Modification of the Levenberg – Marquardt Algorithm for Solving Complex Computational Construction Problems

Valentina Kuzina 1, Alexander Koshev 1
1 Penza State University of Architecture and Construction, German Titov St., 28, Penza, 440028, Russia

kuzina@pguas.ru

Abstract. Systems of nonlinear algebraic equations have a special place in solving modeling and calculation problems in various areas of construction Sciences. These are tasks for calculating building structures, calculating distributed loads, mass and heat transfer tasks, calculating engineering networks, and many other computational and optimization tasks related to the construction industry. As is known, there are a sufficient number of fundamentally different methods for solving systems of nonlinear algebraic equations, which is related to the specifics of the systems, their degree of nonlinearity and conditionality. Non-linear ill-conditioned systems are particularly difficult; an erroneous choice of the method for solving such a system can lead to significantly distorted or simply incorrect results, which is absolutely unacceptable when performing the calculated stages of solving construction problems. In this paper, we propose to use a modification of the well-known Levenberg – Marquardt method developed by the authors, based on the regularization of the Jacobi matrix used in the classical Newton method, to solve complex, ill-conditioned systems of nonlinear equations. The modified method allows solving poorly defined systems and can significantly reduce the amount of calculations needed to ensure the required accuracy by reducing the number of iterative procedures. The paper presents a detailed description of the algorithm, given the solution of the model problem and, as an example the task of modeling and optimization of the port berthing facilities – mooring wall of massive masonry with various add-ons. When composing mathematical equations, the method of limit States was used, which is generally accepted for calculating such port hydraulic structures. The choice of the optimum scheme was carried out by minimizing the cost of structure during limit conditions the reliability of structures when using existing construction materials. The most significant geometric dimensions of the structure were selected using independent parameters to be optimized.

1. Introduction

Systems of nonlinear algebraic equations have a special place in solving problems of modeling and calculation in various areas of development of construction Sciences. These are problems of calculating building structures, calculations of distributed loads, problems of mass and heat transfer, calculations of engineering networks, and many other computational and optimization problems related to the construction sector.

Nonlinear systems either directly describe certain building processes, for example, such as the distribution of loads applied to elements of a building structure [1, 2] or in other tasks of building mechanics, or arise as a result of the transformation of complex integral-differential models, for
example, such as models describing the distribution of hydraulic characteristics in the piping systems of heating plants [3] or models for calculating the distribution of massive impurities in a confined space [2], and in other scientific tasks of construction and construction technologies.

As is known, there are a sufficient number of fundamentally different methods for solving systems of nonlinear algebraic equations, which is related to the specifics of the systems, their degree of nonlinearity and conditionality. Non-linear ill-conditioned systems are particularly difficult; an erroneous choice of the method for solving such a system can lead to significantly distorted or simply incorrect results, which is absolutely unacceptable when performing the calculated stages of solving construction problems.

In this paper, we propose to use a modification of the well-known Levenberg–Marquardt method [4] developed by the authors, based on the regularization of the Jacobi matrix used in the classical Newton method, to solve complex, ill-conditioned systems of nonlinear equations. The modified method allows solving poorly defined systems and can significantly reduce the amount of calculations needed to ensure the required accuracy by reducing the number of iterative procedures.

The paper presents a detailed description of the algorithm, given the solution of the model problem and, as an example the task of modeling and optimization of the port berthing facilities–mooring wall of massive masonry with various add-ons.

