CHAPTER 3

Padé Approximants and Numerical Methods

3.1 Aitken's Δ^2 Method as $[L/1]$ Padé Approximants

One of the best-known and simplest techniques of accelerating the convergence of a sequence is Aitken's Δ^2 method. Given a sequence of real or complex numbers,

$$\mathcal{S}=\{S_n,\ n=0,1,2,\ldots\}, \tag{1.1}$$

such that $S_n \to S$ as $n\to\infty$, the problem is to find a new sequence which converges faster to S.

Define

$$\Delta S_n = S_{n+1} - S_n,$$
$$\Delta^2 S_n = \Delta(\Delta S_n) = S_{n+2} - 2S_{n+1} + S_n,$$

which are the usual forward differences, and the new sequence

$$\mathcal{T}=\{T_n,\ n=0,1,2,\ldots\},$$

where

$$T_n = S_n - \frac{(\Delta S_n)^2}{\Delta^2 S_n}. \tag{1.2}$$

It is clear from (1.2) why Aitken's method is called the Δ^2 method. There are many reasons for expecting in general that $\mathcal{T}$ converges to a limit, that this

ENCYCLOPEDIA OF MATHEMATICS and Its Applications, Gian-Carlo Rota (ed.). Vol. 13: George A. Baker, Jr., and Peter R. Graves-Morris, Padé Approximants: Basic Theory, Part I ISBN 0-201-13512-4

limit is S, and that convergence has been accelerated. But we must add an early word of caution: Aitken's method does not work for any arbitrary convergent sequence $\mathbb{S}$. Like all algorithms of numerical analysis, Aitken's method has its own domain of validity, and in certain circumstances it should not be used. An important example is where all the S_n are identical, so that the T_n are undefined. A more insidious example is the one in which the S_n are equal up to rounding errors, so that the T_n are meaningless noise. This is a notorious situation to beware of. But, in general, the method is safe if it is empirically convergent, and it has wide applicability.

Basically, Aitken's method [1926] is designed to treat sequences with geometric convergence. Suppose that

$$S_n = S - a\alpha^n \tag{1.3}$$

with $a \neq 0$ and $|\alpha| < 1$. Then

$$\Delta S_n = a\alpha^n(1-\alpha),$$
$$\Delta^2 S_n = -a\alpha^n(1-\alpha)^2,$$

and from (2),

$$T_n = (S - a\alpha^n) + (a\alpha^n) = S. \tag{1.4}$$

We see that in this case, Aitken's method yields the exact answer at every stage. More generally, for a sequence $\mathbb{S}$ which is dominated by one geometrically convergent component, we expect that Aitken's method accelerates convergence by "taking out the geometrically convergent part".

As a practical example (see Part II, Section 3.1) we consider the numerical evaluation of

$$S = \int_0^1 x^{-1/2}(1-x)^{-1/2} \exp x \, dx. \tag{1.5}$$

The integrand of (5) is infinite at the end points, but the integral is well defined. We define S_n as the value of the integral obtained by using 2^n equally spaced integration points. These Riemann sums, obtained by doubling the number of integration points at each successive evaluation, converge to S and are obtained with great ease. It turns out that Aitken's algorithm is a very effective technique of estimating S.

The connection between Padé approximants and Aitken's Δ^2 method is made, as in Section 1.3, by using the series derived from the sequence. Define

$$\begin{aligned} c_{n+1} &= \Delta S_n = S_{n+1} - S_n, \qquad n = 0, 1, 2, \ldots, \\ c_0 &= S_0. \end{aligned} \tag{1.6}$$

It follows that S_n are the partial sums of the series, and of course the series converges to S. We form the power series

$$f(z)=\sum_{i=0}^{\infty} c_i z^i. \tag{1.7}$$

Remember that formal power series have a radius of convergence which may be zero, finite, or infinite. We wish to evaluate $f(1)=S$. The method of finding $f(1)$ using the second row of the Padé table is to evaluate

$$[L/1]_f(1), \qquad L=0,1,2,\ldots, \tag{1.8}$$

and to determine the limit as $L\to\infty$.

