

Some Notes on the Sequence Acceleration

Faragó I*

MTA-ELTE NumNet Reseach Group, Institute of Mathematics, Eotvos Lorand University, Hungary

Opinion

In the mathematical problems we seek the solution to the problem.

$$F(u)=0, \quad (1)$$

where F denotes some given operator (function) between two spaces. Typically we are not able to define u^* in an explicit way. Therefore, by use of some numerical method, we replace (1) with the sequence of problems

$$F_n(u_n)=0, n=1,2,\dots, \quad (2)$$

where each problem can be solved, and hence we construct the sequence To our purpose the following requirements are quite natural:

- (u_n) is a convergent sequence;
- $\lim u_n = u^*$.

From a practical point of view, it is also important to have information about the magnitude of the global error $e_n = |u^* - u_n|$, because u_n (for some finite n) is considered as the numerical solution, i.e., the approximation of the exact solution u^* . Clearly, convergence means that e_n tends to zero as n approaches infinity, i.e., e_n is arbitrarily small for sufficiently large values of n . In order to characterize the speed of the convergence, we assume that there exists the $\lim \frac{e_{n+1}}{e_n} = K$. We say that the convergence is linear if $K \in (0,1)$. When $K=1$, then we say that the convergence is sublinear. When $K=0$, the convergence is called superlinear.

For the linear convergence the rate of convergence depends on how far K is from 1: the closer K to 1, the slower the convergence. The linear convergence means that the sequence is similar to the geometric sequence. This means that the convergent sequences, based on Banach's fixed point theorem, has linear convergence.

Therefore the linear and the sublinear convergence might be very slow, and the error quite big for some (even big) value of n . Obviously, from a practical point of view, a numerical method is efficient when the global error is small for relatively small values of n . Hence, in case of slow convergence we need to accelerate the convergence rate.

Aitken's Process

We start with the example, given by the famous Japanese mathematician, Takakazu Seki, which defines an approximation to π in the following way [1]. Let u_n be the perimeter of the polygon with 2^n sides inscribed in a circle of diameter one. Hence

$$u_{15} = 3.1415926487\ 769856708$$

$$u_{16} = 3.1415926523\ 865913571$$

$$u_{17} = 3.1415926532\ 889927759.$$

Seki computed a new value \tilde{u}_{15} as follows:

$$\tilde{u}_{15} := u_{16} + \frac{(u_{16} - u_{15})(u_{17} - u_{16})}{(u_{16} - u_{15}) - (u_{17} - u_{16})} \quad (3)$$

which results in $\tilde{u}_{15} = 3.1415926535897932476$.

Therefore, he concluded that

$$\pi = 3.141592653589. \quad (4)$$

Comparing to the exact decimal digits, we see that all the 12 digits in (4) are exact, while in the expressions for u_{15} , u_{16} and u_{17} only the first 7, 8 and 9 digits are exact. This example shows that with a suitable operation the accuracy can be increased significantly.

The following process was suggested by Alexander Aitken [2].

We assume that (u_n) converges linearly to u^* . Then, for large n we have

$$\lim \frac{u_{n+1} - u^*}{u_n - u^*} \approx K, \text{ and } \lim \frac{u_{n+2} - u^*}{u_{n+1} - u^*} \approx K. \quad (5)$$

Hence,

$$\frac{u_{n+2} - u^*}{u_{n+1} - u^*} \approx \frac{u_{n+1} - u^*}{u_n - u^*}, \quad (6)$$

which yields the relation

$$u^* \approx \frac{u_n u_{n+2} - u_{n+1}^2}{u_n - 2u_{n+1} + u_{n+2}}. \quad (7)$$

Based on (7), we define the new approximation as

$$\tilde{u}_n := \frac{u_n u_{n+2} - u_{n+1}^2}{u_n - 2u_{n+1} + u_{n+2}} = u_n - \frac{(u_{n+1} - u_n)^2}{u_n - 2u_{n+1} + u_{n+2}}. \quad (8)$$

Theorem: Assume that (u_n) converges to u^* linearly. Then

a) The sequence (\tilde{u}_n) converges to the same limit, i.e.,

$$\lim \tilde{u}_n = u^*; \quad (9)$$

b) (\tilde{u}_n) converges faster than (u_n) , i.e.,

$$\lim \frac{\tilde{u}_n - u^*}{u_n - u^*} = 0. \quad (10)$$

Proof: Let

$$e_n = u_n - u^*, \tilde{e}_n = \tilde{u}_n - u^*, n=1,2,\dots \quad (11)$$

Substituting (11) into (8), we get

$$\tilde{e}_n = \frac{e_n e_{n+2} - e_{n+1}^2}{e_n - 2e_{n+1} + e_{n+2}}. \quad (12)$$

In view of linear convergence, we have

*Corresponding author: Faragó I, MTA-ELTE NumNet Reseach Group, Institute of Mathematics, Eotvos Lorand University, Hungary, Tel: +3614116500; E-mail: faragois@cs.elte.hu

Received April 25, 2016; Accepted April 26, 2016; Published April 28, 2016

Citation: Faragó I (2016) Some Notes on the Sequence Acceleration. J Appl Computat Math 5: 303. doi:10.4172/2168-9679.1000303

Copyright: © 2016 Faragó I. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

$$\lim_{n \rightarrow \infty} \frac{e_{n+2}}{e_n} = \lim_{n \rightarrow \infty} \left(\frac{e_{n+2}}{e_{n+1}} \frac{e_{n+1}}{e_n} \right) = K^2. \tag{13}$$

Based on the identity

$$\frac{e_n e_{n+2} - e_{n+1}^2}{e_n - 2e_{n+1} + e_{n+2}} = e_n \frac{e_{n+2} - e_{n+1}}{e_{n+1} - e_n} \frac{e_{n+1}}{1 - 2\frac{e_{n+1}}{e_n} + \frac{e_{n+2}}{e_n}}. \tag{14}$$

which can be verified directly. Since $|K| < 1$, for n tends to infinity (12) implies

$$\lim_{n \rightarrow \infty} \frac{\tilde{e}_n}{e_n} = \lim_{n \rightarrow \infty} \frac{e_{n+1}}{e_n} \frac{e_{n+1} - e_n}{1 - 2\frac{e_{n+1}}{e_n} + \frac{e_{n+2}}{e_n}} = K \frac{K - K}{1 - 2K + K^2} = 0, \tag{15}$$

which proves the statements.