2. Modified Levenberg–Marquardt Algorithm

A system consisting of $n$ nonlinear algebraic equations with the number of unknowns equal to $n$, in vector form can be written as follows:

$$\vec{f}(\vec{x}) = 0; \quad \vec{f} = (f_1, f_2, ..., f_n)^T; \quad \vec{x} = (x_1, x_2, ..., x_n).$$

(1)

It is important to note that the problem of solving system (1) can be reduced to the task of minimizing the function:

$$r(\vec{x}) = \sum_{i=1}^{n} f_i^2(\vec{x}).$$

(2)

An iterative sequence converging to a solution is determined in the Levenberg–Marquardt method as follows:

$$\vec{x}^{(k+1)} = \vec{x}^{(k)} + \Delta \vec{x}^{(k)},$$

(3)

where $\Delta \vec{x}^{(k)}$ is the solution of a system of linear equations of the following form:

$$(M + \mu I_e) \Delta \vec{x}^{(k)} = -\vec{F};$$

(4)

$$M = W^T W, \quad \vec{F} = W^T \vec{f}.$$  

(5)

Here $W$ is the Jacobi matrix of system (1), $I_e$ is the identity matrix, $\mu$ is a numerical parameter.

The introduction of the parameter $\mu$ allows you to expand the convergence area of the Newton–Gauss algorithm and, by shifting to the shortest descent line, prevent divergence. This is especially important when the Jacobi matrix of the system of equations becomes ill-conditioned at some stages of the iterative procedure for calculating the next approximation to the solution.
Note that for \( \mu \geq 0 \), \( \Delta \overline{x}^{(k)}(\mu) \approx -M^{-1} \overline{F} \), which is close to the “step” of the solution by the Newton–Gauss method. With a sufficiently large value of \( \mu \), \( \Delta \overline{x}^{(k)}(\mu) \approx -\overline{F} / \mu \), which corresponds to the increment of the step along the direction of the steepest descent. The proof of the convergence of the iterative scheme (2–4) under certain restrictions on the matrix \( M \) and the parameter \( \mu \) is given, for example, in [5]. However, these limitations cannot be verified a priori, as a rule. Therefore, existing methods for choosing a parameter are based on reasoning that does not use the conditions of the convergence theorem. Levenberg and Marquardt proposed various ways of choosing \( \mu \). All authors sought to make the method monotonously convergent. It seems interesting to us not to require the monotonicity of the decrease of the function \( r(\overline{x}) \) at each step, but to recalculate the factor \( \mu \) at each iteration, using the information from previous iterations. The following algorithm for choosing the parameter \( \mu \) is proposed.

From the expansion of \( r(\overline{x}) \) into a Taylor series at a point \( \overline{x}^{(k)} \), we have:

\[
r(\overline{x}^{(k)} + \Delta \overline{x}) - r(\overline{x}^{(k)}) = 2(\Delta \overline{x}^{(k)})^T \overline{F}^{(k)} + \omega,
\]

where \( \omega \) are the expansion members whose order is higher than the first one.

Let’s make the ratio:

\[
R = \frac{r^{(k)}(\overline{x}^{(k)} + \Delta \overline{x}) - r^{(k)}(\overline{x}^{(k)})}{2(\Delta \overline{x})^T \overline{F}^{(k)}}.
\]

There can be two cases.

1. \( r(\overline{x}^{(k)} + \Delta \overline{x}) - r(\overline{x}^{(k)}) < 0; \)
2. \( r(\overline{x}^{(k)} + \Delta \overline{x}) - r(\overline{x}^{(k)}) \geq 0. \)

Let us consider the first case.

We can highlight the situation when \( R \geq 1 \). In such areas, the function is close to linear, so it is advisable to carry out a shift in the parameter \( \mu \) to the steepest descent method. This requires \( \mu \) to be increased. It is proposed to accept.

If \( \sigma_2 < R \leq 1 + \sigma_3 \), then \( \mu^{(k+1)} = \mu^{(k)} \left( 1 + \gamma |R - \sigma_2| \max \{R - (1 - \sigma_3)\} \right) \); \( \sigma_2 = 0.8; \ \sigma_3 = 0.4; \ \gamma = 1 \).

(Here and in the following formulas, the constants \( \gamma \) and \( \sigma_i \) are chosen on the basis of a numerical solution of a large number of problems.) In other cases, it is advisable either to keep \( \mu \) parameter unchanged or to reduce it. The following formulas were obtained experimentally:

If \( R < \sigma_1 \) (\( \sigma_1 = 0.4 \)), then \( \mu^{(k+1)} = \frac{\mu^{(k)}}{1 + \gamma |R|} \).