From (1.1.8)–(1.1.10),

$$\begin{aligned}
[L/1]_f(1)&=\begin{vmatrix} c_L & c_{L+1} \\ \sum_{i=0}^{L-1} c_i & \sum_{i=0}^{L} c_i \end{vmatrix} \div \begin{vmatrix} c_L & c_{L+1} \\ 1 & 1 \end{vmatrix} \\
&=\frac{(S_L-S_{L-1})S_L-(S_{L+1}-S_L)S_{L-1}}{(S_L-S_{L-1})-(S_{L+1}-S_L)} \\
&=\frac{S_{L-1}(S_{L+1}-2S_L+S_{L-1})-(S_L-S_{L-1})^2}{S_{L+1}-2S_L+S_{L-1}} \\
&=S_{L-1}-\frac{(\Delta S_{L-1})^2}{\Delta^2 S_{L-1}},
\end{aligned}$$

which agrees with (1.2) and shows Aitken's method to be the equivalent of using $[L/1]$ Padé approximants. An even more rapid proof of this result is given by taking $M=1$ in (1.3.8).

A few sequences of numerical analysis are of the special type

$$S_{n+1}=f(S_n) \tag{1.9}$$

The function f is called a one-point iteration function in this context [Traub, 1964]. For example, the geometric sequence

$$S_n=S-a\alpha^n$$

in (1.3) is generated by

$$\begin{aligned}
S_{n+1}&=S-a\alpha^{n+1} \\
&=S+\alpha(S_n-S) \\
&=f(S_n)
\end{aligned}$$

with the identification $f(z)=S+\alpha(z-S)$. We see that in this case we have a convergent sequence when $\alpha=f'(S)$ satisfies $|\alpha|<1$. Further, the geometric sequence corresponds to $f(z)$ being linear. If $\alpha=1$, then $S_n=S-a$, and Aitken's method is inapplicable in this situation. We get further confirmation of the power of Aitken's method and so also of the $[L/1]$ Padé method from the following theorem.

THEOREM 3.1.1 [Henrici, 1964]. *Let $S_{n+1}=f(S_n)$ define a convergent real sequence with limit S, let $f(x)$ be twice differentiable at S, and let $f'(S)\neq 1$. Then, with the definition* (1.2),

$$T_n-S=O\big((S_n-S)^2\big).$$

Proof.

$$\begin{aligned} S_{n+1}-S_n &= f(S_n)-S_n \\ &= f(S)+(S_n-S)f'(S)+(S-S_n)^2 f''(\xi_n)-S_n \end{aligned} \tag{1.10}$$

for some ξ_n lying between S and S_n. Continuity of $f(x)$ and convergence of the sequence imply that $f(S)=S$. Hence (1.10) may be written as

$$\Delta S_n = A(S_n-S)+O\big((S-S_n)^2\big), \tag{1.11}$$

where

$$A=f'(S)-1\neq 0.$$

From (1.10),

$$S_{n+1}-S=(S_n-S)f'(S)+(S-S_n)^2 f''(\xi_n).$$

Similarly, from (1.10) and (1.11),

$$\begin{aligned} S_{n+2}-S_{n+1} &= (S_{n+1}-S)(f'(S)-1)+(S-S_{n+1})^2 f''(\xi_{n+1}) \\ &= A(S_{n+1}-S)+O\big((S-S_n)^2\big). \end{aligned} \tag{1.12}$$

Therefore from (1.11) and (1.12)

$$\Delta^2 S_n = A\,\Delta S_n + O\big((S-S_n)^2\big).$$

Hence

$$\begin{aligned} T_n &= S_n - \frac{(\Delta S_n)\{A(S_n - S) + O((S - S_n)^2)\}}{A\,\Delta S_n(1 + O(S - S_n)^2)} \\ &= S_n - (S_n - S) + O((S - S_n)^2) \\ &= S + O((S - S_n)^2), \end{aligned}$$

so proving the theorem.

The previous theorem makes quite precise the statement that Aitken's method and the $[L/1]$ Padé method accelerate convergence of a sequence dominated by a geometrically convergent component, of the type given by (1.9). In the next section we turn our attention to generalizing these basic ideas. For further details of the general theory, we refer to Brezinski [1977].

If both sequences $\mathfrak{S}$ and $\mathfrak{T}$ given by (1.1) and (1.2) converge, then they converge to the same limit [Lubkin, 1952]. For an account of recent progress with convergence theory, we refer to Cordellier [1979a] and Germain-Bonne [1979].