We remark that that obviously the Seki’s method is an Aitken process.

Steffensen’s Method

The above theorem states that the convergence of the new sequence is faster than that of the original one. We can apply this approach to fixed point iterative methods for nonlinear equations. Since for this case the sequence (u_n) is linearly convergent, and the Aitken’s process accelerates the convergence. Namely, the corresponding (\tilde{u}_n) sequence converges to u^* linearly, but faster than (u_n) .

We give the description of another acceleration method, which is called Steffensen’s method [3]. Which is a good combination of the fixed point iteration and Aitken’s method [3].

The algorithm of this method is as follows. We set an arbitrary x_0 , and then $x_1=f(x_0)$, $x_2=f(x_1)$. Having these points, we compute by (20) the new point following Aitken’s method as

$$\tilde{x}_0 := x_0 - \frac{(x_1 - x_0)^2}{x_2 - 2x_1 + x_0},$$

and with this value we restart the fixed point iteration, i.e.,

$$x_3 = \tilde{x}_0, x_4 = f(x_3), x_5 = f(x_4).$$

Then we compute the approximation again by Aitken’s method as

$$\tilde{x}_3 := x_3 - \frac{(x_4 - x_3)^2}{x_5 - 2x_4 + x_3},$$

and so on.

Steffensen’s method (similarly to Newton’s method) provides quadratic convergence. In Newton’s method we compute one function value and one derivative in each iteration step. In Steffensen’s method we have two function evaluations and a more complicated algebraic expression in each iteration, but no derivative. (However, in order to guarantee quadratic convergence for Steffensen’s method, the function f must be three times continuously differentiable, while in Newton’s method only twice.)

The following example demonstrates the methods, investigated so far.

Example: Find the unique fixed point of the function $f(x)=\cos x$ on the interval $[0,1]$. We use three methods [4]:

1. Fixed point iteration: $x_{n+1}=f(x_n)=\cos x_n$.
2. Aitken’s method for the above sequence.

3. Steffensen’s method

(For each we choose $x_0=0.5$ for the starting value) (Table 1).

The numerical experiment shows the following. To compute the fixed point to ten decimal digits of accuracy, the fixed point iteration requires 57 iterates, Aitken’s method 25 iterates, while Steffensen’s method requires only 3 iterates. (This yields that in the fixed point iteration for Aitken’s method 27 basic points, for Steffensen’s method 11 basic points are needed).

Richardson’s Method

Aitken’s method is applicable for the acceleration of linearly convergent sequences. But, even for this case we cannot predict the magnitude of the increase. To give the characterization of the rate of convergence, we say that a sequence (u_n) converges to u^* with order p if there exists a constant $C \geq 0$ such that the inequality $|u_n - u^*| \leq C.n^{-p}$ holds for all, sufficiently large n . In numerical modelling, the sequence (u_n) is generated by solving the numerical model (2). The discrete models are typically constructed as the formulation of the model (1) on a sequence of discrete meshes (grids), and n depends on the number of grid points. Since in the discretization process the number of grid points is inversely proportional to the grid spacing parameter h , we may assume that $h \approx \frac{1}{n}$.

Hence, the Landau notation can be written as $O(h^p)$, which means that

$$e_n = |u_n - u^*| = O(h^p). \tag{16}$$

This can be written as

$$u^* = u_n + C_p h^p + O(h^{p+1}). \tag{17}$$

The Richardson’s method based on the following observation [5]. When we have 2 different sequences, obtained by the same numerical method but on different meshes, formula (19) gives possibility to eliminate the term $C_p h^p$, and we obtain a method of higher order. E.g., when the sequences of p -th order are generated on the meshes of the step-size $2h$ and h , respectively, based on the corresponding two approximations we define the new approximation as

$$w_h = \frac{2^p u_h - u_{2h}}{2^p - 1}, \tag{18}$$

which will have the order of convergence $p+1$.

Conclusion

While the Aitken’s and Steffensen’s methods serve to accelerate the speed of the convergence of some fixed sequence by use of some transformation of the elements of this sequence, the Richardson’s method does the acceleration by use of two sequences, generated on by the same method but on different meshes.

	1	2	3
0	0.50000	0.73139	0.73139
1	0.87758	0.73609	0.73908
2	0.63901	0.73765	
3	0.80269	0.73847	
4	0.69478	0.73880	
5	0.76820		
6	0.71917		

Table 1: Decimal digits of accuracy.

References

1. Hirayama A, Shimodaira K, Hirose H (1974) Takakazu Seki's collected works. Osaka Kyoiku Toshosha pp: 57-58.
2. Aitken A (1926) On Bernoulli's numerical solution of algebraic equations. Proceedings of the Royal Society of Edinburgh 46: 289-305.
3. Dahlquist G, Björck A, Anderson N (1974) Numerical Methods. Englewood Cliffs, Prentice Hall, NJ pp: 230-231.
4. Jim Lambers (2010) Accelerating Convergence.
5. Richardson LF (1927) The Deferred Approach to the Limit, I—Single Lattice, Philosophical Transactions of the Royal Society of London 226: 299-349.