ACCELERATING THE CONVERGENCE IN ITERATIVE METHODS

Refaat Ahmed El-Attar
Department of Engineering Mathematics and Physics, Faculty of Engineering,
Alexandria University, Alexandria, Egypt.

ABSTRACT

Accelerating the convergence of iterative methods by means of linear methods of extrapolation is considered. New linear methods are proposed. The use of such methods is shown to improve the rate of convergence of an algorithm that solves nonlinear system of Equations.

1. INTRODUCTION

Iterative methods are widely used to find the solution of systems of linear and nonlinear equations. Optimization techniques to find the extremum of functions of several variables are also iterative in nature. The iterative process usually generates a sequence of vectors that under certain conditions will hopefully converge to the solution of the problem. The purpose of this paper is to study some techniques to accelerate the convergence of such methods. More precisely, we consider some extrapolation methods which are obtained with linear interpolation of the sequence of vectors obtained in the iterative process. In Section 2, we study a general iterative process and obtain the equation for constructing the extrapolation scheme. Most of the iterative methods are linear in nature; however, if the method is nonlinear, we show that near the solution vector, the process can be linearized. In Section 3, linear methods of extrapolation are considered. Section 4 deals with linear projection methods of extrapolation obtained by projecting the operator defining the iterative process onto a subspace of the complete space. In Section 5, we apply linear method of extrapolation to accelerate the convergence of an algorithm that solves a system of nonlinear equations. Finally, in Section 6, we show the advantages of implementing the acceleration scheme. This is illustrated through some numerical examples.

2. GENERAL FORM FOR EXTRAPOLATION

Let $\{x_k\}_{k=1}^{\infty}$ be the sequence of vectors, $x_k \in \mathbb{R}^n$, $k = 1, 2, ...,$ generated by an iterative process. Define the error vectors $y_k = x_k - x^*$, where x^* is the limit point of the sequence. Then, in the general case, the iterative process takes the form.

$$y_{k+1} = F(y_k) \tag{1}$$

where $F(y_k)$ is the operator defining the iterative process. We assume that, near the limit point of the problem, the operator $F(y)$ can be expanded in a power series of operators; linear L, quadratic Q, etc...; then

$$y_{k+1} = L(y_k) + Q(y_k) + \dots \tag{2}$$

For linear approximation, we may write y_{k+1} as

$$y_{k+1} = Ly_k \tag{3}$$

If we define the difference vectors as

$$\delta_k = x_k - x_{k-1} = y_k - y_{k-1} \tag{4}$$

then

$$\delta_{k+p} = y_{k+p} - y_{k+p-1} = L^p y_k - L^p y_{k-1} = L^p \delta_k \tag{5}$$

Then, at the $(k + pq)$ iteration, we can write, using (4)

$$x_{k+pq} = x_k + \delta_{k+1} + \ldots + \delta_{k+pq} \quad (6)$$

and, using (5), we get

$$x_{k+pq} = x_k + [I + L^p + L^{2p} + \ldots + L^{p(q-1)}][\delta_{k+1} + \ldots + \delta_{k+p}]$$

$$= x_k + \left[\sum_{i=1}^{q-1} L^{pi}\right][\delta_{k+1} + \ldots + \delta_{k+p}] \quad (7)$$

where I is the identity matrix.

Let $q \to \infty$ in (7), we obtain the vector

$$x^* = \lim_{q\to\infty} x_{k+pq} = x_k + [\sum_{i=1}^{\infty} L^{pi}][\delta_{k+1} + \ldots + \delta_{k+p}] \quad (8)$$

Recall here that a linear approximation was used, then we say that x^* is an approximation to the solution x_∞ (the limit point of the sequence $\{x_k\}_{k=1}^{\infty}$. In other words, if x_k is given, then a linear approximation to the solution is given by equation (8). Moreover, the series $\sum_{i=0}^{\infty} L^{pi}$ is the expansion of the matrix $[\, i - L^p\,]^{-1}$ in powers of L^p; we may write

$$x^* = x_k + [I - L^p]^{-1}[\delta_{k+1} + \ldots + \delta_{k+p}] \quad (9)$$

Substituting for δ_{k+i} from equation (4), we obtain

$$x^* = x_k + [I - L^p]^{-1}[x_{k+p} - x_k] \quad (10)$$

This equation is the basis for constructing linear extrapolation methods. Corresponding to different ways of approximating L^p or $[I - L^p]^{-1}$, we obtain different extrapolation methods. In the following sections, we give some of these choices.

