On Some Iterative-Difference Methods for Solving Nonlinear Least-Square Problem

Stepan Shachno, Oleksandra Gnatyshyn

Faculty of Applied Mathematics and Informatics, Lviv State University of Ivan Franko, 290602, Ukraina, Lviv, Universytetska,1, tel.(0322)794-391, fax (0322)72-36-76,e-mail kom@franko.lviv.ua
Key words. Difference method, nonlinear least-square problem, divided difference, convergence, rate of convergence.

Nonlinear least-square problems often arise when estimating the parameters of functional dependence on the data of experiments or statistical samples. Such data can be obtained, for instance, with the radiolocation trajectory measuring when the flight of airplanes is studied. Solving the nonlinear systems where the amount of nonlinear bonds is larger then the degrees of freedom is also possible only in the sense of the least-squares.

We consider the nonlinear least-square problem:

$$\text{find} \qquad \min_{x \in R^n} \frac{1}{2} F(x)^T F(x) = \frac{1}{2} \sum_{i=1}^{m} F_i^2(x), \qquad (1)$$

where function $F: R^n \to R^m$ - is nonlinear on x, $m \geq n.$. For solving (1) the Gauss-Newton method and modifications [1,4,6,7] are often used. But all of those methods use the operator of derivative $F'(x)$ in the iterative formulas. The calculation of $F'(x)$ is undesirable when the analytical expression of $F'(x)$ is complicated or when we have not the analytical expression of function $F(x)$ but only the computational algorithm.

In this paper we propose the iterative process for solving problem (1) which does not require the evaluation of $F'(x)$:

$$x_{k+1} = x_k - \left[F(u_k, v_k)^T F(u_k, v_k) \right]^{-1} F(u_k, v_k) F(x_k), \qquad k = 0,1,\ldots \qquad (2)$$

where $F(u,v)$ – is the first divided difference for function $F(x)$, which is evaluated by the formula [see 3]:

$$F_{ij}(u,v) = \frac{F_i(v^1,\ldots,v^{j-1},u^j,u^{j+1},\ldots,u^n) - F_i(v^1,\ldots,v^{j-1},v^j,u^{j+1},\ldots,u^n)}{u^j - v^j}, \qquad u_j \neq v_j \qquad (3)$$

where $F_{ij}(u,v)$ denotes the component of the matrix $F(u,v)$ in row i and column j, moreover $F(u,u) = F'(u)$ [8]. Varying choice of u_k, v_k we obtain different analogies of the Gauss-Newton method with the different rates of convergence. When $u_k = v_k = x_k$ method (2) coincides with the known Gauss-Newton method [4].

Let us consider some cases of choice u_k, v_k.

1) Let $u_k = 2x_k - x_{k-1}$, $v_k = x_{k-1}$. Then formula (2) can be written under the form

$$x_{k+1} = x_k - \left[F(2x_k - x_{k-1}, x_k)^T F(2x_k - x_{k-1}, x_k)\right]^{-1} F(2x_k - x_{k-1}, x_k)^T F(x_k), \quad k = 1,2,\dots \tag{4}$$

When $m = n$ method (4) transforms into the method of linear interpolation for nonlinear functional equations [5]. We present the following theorem, which provides the assumptions and rate of convergence of method (4).

Theorem. 1) Let for the initial approximation x_0 there exists the inverse operation $\Gamma_0 = \left[F'(x_0)^T F'(x_0)\right]^{-1}$;

2) $\|F'(x_0)\| \le \alpha$;

3) in the region $R = \{x: \|x - x_*\| \le \xi\}$, $\xi = const > 0$, which contains the unique solution of problem (1),

$$\left\|\left(F'(x) - F'(x_*)\right)^T F(x_*)\right\| \le \sigma \|x - x_*\|; \tag{5}$$

$$\|F''(x)\| \le M; \qquad \|F'''(x)\| \le N;$$

4) $\|F(x_*)\| \le \eta$.

Then for x_0, x_1, sufficiently close to the solution x_* of problem (1), the iterative process (4) converges to the solution with the order which is characterized by the inequality

$$\|x_{n+1} - x_*\| \le C_1 \|x_n - x_*\| + C_2 \|x_n - x_*\|^2 , \tag{6}$$

where C_1, C_2 - are some limited constants.

The proof of theorem is bulky and so we don't give it here, but we would like to note one estimate, obtained in the proof:

$$\| x_{n+1} - x_* \| \le \| x_n - x_* \| \, \|A_n^{-1}\| \left[\left(\tfrac{N}{6} C \| x_n - x_* \| + \alpha\right) \times \left(\tfrac{N}{6} C + \tfrac{M}{2}\right) \| x_n - x_* \| + \tfrac{N}{6}\eta + \sigma \right], \tag{7}$$

From (7) it follows that in the case of null residual when $F(x_*) = 0$ and $F(x)$ is a linear function (then parameter σ=0) the rate of convergence of method (2) is quadratic . Thus, the iterative process (2) converges with the same rate as Gauss-Newton method, but it's advantage is that method (2) doesn't require the evaluation of $F'(x)$.

2) Let $u_k = x_k$, $v_k = x_{k-1}$. Then formula (2) is written under the form

$$x_{k+1} = x_k - \left[F(x_k, x_{k-1})^T F(x_k, x_{k-1})\right]^{-1} F(x_k, x_{k-1})^T F(x_k), \quad k = 1,2,\dots \tag{8}$$

Method (8) converges with the lower order than method (4) . It is easy to prove that in the case of null residual method (8) converges with the order 1.618... But method (8) requires evaluation of vector-function $F(x)$ in one point, meanwhile method (4) requires evaluation of $F(x)$ in two points per iteration. Thus, considering the total amount of evaluations for some problems method (8) is not worse than method (4) and Gauss-Newton method.

Taking into consideration that methods (4) and (8) are simple in realization and require only evaluation of $F(x)$ they can be broadly used in practice.

The numerical experiments confirm the high efficiency of the difference methods (4) and (8). The results of the calculations will be discussed in report.

Note that difference methods for the problems of unconstrained minimization problems are presented in [2].

REFERENCES

1. Бартіш М.Я., Шахно С.М. Деякі методи розв'язування нелінійної задачі найменших квадратів. //Вісн. Львів. ун-ту. Сер. мех.-мат. 1993. Вип.39. С.3-7
2. Bartish M.Ya.,Shachno S.M. On the iterative Steffensen Type Methods//Zeitschrift fur Angewandte
   Mathematik und Mechanik, Berlin 76 (1996),S1, 351-352.
3. Васильев Ф.П. Численные методы решения экстремальных задач. Москва. Наука. 1988. 550с.
4. Дэннис Дж., мл., Шнабель Р. Численные методы безусловной оптимизации и решения нелинейных уравнений. М., Мир. 1988.440 с.
5. Курчатов В.А. Об одном методе линейной интерполяции решения функциональных уравнений. //Докл. АН СССР. Сер. Математика. Физика. 1971. 198. №3. С. 524-526.
6. Ортега Дж., Рейнболдт В. Итерационные методы решения нелинейных систем уравнений со многими неизвестными. М. Мир. 1975. 558 с
7. Shachno S.M. Numerical methods for solving nonlinear least squares problems. 9th Conference of the European Consortium for Mathematics in Industry. Technical University of Denmark Lyngby
   //Copenhagen, Denmark, June 25-29, 1996. Book of abstracts. Pp.543-545.
8. Ульм С. Об обобщенных разделенных разностях. I, II. Известия АН ЭССР. Физика. Математика, 1967.