Exercise 1. Show that Newton's method of finding a zero of a function $\phi(z)$ takes the form (1.9), with the one-point iteration function

$$f(z) = z - \frac{\phi(z)}{\phi'(z)}.$$

Exercise 2. Is it a good idea to accelerate convergence of the sequence generated by Newton's method of the previous exercise if

(i) z_0 is a simple root of $\phi(z)$?

(ii) z_0 is a multiple root of $\phi(z)$?

Exercise 3. Consider the series

$$\sum_{i=0}^{\infty} c_i = \frac{1}{2} + \frac{1}{3} - \frac{5}{6} + \frac{1}{4} + \frac{1}{5} - \frac{9}{20} + \cdots$$

where

$$c_{3m-3} = \frac{1}{2m}, \qquad c_{3m-2} = \frac{1}{2m+1}, \quad \text{and} \quad c_{3m-1} = -\frac{4m+1}{4m^2+2m}$$

for $m=1,2,3,\ldots$ [Marx, 1963]. Define $S_n=\sum_{i=0}^{n} c_i$, and T_n by (1.2). Prove that

(i) $S_{3m-1}=0$ for $m=1,2,3,\ldots,$
(ii) $S_n \to 0$ as $n\to\infty$,
(iii) $T_{3m}\to 0$, $T_{3m-1}\to 1$, and $T_{3m-2}\to 0$ as $m\to\infty$.

Notice the implication that $\{S_n\}$ converges, yet $\{T_n\}$ diverges by oscillation.

3.2 Acceleration and Overacceleration of Convergence

It is natural to ask how the accelerated sequence derived from Aitken's Δ^2 method may be improved upon. The natural answer is to iterate Aitken's method. This answer is not entirely satisfactory, because of the lack of justification based on principle, as the following remarks will make clear.

Aitken's scheme works well if the original sequence converges geometrically; the accelerated sequence takes full account of the dominant terms in the original sequence, and one should wonder what is the reason for accelerating again. Let us suppose that the original sequence is a geometric sequence rounded to given accuracy. Then the accelerated sequence, according to (1.4), is the limit but contains the small rounding errors. Further acceleration by Aitken's method (and the Padé method for that matter) requires differencing, and consequently, the results depend entirely on rounding errors in the original sequence. Thus, some sort of theoretical basis or an empirical numerical criterion is an essential prerequisite before iterating acceleration schemes. It is all too tempting to try to extract too much information by accelerating a few terms of a sequence too fast.

Consider the partial sums

$$S_n=\sum_{r=0}^{n}(1+\varepsilon_r)(0.5)^r,\qquad n=1,2,3,\ldots,$$

where the numbers $\{\varepsilon_r\}$ represent floating-point rounding errors with $|\varepsilon_i|\leqslant\varepsilon$. We consider estimation of the quantity S_∞ using the Padé-approximant method by forming diagonal approximants to

$$f(z)=\sum_{r=0}^{\infty}(1+\varepsilon_r)z^r$$

and evaluating the approximants at $z=0.5$. (Note the trivial variation on the method described in Sections 3.1, and 1.1, where the approximants are evaluated at $z=1$.) Working to first order in ε in (1.1.8) and (1.1.9), we find the error bounds

$$[1/1]=2\pm 7\varepsilon,$$

showing that this approximant is sensitive, but not unduly sensitive, to rounding errors in the data coefficients. However, when we come to consider the [2/2] Padé approximant, we find that

$$Q^{[2/2]}(z)=z^2(\varepsilon_2+\varepsilon_4-2\varepsilon_3)-z(\varepsilon_1+\varepsilon_4-\varepsilon_2-\varepsilon_3)+(\varepsilon_1+\varepsilon_3-2\varepsilon_2)$$

to first order. We see that the value of $Q^{[2/2]}(z)$ is completely controlled by rounding error in this case; the zeros of $Q^{[2/2]}(z)$, which are the poles of the approximant, are distributed all over the complex plane. (We do not suggest that the distribution is random, and we would expect more zeros of $Q^{[2/2]}(z)$ near $|z|=1$ than near $z=0$, for example.) Since the value of the [2/2] Padé approximant depends primarily on rounding error, whereas the [1/1] approximant is accurate within errors, use of the [2/2] approximant is an example of overacceleration of convergence. In this case, the moral is that we should use the lower-order approximant.

