Applied Numerical Mathematics 90 (2015) 38–54

Contents lists available at ScienceDirect

Applied Numerical Mathematics

www.elsevier.com/locate/apnum

Generalizations of Aitken's process for a certain class of sequences

Davide Buoso, Anna Karapiperi, Stefano Pozza *

Dipartimento di Matematica, Università degli Studi di Padova, Via Trieste 63, 35121, Padova, Italy

ARTICLE INFO

Article history:
Received 18 March 2013
Received in revised form 26 July 2014
Accepted 4 December 2014
Available online 9 December 2014

Keywords:
Convergence acceleration
Aitken's process
Extrapolation
Semi-convergence

ABSTRACT

In this paper, we construct several sequence transformations whose kernels contain sequences of the form $S_n = S + a_n\lambda^n$, $n = 0, 1, \ldots$, where S and λ are unknown parameters, and (a_n) is a known sequence. These transformations generalize Aitken's Δ^2 process. We provide certain sufficient conditions under which one of our transformations accelerates the convergence of certain types of sequences. Finally, we illustrate these theoretical results through several numerical experiments using diverging and converging sequences.

1. Introduction

In numerical analysis and in theoretical physics, sequences which converge slowly to their limit occur frequently. These require too many terms for a good approximation. For this reason sequence transformations play a fundamental role as they potentially accelerate the convergence of a sequence or power series. In other words, a smaller number of terms needs to be computed and this enables to estimate the limit more efficiently. Moreover, such sequence transformations are used for summing up divergent sequences and series.

It has been proved that a transformation able to accelerate any sequence does not exist (see [11,12]). Thus, it is important in practical situations to develop new algorithms depending on the class of sequences of interest. Seemingly, if this class is too small, such transformation will be very specific and hence it will be useful only in particular cases; on the other hand, such specialization typically provides better acceleration properties. For this reason, the notion of *kernel* is of particular interest. The kernel of a transformation $T : (S_n) \longmapsto (T^{(n)})$ is the set of all sequences (S_n) which are transformed by T into a constant sequence, i.e. $T^{(n)} = S$ for all n, usually the limit of the sequence S_n (if it exists). When T is applied to a sequence which does not belong to its kernel, it produces a sequence $(T^{(n)})$ which could converge, under certain assumptions, to S faster than (S_n), that is

$$\lim_{n\to\infty} \frac{T^{(n)} - S}{S_n - S} = 0.$$

In this case, we say that the transformation accelerates the convergence of the sequence. Although it has not yet been proven, numerical experiments have always confirmed that the "closer" a sequence is to the kernel, the faster the transformed sequence will converge (to the same limit).

* Corresponding author. Tel.: +39 049 827 1304.
E-mail addresses: dbuoso@math.unipd.it (D. Buoso), akarapi@math.unipd.it (A. Karapiperi), spozza@math.unipd.it (S. Pozza).

http://dx.doi.org/10.1016/j.apnum.2014.12.002
0168-9274/

Sequence transformations have nowadays become an extremely useful numerical tool, about which a vast literature has been developed (see e.g. [8,5,7,11,19,23,27]). Moreover, in the last decade many works have appeared on how to effectively use sequence transformations (or other numerical tools) in practical situations. See for instance [4,16,18], and [21]. See also [14], where the efficiency of several numerical techniques is discussed in order to facilitate the evaluation of power series expansions for special functions.

The standard way of constructing a transformation is to start from the definition of the kernel which is usually characterized by a property. For instance, the kernel of Aitken's Δ^2 process[1] is the set of sequences of the form

$$S_n = S + a\lambda^n, \quad n = 0, 1, \ldots, \tag{1}$$

where $a \in \mathbb{C}$ is different from 0 and $\lambda \in \mathbb{C}$ is different from 0 and 1. If $|\lambda| < 1$, then S is the limit of the sequence; otherwise, for (S_n) diverging, S is called the *antilimit* of the sequence.

In the sequel we present some formal aspects of sequence transformations as discussed in [22, Section 2]. It is known that a computational algorithm can only involve a finite number of arithmetic operations, thus a sequence transformation T can only use *finite* subsets of the original sequence $(S_n)_{n=0}^{\infty}$ for the computation of new sequence elements. These finite subsets normally consist of consecutive elements, i.e. $\{S_n, S_{n+1}, \ldots, S_{n+\ell}\}$.

