NUMERICAL METHOD FOR LEAST SQUARE SOLVING OF NONLINEAR EQUATIONS

GY. POPPER

Department of Civil Engineering Mechanics,
Technical University, H-1521 Budapest

Received July 10, 1984

Presented by Prof. Dr. S. Kaliszky

A new numerical method for the least square solutions of nonlinear equation systems has been suggested. The method does not need the calculation of the derivatives.

It is well-known, that if A is a (real) $m \times n$ matrix with $m > n$ and b denotes a (real) vector with m elements, the set of linear equations

$$\mathrm{A}x = b \tag{1}$$

is "overdetermined" and in general its solution in the classical sense does not exist. The generalized solution — or the least square solution — of Eq. (1) is defined by the solution of the linear least square problem

$$||b - \mathrm{A}x||_2 = \min. \tag{2}$$

Since $||b - \mathrm{A}x||_2^2 = (b - \mathrm{A}x)^T(b - \mathrm{A}x)$, the condition for minimum $\frac{\mathrm{d}}{\mathrm{d}x}||b - \mathrm{A}x||_2^2 = 0$ leads to

$$\mathrm{A}^T\mathrm{A}x = \mathrm{A}^T b. \tag{3}$$

Thus the solution of system (1) can be obtained via solving system (3).

If the rank of A is less than n, then there is no unique solution. Thus we require amongst all x which satisfy (2) that $||x||_2 = \min.$ and this solution is unique (see e.g. [2]).

The explicit formation of the symmetric matrix $\mathrm{A}^T\mathrm{A}$ involves unnecessary numerical inaccuracy and can be avoided e.g. using the method given by Businger and Golub in [1].

Let us consider now the problem of solving the nonlinear system of equations

$$f_i(x_1, \ldots, x_n) = 0, \qquad i = 1, \ldots, m$$

with $m > n$, which we usually write in the vector form

$$f(x) = O, \tag{4}$$

where

$$f(x) = [f_1(x), \ldots, f_m(x)]^T.$$

Similarly to the linear problem (1), the solution of nonlinear equations (4) in general also does not exist in the classical sense and we will define the generalized solution of Eq. (4) by the solution of the nonlinear least square problem

$$\|f(x)\|_2 = \min.$$

One of the basic iteration procedures for approximating a solution of Eq. (4) for $m = n$ is the secant method [3]: Starting with $n + 1$ approximating vectors (with n elements)

$$x^{(1)}, \ldots, x^{(n)}, x^{(n+1)},$$

the further approximation is given by the linear combination

$$x^{(n+2)} = q_1 x^{(1)} + \ldots + q_n x^{(n)} + q_{n+1} x^{(n+1)},$$

where the weighting coefficients $q_1, \ldots, q_{n+1}$ are computed by solving the system of linear equations

$$\begin{bmatrix} f(x^{(1)}) & , \ldots, f(x^{(n)}), & f(x^{(n+1)}) \\ 1, & \ldots, 1 \quad , & 1 \end{bmatrix} \begin{bmatrix} q_1 \\ \vdots \\ q_n \\ q_{n+1} \end{bmatrix} = \begin{bmatrix} 0 \\ \hline 1 \end{bmatrix} \tag{5}$$

Now we have $n + 2$ approximating vectors

$$x^{(1)}, \ldots, x^{(n+1)}, x^{(n+2)}.$$

Dropping the one for which norm

$$\|f(x)\|_2$$

is the maximal one, we obtain again $n + 1$ starting vectors necessary for the next step of the iteration process.

The secant method can be extended to the case $m > n$. Then the matrix of coefficients in Eq. (5) is of type $(m + 1)\times(n + 1)$, and the previous reduction of Eq. (1) to Eq. (3) can be applied to the system given by the Eq. (5).

To guarantee the exact satisfaction of the relation

$$q_1 + \ldots + q_n + q_{n+1} = 1$$

[which is expressed by the last equation in (5)], the system (5) must be rewritten in the form

$$[f(x^{(n+1)}) - f(x^{(1)}), f(x^{(n+1)}) - f(x^{(2)}), \ldots, f(x^{(n+1)}) - f(x^{(n)})] \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_n \end{bmatrix} = f(x^{(n+1)}) \tag{6}$$

with

$$q_{n+1} = 1 - q_1 - q_2 - \ldots - q_n$$

and the reduction will be applied only to the system (6).

Note: If the suggested method is applied to linear equation systems, the least square solution is obtained by the first step of the iteration.

Example

Let us consider the system of nonlinear equations

$$f(x) = \begin{bmatrix} x_1 - 1 \\ x_2 - 1 \\ x_1^2 + x_2 - 1 \end{bmatrix} = 0,$$

with exact least square solution (rounded up to five decimals)

$$x_1 = 0.68233, \quad x_2 = 0.76721$$

with

$$\|f(x)\|_2^2 = 0.20923.$$

Starting with $n+1=3$ initial approximating vectors

$$x^{(1)} = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad x^{(2)} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad x^{(3)} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

with

$$\|f(x^{(i)})\|_2 = 1, \quad i = 1, 2, 3$$

the results of the computation are as follows:

iter. step	starting vectors	computed vector: $x^{(i)}$	$\|f(x)^{(i)}\|_2^2$
1.	$x^{(1)}, x^{(2)}, x^{(3)}$	$x^{(4)} = [0.66667, 0.66667]^T$	0.23457
2.	$x^{(2)}, x^{(3)}, x^{(4)}$	$x^{(5)} = [0.79070, 0.65116]^T$	0.24187
3.	$x^{(3)}, x^{(4)}, x^{(5)}$	$x^{(6)} = [0.67822, 0.74185]^T$	0.21092
4.	$x^{(4)}, x^{(5)}, x^{(6)}$	$x^{(7)} = [0.67086, 0.77756]^T$	0.20962
5.	$x^{(4)}, x^{(6)}, x^{(7)}$	$x^{(8)} = [0.68448, 0.76584]^T$	0.20930

As it appears already from the first five steps of the iteration, the convergence is rapid.

References

1. BUSINGER P.—GOLUB G. H.: Linear Least Squares Solutions by Householder Transformations, Numerische Mathematik, *7*, 269—276 (1965)
2. POPPER, GY.: Singular Value Decomposition of Matrices and Its Application in Numerical Analysis. Periodica Polytechnica, *25*, 201—209 (1981)
3. WOLFE P.: The Secant Method for Simultaneous Nonlinear Equations, Communications of the ACM, *2*, 12 (1959)

Associate Prof. Dr. György POPPER, H-1521 Budapest