Jaromír Janko

Solution of non-linear equations systems by Newton's method and the gradients method

Aplikace matematiky, Vol. 10 (1965), No. 3, 230–234

Persistent URL: http://dml.cz/dmlcz/102955

SOLUTION OF NONLINEAR EQUATIONS SYSTEMS BY NEWTON'S METHOD AND THE GRADIENTS METHOD

JAROMÍR JANKO

(to topic a)

For the solution of nonlinear equations systems

$$F(X) = 0\,,$$

where F, X are vectors of components f_i, x_i, $i = 1, 2, \ldots, n$, a number of methods has been elaborated. On the whole these methods do not perform a direct computation, but on the basis of root approximation (zero approximation) gained by experience, from physical reasons or otherwise, the residues of the equations are computed, and it is then attempted to vary the root approximation so that all residues move towards zero. This new approximation is then taken as the starting point of a second analogous step; such steps are performed successively until all residues decrease under chosen limits. This is considered to be the physically appropriate solution. Since such problems often occur in our Institute, empirical tests were performed; their aim was the selection of an optimal method for the solution of the simultaneous nonlinear equations.

Two known methods were treated: Newton's method uses the Taylor expansion and by solving the non-homogeneous linear equations obtains the root correction; and the method of gradients which solves another equivalent system of equations and obtains new approximations of roots by means of matrix computations. Both methods were also modified in the sense that in all steps the matrix of the partial derivatives was left constant from the 0-th approximation, so that an enormous amount of computatial effort is saved; of course, usually the number of necessary iterations is increased, so that it is not evident a priori whether the modification is or not an improvement of the method. The convenience of the method used was assessed directly from the most practical parameter, viz. the time needed for machine computation. This may be given directly, or estimated from the time of duration of one approximate cycle and the total number of cycles necessary for obtaining the appropriate solution.

On the digital computer LGP 30 the following two systems were solved: the system consisting of the three equations

$$(1) \qquad \begin{aligned} y_1 + y_1^2 - a_2 y_2 y_3 - a_4 &= 0\,, \\ y_2 - y_2^2 + a_3 y_1 y_3 + a_5 &= 0\,, \\ y_3 + y_3^2 + a_2 y_1 y_2 - a_6 &= 0 \end{aligned}$$

and the system consisting of the six equations

$$(2) \qquad \begin{aligned} \frac{y_5 + a_3 y_6}{y_2 + a_3 y_3} \left(- a_1 + a_3 y_1\right) + a_2 - a_3 y_4 &= 0\,, \\ \frac{y_5 + a_6 y_6}{y_2 + a_6 y_3} \left(- a_4 + a_6 y_1\right) + a_5 - a_6 y_4 &= 0\,, \\ \frac{y_5 + a_9 y_6}{y_2 + a_9 y_3} \left(- a_7 + a_9 y_1\right) + a_8 - a_9 y_4 &= 0\,, \\ \frac{y_4 + a_{10} y_6}{y_1 + a_{10} y_3} \left(- a_1 + a_{10} y_2\right) + a_2 - a_{10} y_5 &= 0\,, \\ \frac{y_4 + a_{11} y_6}{y_1 + a_{11} y_3} \left(- a_4 + a_{11} y_2\right) + a_5 - a_{11} y_5 &= 0\,, \\ \frac{y_4 + a_{12} y_6}{y_1 + a_{12} y_3} \left(- a_7 + a_{12} y_2\right) + a_8 - a_{12} y_5 &= 0 \end{aligned}$$

where a_i are constants, y_i unknowns. The results of solution are summarised in following tables and diagrams.

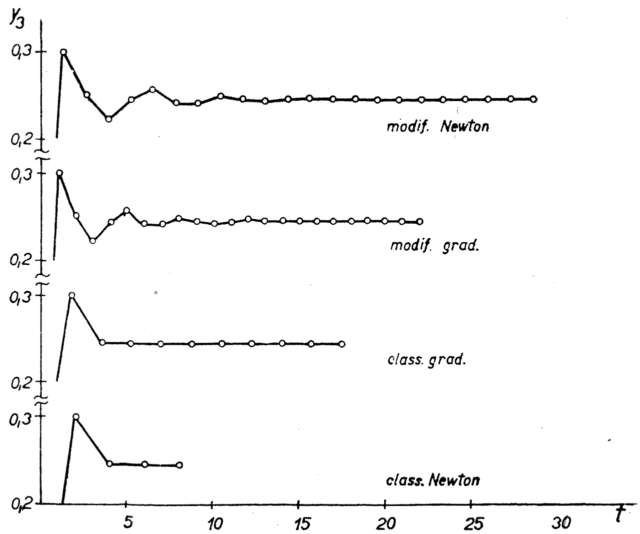

Fig. 1.

