Numer. Math. 40, 39 – 46 (1982)

Numerische Mathematik

The ε-Algorithm and Multivariate Padé-Approximants

Annie A.M. Cuyt[1]

University of Antwerp, Department of Mathematics, Universiteitsplein 1, B-2610 Wilrijk, Belgium

Summary. In the univariate case the ε-algorithm of Wynn is closely related to the Padé-table in the following sense: if we apply the ε-algorithm to the partial sums of the power series $f(x) = \sum_{i=0}^{\infty} c_i x^i$ then ε_{2m}^{l-m} is the (l, m) Padé-approximant to $f(x)$ where l is the degree of the numerator and m is the degree of the denominator [1 pp. 66–68].

Several generalizations of the ε-algorithm exist but without any connection with a theory of Padé-approximants.

Also several definitions of the Padé-approximant to a multivariate function exist, but up till now without any connection with the ε-algorithm.

In this paper, we see that the multivariate Padé-approximants introduced in [3], satisfy the same property as the univariate Padé-approximants: if we apply the ε-algorithm to the partial sums of the power series

$$f(x_1, \ldots, x_n) = \sum_{i_1+\ldots+i_n=0}^{\infty} c_{i_1\ldots i_n} x_1^{i_1} \ldots x_n^{i_n}$$

then $\varepsilon_{2m}^{(l-m)}$ is the (l, m) multivariate Padé-approximant to $f(x_1, \ldots, x_n)$.

Subject classification: AMS(MOS): 65D15, CR: 5.13

1. Univariate Padé-Approximants

We briefly repeat the concepts which will be generalized to the multivariate case in the third section.

Let $f(x) = \sum_{k=0}^{\infty} c_k x^k$, $p(x) = \sum_{i=0}^{l} a_i x^i$ and $q(x) = \sum_{j=0}^{m} b_j x^j$. We say that $f(x) = O(x^k)$ if non-negative constants K and r with $r<1$ exist such that $|f(x)| \leqq K|x|^k$ for $|x|<r$.

[1] Aspirant of the Belgian National Foundation for Scientific Research (NFWO)

0029-599X/82/0040/0039/$01.60

Definition 1.1. If $p(x)$ and $q(x)$ satisfy

$$f(x)\cdot q(x)-p(x)=O(x^{l+m+1})$$

then the irreducible form $R_{lm}(x)=\frac{p_*(x)}{q_*(x)}$ of $\frac{p(x)}{q(x)}$ is called the (l, m) Padé-approximant to f.

A normalization such as $q_*(0)=1$ is generally imposed in order to guarantee the unicity of the Padé-approximant.

The following property for $p(x)$ and $q(x)$ satisfying Definition 1.1. can easily be proved [4, pp. 243].

Theorem 1.1. *If the following determinant*

$$\begin{vmatrix} 1 & \cdots & 1 \\ c_{l+1}x^{l+1} & \cdots & c_{l+1-m}x^{l+1-m} \\ \vdots & & \vdots \\ c_{l+m}x^{l+m} & \cdots & c_l x^l \end{vmatrix} \neq 0$$

then

$$\frac{p(x)}{q(x)}=\frac{\begin{vmatrix} f_l(x) & f_{l-1}(x) & \cdots & f_{l-m}(x) \\ c_{l+1}x^{l+1} & c_l x^l & \cdots & c_{l+1-m}x^{l+1-m} \\ \vdots & & & \vdots \\ c_{l+m}x^{l+m} & & \cdots & c_l x^l \end{vmatrix}}{\begin{vmatrix} 1 & & & 1 \\ c_{l+1}x^{l+1} & c_l x^l & \cdots & c_{l+1-m}x^{l+1-m} \\ \vdots & & \ddots & \vdots \\ c_{l+m}x^{l+m} & & & c_l x^l \end{vmatrix}}$$

where $f_i(x)=\sum_{k=0}^{i} c_k x^k$.

In numerator and denominator of the quotient of determinants in Theorem 1.1. a factor x^{lm} can immediately be cancelled, but we prefer this expression which can easily be generalized to the multivariate case.

2. The ε-Algorithm

The ε-algorithm perfoms the following caluclations; the input are the elements of a row $\{s_i \mid i=0, 1 \ldots\}$.