If \( R \geq 2 - \sigma_4 \) (\( \sigma_4 = 0.4 \)), \( \mu^{(k+1)} = \frac{\mu^{(k)}}{1 + \gamma |R - 2 - \sigma_4|} \).

If \( \sigma_1 \leq R \geq \sigma_2 \) or \( 1 + \sigma_3 \leq R \geq 2 - \sigma_4 \), then \( \mu^{(k+1)} = \mu^{(k)} \).
In the second case, it is advisable to decrease $\mu$ by the following formula: $\mu^{(k+1)} = \mu^{(k)} \frac{1}{1 + |R|}$. In addition, the parameter $\mu$ at each step is suggested to be multiplied by a certain weighting factor $\varphi(x^{(k+1)}, x^{(k)})$. We took: $\varphi(x^{(k+1)}, x^{(k)}) = \|x^{(k+1)} - x^{(k)}\|$.

Using a weighting factor allows you to increase the rate of convergence of the method. For example, when the gradient of a function is close to zero and the value of the function itself is far from the minimum, the method of steepest descent works poorly, so it is necessary to make a shift to the Newton – Gauss method, and since $\Delta x^{(k)}$ is quite large ($\Delta x^{(k)} > 1$), and due to the weight, you can increase the parameter $\mu$ and make a large step, that is, it will increase the shift to the Newton – Gauss method. On the other hand, in areas where the function is close to linear, the steps at many iterations are quite large ($|\Delta x^{(k)}| > 1$), and due to the weight, you can increase the parameter $\mu$ and make a large step, that is, it will increase the shift to the Newton – Gauss method.

The parameter is constrained $\mu > \mu_c$. This restriction regularizes the matrix $M$ and allows one to find the inverse of it even in the case of poor conditionality of the matrix. The rate of convergence is also affected by the initial value of the parameter $\mu_0$. If the initial approximation $x^{(0)}$ is far from the solution $x^*$, then we can take $\mu_0 > 1$, if close to, then $\mu_0 < 1$. We took $\mu_0 = 1$ in solving practical problems. In the algorithm, an automatic transition to the Newton method is provided for the refinement of the solution.

3. Test example of numerical implementation of the method

The efficiency of the algorithm is confirmed by the solution of a number of different systems of nonlinear algebraic equations. As you know, the numerical solution of control examples by any method allows you to judge the scope of its applicability. In this paper, we give an example of solving a system of equations obtained by taking partial derivatives of a well-known Rosenbrock test function [6, p. 118]:

$$
\begin{align*}
\begin{cases}
f_1 = 200(x_2 - x_1^2); \\
f_2 = -400(x_2 - x_1^2)x_1 - 2(1 - x_1).
\end{cases}
\end{align*}
$$

(8)

The initial approximation is chosen from the same test example: $x_1 = -1.2; \quad x_2 = 1$.

We note that the condition number of the Jacobi matrix of this system allows us to consider this matrix as ill-conditioned. This circumstance does not allow us to obtain a solution to the system, for example, by the Newton method. A computer procedure written according to the proposed algorithm, for 12 iterations, finds the exact solution to the system. Note that the modified Newton – Raphson method with automatic parameter selection solves the system in 150 iterations. Figure 1 shows a view of the surface described by the Rosenbrock function constructed in the MathCAD system.
5

Figure 1. The Rosenbrock function constructed in the MathCAD system

It is easy to see that the function under study near the minimum point is close to linear in both coordinates, which makes it much more difficult to find the solution point by Newton's method.