On fig. 1 there is a graphical representation of the results of the solution of system (1) with the initial approximation 0, 0, 0.

On the horizontal axis the solution time in minutes is plotted; the individual points mean the iteration steps. On the vertical axis the value of the third root y_3 is plotted.

Table I

Method	Modification	0-th approx.	Roots	Residues	Number of iterations	Computing time per iteration [min]	Total computing time [min]	Remarks, Order
Newton	classical	0 0 0	0,01282412 −0,1778006 0,2446881	0 0 0	4	2	8	optimal 1.
Grad.	classical	0 0 0	0,01282456 −0,1778004 0,2446883	0 0 0	10	1,75	17,5	2.
Newton	modified	0 0 0	0,01282784 −0,1778697 0,2447190	10^{-5} 10^{-5} 10^{-5}	22	1,3	28,6	4.
Grad.	modified	0 0 0	0,01282784 −0,1778697 0,2447190	10^{-5} 10^{-5} 10^{-5}	22	1	22	3.
Newton	classical	0,8 3 −5	0,05788025 0,01562790 −1,240399	0 0 0	7	2	14	optimal 1.
Grad.	classical	0,8 3 −5	0,05793612 0,01560795 −1,240848	10^{-4} 10^{-4} 10^{-4}	26	1,75	45,5	2.
Newton	modified	0,8 3 −5	0,06017865 0,01768135 −1,246166	$8 . 10^{-3}$ $8 . 10^{-3}$ $8 . 10^{-3}$	35	1,3	45,5	4.
Grad.	modified	0,8 3 −5	0,05765095 0,02121230 −1,243565	$8 . 10^{-3}$ $8 . 10^{-3}$ $8 . 10^{-3}$	36	1	36	3.
Grad.	modified	0,8 3 −5	0,05788103 0,01562717 −1,240394	10^{-6} 10^{-6} 10^{-6}	106	1	106	

The superiority of the classical Newton method is immediately apparent; although one step lasts two full minutes, the solution is reached in the shortest time. In the other methods and modifications the individual steps are faster, but the number of necessary iterations is too large. By plotting the other roots, possibly with different initial approximations, very similar diagrams are obtained.

From table I, where the results of solution of system (1) are summarised, there follow these conclusions: The classical method of Newton is optimal, the classical method of gradients is twice slower. There then follow the two modificated

Table II

Method	Modification	0-th approx.	Roots	Residues	Number of iterations	Computing time per iteration [min]	Total computing time [min]	Remarks, Order
Newton	modified	1 15 -4 12 -4 0	1,010202 15,02380 -4,098622 12,00001 -3,999999 -0,111461	0 0 0 $2 . 10^{-5}$ $3 . 10^{-5}$ 0	3	5,25	15,75	2.
Grad.	modified	”				6		diverges
Newton	classical	”	1,010201 15,02382 -4,098642 12,00001 -4,000003 10^{-7}	0 0 0 10^{-5} -10^{-5}	2	6	12	optimal 1.
Grad.	classical	”				3,25		diverges
Newton	classical	1,5 8 -3 18 -2 0	1,010200 15,02382 -4,098637 12,00002 -3,999975 $-6 . 10^{-5}$	0 0 0 10^{-5} $2 . 10^{-5}$ $3 . 10^{-5}$	10	5,25	52,5	optimal
Newton	modified	”						diverges
Grad.	modified	”						diverges
Grad.	classical	”						diverges

methods — here the phenomenon occured, that even though the time duration of one step was shortened appreciably, the number of needed steps increased so much that the modification of the method became valueless.

The modified methods follow in reverse order — the method of gradients is a little quicker than Newton's. The second half of this table corresponds to a different choice of initial approximation (and other root approximations are obtained) and verifies the above classification of methods. Table II, applying to system (2), is in accordance with these results; however, the method of gradients does not converge in this case.

Jaromír Janko, Výzkumný ústav zvukové, obrazové a reprodukční techniky, Plzeňská 66, Praha 5, ČSSR.