The Padé method has the following interpretation (among others): the given sequence has a certain number of geometric components which dominate. Let us suppose that

$$S_n=a\sum_{r=0}^{n}\alpha^r+b\sum_{r=0}^{n}\beta^r+\text{much smaller terms} \tag{2.1}$$

with $a\neq0$, $b\neq0$, $|\alpha|<1$, $|\beta|<1$, and $n=0,1,2,\ldots$. This expression may be rewritten as

$$S_n=\frac{a}{1-\alpha}(1-\alpha^{n+1})+\frac{b}{1-\beta}(1-\beta^{n+1})+\text{much smaller terms.} \tag{2.2}$$

Note that S_n is derived from the series

$$\begin{aligned} c_0&\equiv S_0=a+b+\text{a much smaller term},\\ c_n&\equiv S_n-S_{n-1}=a\alpha^n+b\beta^n+\text{much smaller terms}, \qquad n=1,2,3,\ldots \end{aligned} \tag{2.3}$$

The third row of the Padé table takes account of the explicit leading terms in (2.1), (2.2), and (2.3), whereas direct calculation shows that the once iterated Aitken method does not. We assumed in (2.1) that $a\neq0$ and $b\neq0$, so that there are genuinely two geometric components which dominate the remainder, and this assumption is crucial. We are a bit vague about the size of the remainder terms, so as not to prejudice the development, and to admit possibilities such as $c_n=\alpha^n+(-\alpha)^n$ for which the odd terms vanish. We assume that $|\alpha|<1$ and $|\beta|<1$ which is conventional but not entirely necessary. The Padé method does make sense of well-posed problems with divergent sequences, such as (2.1) with $|\alpha|>1$ or $|\beta|>1$. If $\alpha=1$ or $\beta=1$, corresponding to a divergent sequence with one component derived from an arithmetic progression, the Padé method gives $S=\infty$ with an obvious interpretation.

As in Section 1.3, to justify the Padé method, we form the function

$$f(z)\equiv\sum_{r=0}^{\infty}c_r z^r=\frac{a}{1-\alpha z}+\frac{b}{1-\beta z}+\text{correction terms}, \tag{2.4}$$

which we wish to evaluate at $z=1$. Formation of $[L/2]$ approximants is suggested by the explicit dominant terms of (2.4), and we apply de Montessus's theorem. Borrowing from Section 6.2, we quote the theorem in context here.

Let $R>|\alpha|^{-1}$ and $R>|\beta|^{-1}$. We assume that the remainder terms in (2.3) are small, and explicitly we require that $c_n=o(R^{-n})$. Hence, $f(z)$ is meromorphic with precisely two poles within $|z|<R$. Since $|\alpha|>1$, then $R>1$. Now de Montessus's theorem asserts that $[L/2]$ approximants converge to $f(z)$ at $z=1$, which is not a pole of $f(z)$, by assumption. Hence, the $[L/2]$ approximants converge for sequences such as (2.1) or series such as (2.3), with the stated hypothesis about the residuals.

3.3 The ε-Algorithm and the η-Algorithm

In this section we describe the ε-algorithm for sequence transformations and show that one of its columns is the sequence of Aitken's Δ^2 method. Then we describe the η-algorithm, which is the corresponding algorithm for series transformations.

The ε-algorithm originates with Shanks [1955] and Wynn [1956]. It involves the two-dimensional array called the ε-table (Table 1). The subscript k of $\varepsilon_k^{(j)}$ denotes the column, and the superscript j measures the progression down the column. The table is constructed iteratively from its first two elements. Define $\varepsilon_{-1}^{(j)}$ to be zero and $\varepsilon_j^{(0)}$ to be the given sequence, for $j=0,1,2,\ldots$. Then all the other elements may be calculated from the

Table 1. The ε-Table.

$\varepsilon_{-1}^{(0)}$			
	$\varepsilon_0^{(0)}$		
$\varepsilon_{-1}^{(1)}$		$\varepsilon_1^{(0)}$	
	$\varepsilon_0^{(1)}$		$\varepsilon_2^{(0)}$
$\varepsilon_{-1}^{(2)}$		$\varepsilon_1^{(1)}$	$\vdots$
	$\varepsilon_0^{(2)}$	$\vdots$	
$\varepsilon_{-1}^{(3)}$	$\vdots$		
$\vdots$			

ε-algorithm, which is

$$\varepsilon_{k+1}^{(j)} = \varepsilon_{k-1}^{(j+1)} + \left[\varepsilon_k^{(j+1)} - \varepsilon_k^{(j)}\right]^{-1}. \tag{3.1}$$