3. METHODS FOR LINEAR EXTRAPOLATION

Let L be a diagonal matrix G, then equation (10) takes the form

$$x^* = x_k + [I - G^p]^{-1}[x_{k+p} - x_k] \quad (11)$$

and the component wise of the vector x^* is

$$x_j^* = x_{k,j} + (1 - g_{j,j}^p)^{-1}[x_{k+p,j} - x_{k,j}] \quad (12)$$

To eliminate $g_{j,j}^p$ we have at iteration $(k + p + 1)$

$$x_j^* = x_{k+1,j} + (1 - g_{j,j}^p)^{-1}[x_{k+p+1,j} - x_{k+1,j}] \quad (13)$$

From equations (12) and (13), we obtain

$$x_j^* = \frac{x_{k+p+1,j}x_{k,j} - x_{k+p,j}x_{k+1,j}}{x_{k+p+1,j} - x_{k+p,j} - x_{k+1,j} + x_{k,j}} \qquad p = 1, 2, \ldots \quad (14)$$

In the general case, at iteration $(k + p + m)$, we have

$$x_j^* = x_{k+m,j} + (1 - g_{j,j}^p)^{-1}[x_{k+p+m,j} - x_{k+m,j}] \quad (15)$$

and from equations (12) and (15) we get

$$x_j^* = \frac{x_{k+p+m,j}x_{k,j} - x_{k+p,j}x_{k+m,j}}{x_{k+p+m,j} - x_{k+p,j} - x_{k+m,j} + x_{k,j}} \qquad p = 1, 2, \ldots, \quad m = 1, 2, \ldots \quad (16)$$

This relation defines the linear Δ^2- method of extrapolation; it can be transformed to the following form

$$x_j^* = x_{kj} + \left(1 - \frac{x_{k+p+mj} - x_{k+pj}}{x_{k+mj} - x_{kj}}\right)^{-1} [x_{k+pj} - x_{kj}] \quad (17)$$

If we compare relations (17) and (10), we see that the method defined by relation (16) is obtained from (10) by letting the matrix L be of the form

$$(L)^{p}_{i,j}=\begin{cases}\dfrac{x_{k+p+mj}-x_{k+pj}}{x_{k+mj}-x_{kj}}, & i=j\\ 0, & i\neq j\end{cases} \tag{18}$$

If we let $p = m = 1$, the linear extrapolation method defined by (16) reduces to Aitken's familiar three-point δ^2-method of extrapolation,

$$x_j^{*}=\frac{x_{k+2,j}-(x_{k+1,j})^2}{x_{k+2,j}-2x_{k+1,j}+x_{k,j}} \tag{19}$$

and the matrix L is given by

$$L_{i,j}=\begin{cases}\dfrac{x_{k+2+mj}-x_{k+1j}}{x_{k+1j}-x_{kj}}, & i=j\\ 0, & i\neq j\end{cases} \tag{20}$$

For $p = m = 2$, we get

$$x_j^{*}=\frac{x_{k+4,j}x_{k,j}-(x_{k+2,j})^2}{x_{k+4,j}-2x_{k+2,j}+x_{k,j}} \tag{21}$$

This method is of great interest when solving nonlinear equations. Oscillations may arise during the iteration process. Certain components of the vector x_k changes sign at adjacent steps in the process. In this case, it is obvious that relation (21) best suits this particular case.

Now, if we approximate the matrix L by $\alpha_k I$, then relation (10) will be

$$x^{*}=(1-\alpha_k^{p})^{-1}x_{k+p}+\{1-(1-\alpha_k^{p})^{-1}\}x_k \tag{22}$$

If $p = 1$, relation (22) becomes

$$x^{*}=(1-\alpha_k)^{-1}x_{k+1}+\{1-(1-\alpha_k)^{-1}\}x_k \tag{23}$$

Moreover, in the special case, when $\alpha_k = \alpha$, relation (23) reduces to

$$x^{*}=\beta\, x_{k+1}+(1-\beta)\,x_k\ ,\ \beta=(1-\alpha)^{-1} \tag{24}$$

The convergence of the method defined by (24) was studied in [1] and [2].

4. METHODS FOR LINEAR PROJECTION EXTRAPOLATION

We consider here the iterative process defined in Hilbert space H. The linear operator defining the process can be restored from the vectors obtained during the process. The approximate form of the linear operator can then be obtained using a limited number of vectors.

Let the operator $P = [I - L^{p}]^{-1}$ be approximated by $\hat{P}$ in the subspace $\hat{H}$ of the complete space H. Let $\hat{Q}$ be the projector onto subspace $\hat{H}$. Then the operator $\hat{P}$ approximating P has the form

$$\hat{P}=\hat{Q}P\hat{Q}=\hat{Q}[\hat{Q}I\hat{Q}-\hat{Q}L^{p}\hat{Q}]^{-1}\hat{Q} \tag{25}$$

Here we note that $\hat{Q}^2 = \hat{Q}$. Then, the general form for linear projection methods of extrapolation is

$$x^{*}=x_k+\hat{Q}[\hat{Q}I\hat{Q}-\hat{Q}L^{p}\hat{Q}]^{-1}\hat{Q}[x_{k+p}-x_k] \tag{26}$$