Put

$$\begin{aligned} &\varepsilon_{-1}^{(i)}=0 && i=0,1,\ldots, \\ &\varepsilon_0^{(i)}=s_i && i=0,1,\ldots, \\ &\varepsilon_{2j}^{(-j-1)}=0 && j=0,1,\ldots, \end{aligned}$$

and compute

$$\varepsilon_{j+1}^{(i)}=\varepsilon_{j-1}^{(i+1)}+\frac{1}{\varepsilon_j^{(i+1)}-\varepsilon_j^{(i)}}\qquad \begin{array}{l} j=0,1,\ldots \\ i=-j,-j+1,\ldots. \end{array}$$

The $\varepsilon_j^{(i)}$ can be ordered in a table as follows:

$$\begin{array}{cccccc}
 & \varepsilon_0^{(-1)}=0 & & \varepsilon_2^{(-2)}=0 & \ldots \\
\varepsilon_{-1}^{(0)}=0 & & \varepsilon_1^{(-1)} & & \ldots \\
 & \varepsilon_0^{(0)}=s_0 & & \varepsilon_2^{(-1)} & \ldots \\
\varepsilon_{-1}^{(1)}=0 & & \varepsilon_1^{(0)} & & \ldots \\
 & \varepsilon_0^{(1)}=s_1 & & \varepsilon_2^{(0)} & \ldots \\
\varepsilon_{-1}^{(2)}=0 & & \varepsilon_1^{(1)} & & \ldots \\
 & \varepsilon_0^{(2)}=s_2 & & \varepsilon_2^{(1)} & \ldots \\
\varepsilon_{-1}^{(3)}=0 & \vdots & \varepsilon_1^{(2)} & \vdots & \ldots \\
\vdots & & \vdots & &
\end{array}$$

So the index j refers to a column while i refers to a diagonal. The following property can be proved for the ε-algorithm. The proof is very technical and can be found in [2 pp. 44–46].

Let us denote by $\Delta s_i = s_{i+1} - s_i$.

Theorem 2.1.

$$\varepsilon_{2j}^{(i)}=\frac{\begin{vmatrix} s_{i+j} & \cdots & & s_i \\ \Delta s_{i+j} & \cdots & \Delta s_{i+1} & \Delta s_i \\ \vdots & \ddots & \vdots & \vdots \\ \Delta s_{i+2j-1} & & \Delta s_{i+j} & \Delta s_{i+j-1} \end{vmatrix}}{\begin{vmatrix} 1 & \cdots & 1 \\ \Delta s_{i+j} & \cdots & \Delta s_i \\ \vdots & & \vdots \\ \Delta s_{i+2j-1} & \cdots & \Delta s_{i+j-1} \end{vmatrix}}$$

with $s_i = 0$ *for* $i < 0$.

Now it is easy to see that for $s_i = \sum_{k=0}^{i} c_k x^k$

$$\varepsilon_{2m}^{(l-m)}=\frac{\begin{vmatrix} f_l(x) & \cdots & f_{l-m}(x) \\ c_{l+1}x^{l+1} & \cdots & c_{l-m+1}x^{l-m+1} \\ \vdots & & \vdots \\ c_{l+m}x^{l+m} & \cdots & c_l x^l \end{vmatrix}}{\begin{vmatrix} 1 & \cdots & 1 \\ c_{l+1}x^{l+1} & \cdots & c_{l-m+1}x^{l-m+1} \\ \vdots & & \\ c_{l+m}x^{l+m} & \cdots & c_l x^l \end{vmatrix}}.$$

Of course, we restrict ourselves always to finite $\varepsilon_j^{(i)}$. Since the algorithm is a nonlinear algorithm, it can always happen that $\varepsilon_j^{(i)}$ is indefinite (when $\varepsilon_{j-1}^{(i+1)} = \varepsilon_{j-1}^{(i)}$).

3. Multivariate Padé-Approximants

Let $f(x) = \sum_{k=0}^{\infty} c_k x^k$ where $x = (x_1, \ldots, x_n)$,

$$c_k x^k = \sum_{k_1 + \ldots + k_n = k} c_{k_1 \ldots k_n} x_1^{k_1} x_2^{k_2} \ldots x_n^{k_n}.$$

In other words $c_k x^k$ is a homogeneous polynomial of degree k in $x_1, \ldots, x_n$.

Let $p(x) = \sum_{i=lm}^{lm+l} a_i x^i$ and $q(x) = \sum_{j=lm}^{lm+m} b_j x^j$ where

$$a_i x^i = \sum_{i_1 + \ldots + i_n = i} a_{i_1 \ldots i_n} x_1^{i_1} \ldots x_n^{i_n}$$

and

$$b_j x^j = \sum_{j_1 + \ldots + j_n = j} b_{j_1 \ldots j_n} x_1^{j_1} \ldots x_n^{j_n}.$$

We say that $f(x) = O(x^k)$ if non-negative constants K and r with $r < 1$ exist, such that $|f(x)| \leq K \|x\|^k$ for $\|x\| < r$ [3 pp. 66].

Definition 3.1. If $p(x)$ and $q(x)$ satisfy

$$f(x) \cdot q(x) - p(x) = O(x^{lm+l+m+1}) \quad \text{and } q(x) \not\equiv 0,$$

then the irreducible form $R_{lm}(x) = \frac{p_*(x)}{q_*(x)}$ of $\frac{p(x)}{q(x)}$ is called the (l, m) multivariate Padé-approximant to f.