To see more clearly how this rule should be applied, we note that it connects the elements in the rhombus pattern of Figure 1, which shows how the

$$\begin{array}{ccc} & \varepsilon_k^{(j)} & \\ \varepsilon_{k-1}^{(j+1)} & & \varepsilon_{k+1}^{(j)} \\ & \varepsilon_k^{(j+1)} & \end{array}$$

Figure 1. A rhombus pattern.

right-hand member $\varepsilon_{k+1}^{(j)}$ is derived from the other three members. It is now plain that the ε-algorithm allows the whole ε-table to be calculated. It is further plain that if $\varepsilon_k^{(j)} = \varepsilon_k^{(j+1)}$, i.e. two successive members of the same column are equal, the element $\varepsilon_{k+1}^{(j)}$ does not exist. We assume, unless explicitly stated otherwise, that all elements exist. Otherwise the table is said to be degenerate. We will show that the sequence of the fourth column, namely $\{\varepsilon_2^{(j)}, j=0,1,2,\ldots\}$, is the same as that obtained from Aitken's Δ^2 rule.

From (3.1),

$$\begin{aligned} \varepsilon_1^{(j)} &= \left[\varepsilon_0^{(j+1)} - \varepsilon_0^{(j)}\right]^{-1} \\ &= \left[S_{j+1} - S_j\right]^{-1} = \left[\Delta S_j\right]^{-1}. \end{aligned}$$

Again from (3.1),

$$\begin{aligned} \varepsilon_2^{(j)} &= \varepsilon_0^{(j+1)} + \left[\varepsilon_1^{(j+1)} - \varepsilon_1^{(j)}\right]^{-1} \\ &= S_{j+1} + \frac{1}{\left[\Delta S_{j+1}\right]^{-1} - \left[\Delta S_j\right]^{-1}} \\ &= S_j + \Delta S_j + \frac{\Delta S_j \Delta S_{j+1}}{\Delta S_j - \Delta S_{j+1}} \\ &= S_j - \frac{(\Delta S_j)^2}{\Delta^2 S_j}, \qquad \text{where} \quad \Delta^2 S_j = \Delta S_{j+1} - \Delta S_j. \end{aligned}$$

This formula is precisely Aitken's Δ^2 method (1.2) applied to the sequence $\{S_j, j=0,1,2,\ldots\}$, and is also the result of using the second row of the Padé table as a Padé method for sequence acceleration.

After we have established Wynn's identity in Section 3.4, we then show in Section 3.5 that the ε-table and the Padé table are identified by the formula

$$\varepsilon_{2k}^{(j)}=[k+j/k]_f(1).$$

What we have just achieved is the proof of this result for $k=1$ and $j=0,1,2,\ldots$.

Table 2. Even Columns of an ε-Table for π.

$n=0$	4.0000000				
1	2.6666667	3.1666667			
2	3.4666667	3.1333333	3.1423423		
3	2.8952381	3.1452381	3.1413919	3.1416149	
4	3.3396825	3.1396825	3.1416627	3.1415873	3.1415933
5	2.9760462	3.1427129	3.1415634	3.1415943	3.1415925
6	3.2837385	3.1408813	3.1416065	3.1415921	
7	3.0170718	3.1420718	3.1415854		
8	3.2523659	3.1412548			
9	3.0418396				

Example 1. We consider Gregory's notoriously slowly convergent series for π,

$$\pi=4-\tfrac{4}{3}+\tfrac{4}{5}-\tfrac{4}{7}+\cdots.$$

In Table 2, we exhibit the even columns of the ε-table for this series. The first column seems scarcely convergent, whereas the correctness of the final extrapolations is instantly recognizable.

Table 3. Even Columns of an ε-Table for ln 3.

$n=0$	.000000				
1	2.000000	1.000000			
2	.000000	1.142857	1.090909		
3	2.666667	1.066667	1.101449	1.098039	
4	−1.333333	1.128205	1.097046	1.098805	1.098570
5	5.066667	1.066667	1.099725	1.098521	1.098626
6	−5.600000	1.136842	1.097674	1.098667	
7	12.685714	1.049351	1.099507		
8	−19.314286	1.165714			
9	37.574603				

Example 2. We consider a familiar divergent series, namely the one given by the Maclaurin series of $\ln(1+z)$ with $z=2$:

$$\ln 3=2-2+\frac{8}{3}-4+\cdots.$$