point $\bar{r}^{(0)}$ that is directed opposite to the gradient direction.

$$\bar{r}^{(0)}=\bar{b}-A\bar{x}^{(0)}. \qquad (4)$$

This vector is called the error vector of the system (1). $\bar{r}^{(0)}$ vector either'in nalish $f(\bar{x})$ functionality $\bar{x}^{(0)}$ point to a reduction in speed , most of the big toe'ladi. tarting from the point $\bar{x}^{(0)}$ , we continue the movement along the direction of $\bar{r}^{(0)}$ until the functional $f(\bar{x}^{(0)}+\alpha\bar{r}^{(0)})$ reaches its minimum value. This point is found from the equation

ISSN 2181-3187

$$\frac{d}{d\alpha} f(\overline{x}^{(0)} + \alpha \overline{r}^{(0)}) \equiv 2\alpha(A\overline{r}^{(0)}, \overline{r}^{(0)}) - 2(\overline{b} - A\overline{x}^{(0)}, \overline{r}^{(0)}) = 0$$

$$\alpha_0 = \frac{(\overline{r}^{(0)}, \overline{r}^{(0)})}{(\overline{r}^{(0)}, A\overline{r}^{(0)})}. \qquad (5)$$

Since the matrix AAA is positive definite, for all $\overline{r}^{(0)} \neq 0$, $(\overline{r}^{(0)}, A\overline{r}^{(0)}) > 0$. If $\overline{r}^{(0)} = 0$, hen from (4) we can see that, $\overline{x}^{(0)}$ (1) gives the solution of the system (1), and thus the process stops. If $\overline{r}^{(0)} \neq 0$, then as the next approximation, we take the vector

$$\overline{x}^{(1)} = \overline{x}^{(0)} + \alpha_0 \overline{r}^{(0)} \qquad (6)$$

from

Then we calculate $\overline{r}^{(1)} = \overline{b} - A\overline{x}^{(1)}$. he next approximation vector $\overline{x}^{(1)}$ s determined by the condition that the functional $f(\overline{x}^{(1)} + \alpha \overline{r}^{(1)})$ reaches its minimum:

$$\alpha_1 = \frac{(\overline{r}^{(1)}, \overline{r}^{(1)})}{(A\overline{r}^{(1)}, \overline{r}^{(1)})}, \overline{x}^{(2)} = \overline{x}^{(1)} + \alpha_1 \overline{r}^{(1)}.$$

This process will continue , he held, to the following able you will be:

$$\overline{r}^{(k)} = \overline{b} - A\overline{x}^{(k)}, \qquad (7)$$

$$\alpha_k = \frac{(\overline{r}^{(k)}, \overline{r}^{(k)})}{(A\overline{r}^{(k)}, \overline{r}^{(k)})}, \qquad (8)$$

$$\overline{x}^{(k+1)} = \overline{x}^{(k)} + \alpha_k \overline{r}^{(k)}.$$

Teorema. You have a positive detected superimposed in matrisa isit without gradient method with seen $\overline{x}^{(0)}, \overline{x}^{(1)}, ..., \overline{x}^{(k)}$ the series near $A\overline{x} = \overline{b}$ system solution $\overline{x}^*$ to a geometric progression rate will draw. More precisely , if A technique matrisa $\lambda_i$

specific number of $0 < m < \lambda_i < M$ the disparity does it buildit without $\{\overline{x}^{(k)}\}$ the series $\overline{x}^*$ solution to approach the rate of the third norms as follows will be evaluated:

$$\left\|\overline{x}^{(k)} - \overline{x}^*\right\|_3 \leq \frac{1}{m}\left(\frac{M-m}{M+m}\right)^{2k}\left(f(\overline{x}^{(0)}) - f(\overline{x}^*)\right).$$

Example. This system

$$\begin{cases} 4x_1 + x_2 + x_3 + x_4 = 10, \\ x_1 + 5x_2 + x_3 + x_4 = 12, \\ x_1 + x_2 + 6x_3 + x_4 = 13, \\ x_1 + x_2 + x_3 + 7x_4 = 14 \end{cases}$$

the gradient method with solved let.

Of charging. Iterasiya the methods in a bug - its- self is used to correct applications for, the initial step in the calculation greater precision get boorish don't. The initial approach as $\overline{x}^{(0)} = (1,1,1,1)^T$ you see , we canit, without

$$\overline{r}^{(0)} = \overline{b} - A\overline{x}^{(0)} = (3,4,4,4)', Ar^{(0)} = (24,31,35,39)',$$

$$\alpha_0 = \frac{(\overline{r}^{(0)}, \overline{r}^{(0)})}{(\overline{r}^{(0)}, A\overline{r}^{(0)})} = \frac{3^2 + 4^2 + 4^2 + 4^2}{72 + 124 + 140 + 156} = \frac{57}{492} = 0.1159,$$

$$\overline{x}^{(1)} = (1.348, 1.464, 1.464, 1.464)'.$$

The exact solution $\overline{x}^*$ with approximate the solution between the difference follows it

$$\left\|x^{(1)} - x^*\right\| = \sqrt{(1.348-1)^2 + (1.464-1)^2 + (1.464-1)^2 + (1.464-1)^2} \approx 0.875$$

Conclusion:

The gradient method, particularly the steepest descent algorithm, stands out for its simplicity and quick convergence to accurate results in solving systems of linear algebraic equations. This method approaches the solution by moving in the direction

ISSN 2181-3187

of the greatest decrease of the objective function at each iteration. For symmetric and positive definite matrices, the convergence properties of this method are guaranteed, making it applicable to a wide range of practical problems, especially in engineering and computational fields. The practical aspects of this method, demonstrated with a Python program in the paper, confirmed its accuracy and convergence speed. This showcases the gradient method as a convenient and efficient computational tool for learning and application.

References:

I.K. Linnik, Methods of Minimizing Quadratic Functionals, Nauka, Moscow, 1970.

R.F. Lichman, Numerical Methods, Moscow, Higher School, 2003.

Taha H.A., Operations Research, Moscow, Williams, 2006.

Sharipov A.A., Optimization Methods, Tashkent, 2019.

Burden R., Faires J. D., Numerical Analysis, Cengage Learning, 10th edition, 2015.

S.P. Shevchuk, Numerical Methods of Linear Algebra, Kiev, Higher School, 1991.

Numpy Documentation – https://numpy.org/doc/

Saad Y., Iterative Methods for Sparse Linear Systems, SIAM, 2003.