All the commonly used sequence transformations T can be represented by infinite sets of double indexed quantities $T_k^{(n)}$ with $k, n \in \mathbb{N}_0$, which can be displayed in a two-dimensional array called the *table* of T. Following the commonly used notation, we use the convention that the superscript n indicates the minimal index occurring in the finite subset of sequence elements used for the computation of a given $T_k^{(n)}$, whilst the subscript k – usually called the *order* of the transformation – is a measure for the complexity of the transformation process.

The elements $T_k^{(n)}$ of the table of T are gauged in such a way that $T_0^{(n)}$ corresponds to an untransformed sequence element, i.e. $T_0^{(n)} = S_n$, $n \in \mathbb{N}_0$. An increasing value of k implies that the complexity of the transformation increases. In that case, $\ell = \ell(k)$ increases too. This means that for every $k, n \geq 0$ the sequence transformation T produces a new transform according to $T_k^{(n)} = T(S_n, \ldots, S_{n+\ell(k)})$. The exact relationship connecting k and ℓ is specific for a given sequence transformation.

In order to obtain a good approximation to S we can proceed following several different *paths* in the table of T. Here we mention the two predominantly used in practical applications paths, namely the *order-constant* or *horizontal* path and the *index-constant* or *vertical* path.[2] On an *order-constant* path, the transformation order k is kept fixed and $n \to \infty$, while if we keep fixed the index n and $k \to \infty$, then we have an *index-constant* path. In other words, on an index-constant path, $T_k^{(n)}$ is computed with the highest possible transformation order k for a given set of input data. This is the reason why the index-constant approach is normally more efficient.

However, in this paper we follow the order-constant path. Aitken's Δ^2 process, which belongs to the order-constant case, is the following transformation

$$T^{(n)} = \frac{S_n S_{n+2} - S_{n+1}^2}{S_{n+2} - 2S_{n+1} + S_n}, \quad n = 0, 1, \ldots.$$

Since the previous formula is unstable (see for example [8, pp. 34–35, pp. 400–403], [9, p. 173]), in numerical applications other equivalent formulas are used such as

$$T^{(n)} = S_n - \frac{(\Delta S_n)^2}{\Delta^2 S_n} = S_{n+1} - \frac{\Delta S_n \Delta S_{n+1}}{\Delta^2 S_n} = S_{n+2} - \frac{(\Delta S_{n+1})^2}{\Delta^2 S_n}, \quad n = 0, 1, \ldots, \tag{2}$$

where Δ is the forward difference operator defined as $\Delta S_n = S_{n+1} - S_n$. Note that the denominator is $\Delta^2 S_n = \Delta S_{n+1} - \Delta S_n$, which explains the name.

In this special case it has been proved that the kernel consists of all the sequences of the form of Eq. (1) and only them. In general, sufficiency is difficult to prove. We refer to [8,10] for a detailed discussion of this transformation.

In [9], Brezinski and Redivo-Zaglia considered a more general kernel, namely one consisting of sequences of the form

$$S_n = S + (a + bx_n)\lambda^n, \quad n = 0, 1, \ldots, \tag{3}$$

or

$$S_n = S + (a + bx_n)^{-1}\lambda^n, \quad n = 0, 1, \ldots, \tag{4}$$

where S, a, b and λ are unknown numbers and (x_n) a known sequence. These kernels obviously contain (1).

[1] Alexander Craig Aitken (1895–1967) used this method in [3] (1926), so it is named after Aitken. However, Aitken's Δ^2 process was actually discovered by Japanese Mathematician Takakazu Seki (?–1708) before 1680. The same method was obtained by Hans Eduard von Naegelsbach (1838–1899?) in 1876 and by James Clerk Maxwell (1831–1879) in 1873 but neither of them used it for the purpose of acceleration (see e.g. [6,17] and the references therein).

[2] For a formal definition of the *path* refer to [27, Definition 3, p. 3]. Therein the terms *horizontal* and *vertical* path are used. The terminology *order-constant* and *index-constant* path respectively was introduced in [22, Section 2].