4. Calculation of optimal parameters of a berth port facility

We used the modified Levenberg – Marquardt algorithm to solve the problem of modeling and optimizing a port mooring structure – a mooring wall made of massive masonry with various types of superstructures. The main problem in the design and construction of such a massive structure is the high material consumption, which causes its significant cost. In mathematical modeling of the technological and technical formulation of the problem, the method of limit States, commonly used for calculating port hydraulic structures, was used. The choice of the optimal scheme was made from the condition of the minimum cost of the structure when the restrictions set by the construction conditions were met [1].

The most significant geometric dimensions of the structure were selected as independent parameters to be optimized: the width of the underwater part $x_1$ and the width of the superstructure $x_2$. In accordance with the current regulatory documents, the following restrictions were introduced: 1) stability of the superstructure on the basis of the scheme of flat-shift structures on the diagram of plane shear on the contact structure on the bed; 2) limit values of the compressive and tensile stresses along the contact add-on – base; the magnitude of the compressive and tensile stresses at the contact a construction basis.

The mathematical formulation of the problem is as follows: minimize the total cost of the structure $S(x_1, x_2) \rightarrow \min$ with restrictions $f_i(x_1, x_2) \leq 0$, $i = 1, \ldots, m$. Note that the total cost of construction is the cost of each of the elements of a design: cost, cost of add-ons and the cost of backfill. Type of functions $f_i(x_1, x_2), \ldots, f_m(x_1, x_2)$ is extremely bulky, and cannot be presented in analytical form. The functions $S(x_1, x_2), f_i(x_1, x_2)$ are not linear with respect to the coordinates $x_1, x_2$ and, therefore, the problem to be solved is a nonlinear programming problem. To form an objective function

$$f(x_1, x_2) := \left[200 \cdot (x_2 - x_1^2) \right]^2 + \left[-400 \cdot (x_2 - x_1) \cdot x_1 - 2 \cdot (1 - x_1)\right]^2$$
that is subject to minimization, the method of penalty functions was used, reducing the problem to an unconditional minimization of a function of the following type:

\[ F(x_1, x_2) = S(x_1, x_2) + t_k \sum [f_i(x_1, x_2)]^2. \]  

(9)

To implement the above algorithm, the problem was reduced to a system of nonlinear algebraic equations obtained by equating to zero the partial derivatives of the objective function with variables \( x_1, x_2 \). Note that the function \( S(x_1, x_2) \) is strictly increasing relative to variables, and, therefore, it is advisable to take the initial values \( x_1^0, x_2^0 \) as minimal as possible for this structure. In this case, as a rule, the initial approximation does not fall within the scope of restrictions, which was the main criterion for choosing a method for solving the problem of the “outer point method”.

5. Conclusions

To note in conclusion that the use of mathematical modeling and calculations allows us to obtain significant results in engineering structures. In particular, the application of the described optimization and calculation methods in designing is not only an increase in labor productivity, but also a decrease in construction costs by at least 10% [1].

The efficiency of the algorithm is confirmed by the solution of a number of various test and practically important minimization problems that arise during mathematical modeling of various objects and processes.

References

[1] A.N. Koshev, V.V. Kuzina, “Mathematical modeling in tasks of construction,” Procedia Engineering, 161, “Elsevier BV”, pp. 1874–1878, 2016. Doi: 10.1016 / j.proeng.2016.08.726.

[2] A.N. Koshev, V.V. Kuzina, “Mathematical modeling as an element of management in the tasks of the construction industry,” Regional architecture and construction, no. 2 (13). Scientific and technical magazine, Penza, PGUAS, pp. 167–176, 2012.

[3] R.F. Fayzullin, “On the solution of systems of non-linear algebraic equations of a special form,” Mathematical structures and modeling, no. 7, pp. 79-84, 2001.

[4] B.D. Marquardt, “An algorithm for least squares estimation,” An apple math, vol. 11, p. 431, 1969.

[5] J. Ortega, V. Reinboldt, “Iterative methods for solving nonlinear systems of equations with many unknowns,” Moscow: Mir, 558 p., 1975.

[6] H. Rosenbrock, S. Story, “Computational methods for chemical engineers,” Moscow: Mir, 443 p., 1968.