Now, Let H be the finite dimensional vector space $\mathbb{R}^n$, and L the Hermitian operator, i.e., L is represented by a symmetric matrix in $\mathbb{R}^n$. Then, it is convenient to choose the space formed by the first m moments of the matrix L. The eigenvalues and corresponding eigenvectors of $\hat{Q}L\hat{Q}$ can be easily found by standard techniques. Knowing the eigenvalues λ_i and the projectors t_i onto the subspace, corresponding to λ_i of $\hat{Q}L\hat{Q}$, we can obtain from (26) the final form of the linear projection extrapolation as

$$x^{*}=x_k+\left\{\sum_{i=1}^{m}\frac{t_i}{1-\lambda_i^{p}}\right\}[x_{k+p}-x_k] \tag{27}$$

5. APPLICATION TO THE SOLUTION OF NONLINEAR SYSTEM OF EQUATIONS

Let $x \in \mathbb{R}^n$ and assume that $f(x)$ is a given real-valued vector function. If $f(x)$ is differentiable at a point x_j, we denote its Jacobian by $J(x_j)$. The problem is to find x^* such that $f(x^*) = 0$. In [3], an algorithm was proposed to find the solution of such a problem. Numerical results showed that the proposed algorithm is stable and superior in terms of computation time than Broyden's method [4].

We introduce here together with this algorithm the linear extrapolation method described in section 3. The steps of the enhanced procedure are given below.

Step 1 Given x_o, H_o, set $i = 0$, compute $f(x_o)$ and

$$\Delta x_o = -H_o f_o \quad \text{and} \quad x_1 = x_o + \Delta x_o$$

Compute and store $(\Delta x_o^T \Delta x_o)$ and x_o.

Step 2 Set $i \rightarrow i + 1$. Compute

$$f_i = f(x_i); \quad v_{oi} = -H_o f_i;$$

$$w_{oi} = (\Delta x_o^T v_{oi}) / (\Delta x_{oi}^T \Delta x_o)$$

If $i \rightarrow 1$, GO TO Step 4.

Step 3 For $j = 1, 2, ..., (i - 1)$

$$v_{ji} = v_{j-1,i} + w_{j-1,i} \Delta x_j$$

$$w_{ji} = (\Delta x_j^T v_{ji}) / (\Delta x_{ji}^T \Delta x_j)$$

Step 4 $\Delta x_i = (v_{i-1,i}) / (1 - w_{i-1,i})$

Compute and store $(\Delta x_i^T \Delta x_i)$

$$x_{i+1} = x_i + \Delta x_i$$

Step 5 If $\|\Delta x_i\| < \epsilon_1$, and $\|f_i\| < \epsilon_2$, STOP.

Step 6 If $i \geq 4$ then

For $j = 1$ to n

$$\hat{x}_{i+1,j} = \frac{x_{i-3,j} x_{i+1,j} - (x_{i-1,j})^2}{x_{i-3,j} - 2x_{i-1,j} + x_{i+1,j}}$$

$$x_{i+1} = \hat{x}_{i+1}.$$

Step 7 GO TO Step 2.

Remarks

1- The first five steps state the algorithm proposed in [3].
2- Step 6 of the procedure is the linear extrapolation method described by equation (21) and uses a four-point extrapolation. Several other choices could have been made, e.g., equation (16) with different values for p and m.
3- The matrix H_i is an approximation to the inverse of the Jacobian matrix J_i of $f(x_i)$.
4- By introducing the acceleration scheme only extra four vectors need to be stored; however the results showed that the enhanced convergence of the sequence to the solution of the problem far outweigh this small disadvantage.

6. NUMERICAL RESULTS

We present here the results of some computational tests in which the Quasi-Newton method described in [3] is compared to the accelerated method proposed in this paper. The test problems in [3] are used to illustrate the effectiveness of the new technique. The following table summarizes the results of the computation.

Table I. Results of the Computation.

	TOTAL CPU TIME (sec)	
	Method in [3]	Accelerated Method
Example E1 [3]	0.75	0.64
Example E2 [3]	1.31	1.02
Example E3 [3]	1.08	0.91
Example E4 [3]	8.32	6.34

7. CONCLUSION

The introduction of an accelerating scheme has greatly improve the convergence in iterative processes. A linear extrapolation technique was used in conjunction with a Quasi-Newton method. Several other choices for linear and linear projection extrapolation were also given. Numerical results proved that the new technique greatly enhances the rate of convergence of iterative process. The new scheme can also be used in function minimization and nonlinear programming.

REFERENCES

[1] A. Hadjidimos and A. Yeyios, "The principle of Extrapolation in Connection with the Accelerated Overrelaxation Method", *Linear Algebra Appl.*, vol. 30, pp. 115-128, 1980.

[2] A. Hadjidimos and A. Yeyios, "How to Improve the Convergence Rates of First Order Scheme", *Internat. J. Computer Math.*, vol. 10, pp. 283-294, 1982.

[3] R.A. El-Attar, "A New Quasi-Newton Algorithm for the Solution of Nonlinear Equations", *Proceedings of the 23rd ISCIE Conference*, pp. 133-138, Osaka, Japan, 1991.

[4] C.G. Broyden, "A Class of Methods for Solving Nonlinear Simultaneous Equations", *Mathematics of Computation*, vol. 19, pp. 577-593, 1965.