Polynomials $p(x)$ and $q(x)$ satisfying Definition 3.1 can be calculated by solving the following systems of equations:

$$\begin{cases} c_0 \cdot b_{lm} x^{lm} = a_{lm} x^{lm} \\ \vdots \\ c_l x^l \cdot b_{lm} x^{lm} + \ldots + c_0 \cdot b_{lm+l} x^{lm+l} = a_{lm+l} x^{lm+l} \end{cases} \qquad \text{for all } x \text{ in } \mathbb{R}^n$$

with $b_{lm+j} x^{lm+j} \equiv 0$ if $j > m$

$$\begin{cases} c_{l+1} x^{l+1} \cdot b_{lm} x^{lm} + \ldots + c_{l+1-m} x^{l+1-m} \cdot b_{lm+m} x^{lm+m} = 0 \\ \vdots \\ c_{l+m} x^{l+m} \cdot b_{lm} x^{lm} + \ldots + c_l x^l \cdot b_{lm+m} x^{lm+m} = 0 \end{cases} \qquad \text{for all } x \text{ in } \mathbb{R}^n$$

with $c_k x^k \equiv 0$ if $k < 0$.

The homogenous system contains m equations in the $m+1$ unknowns $b_{lm} x^{lm}, \ldots, b_{lm+m} x^{lm+m}$.

A solution is given by

$$b_{lm}x^{lm} = \begin{vmatrix} c_l x^l & \ldots c_{l+1-m}x^{l+1-m} \\ \vdots & \vdots \\ c_{l+m-1}x^{l+m-1} \ldots & c_l x^l \end{vmatrix},$$

$$b_{lm+j}x^{lm+j} = \begin{vmatrix} c_l x^l & \ldots & \boxed{\begin{matrix} -c_{l+1}x^{l+1} \ldots \\ \vdots \\ -c_{l+m}x^{l+m} \end{matrix}} & c_{l+1-m}x^{l+1-m} \\ \vdots & & & \vdots \\ c_{l+m-1}x^{l+m-1} & & & c_l x^l \end{vmatrix} \quad \text{for } j=1,\ldots,m$$

↑

j^{th} column in $b_{lm}x^{lm}$ replaced by this column.

Now it is easy to see that a representation of $p(x)$ and $q(x)$ satisfying Definition 3.1 is given in Theorem 3.1.

Theorem 3.1. *If*

$$\begin{vmatrix} 1 & \ldots & 1 \\ c_{l+1}x^{l+1}\, c_l x^l & \ldots & c_{l+1-m}x^{l+1-m} \\ \vdots & & \vdots \\ c_{l+m}x^{l+m} & \ldots & c_l x^l \end{vmatrix} \not\equiv 0$$

then

$$\frac{p(x)}{q(x)} = \frac{\begin{vmatrix} f_l(x)\, f_{l-1}(x) & \ldots & f_{l-m}(x) \\ c_{l+1}x^{l+1}\, c_l x^l & \ldots & c_{l+1-m}x^{l+1-m} \\ \vdots & & \vdots \\ c_{l+m}x^{l+m} & \ldots & c_l x^l \end{vmatrix}}{\begin{vmatrix} 1 & \ldots & 1 \\ c_{l+1}x^{l+1}\, c_l x^l \ldots & & c_{l+1-m}x^{l+1-m} \\ \vdots & & \vdots \\ c_{l+m}x^{l+m} & \ldots & c_l x^l \end{vmatrix}}$$

where $f_i(x) = \sum_{k=0}^{i} c_k x^k$.

We give a few examples to illustrate Theorem 3.1 and the translation of degrees in $p(x)$ and $q(x)$ by $l.m$.

Let

$$f(x_1, x_2) = 1 + \frac{x_1}{0.1 - x_2} + \sin(x_1 \cdot x_2)$$

$$= 1 + 10x_1 + 101x_1x_2 + \sum_{k=3}^{\infty} 10^k x_1 x_2^{k-1} + \sum_{k=1}^{\infty} (-1)^k \frac{(x_1x_2)^{2k+1}}{(2k+1)!}.$$

Take $l = 1 = m$.

The determinant

$$\begin{vmatrix} 1 & 1 \\ c_2 x^2 & c_1 x \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 101x_1x_2 & 10x_1 \end{vmatrix} \not\equiv 0.$$

So

$$\frac{p(x)}{q(x)}=\frac{\begin{vmatrix}1+10x_1 & 1\\ 101x_1x_2 & 10x_1\end{vmatrix}}{\begin{vmatrix}1 & 1\\ 101x_1x_2 & 10x_1\end{vmatrix}}=\frac{10x_1+100x_1^2-101x_1x_2}{10x_1-101x_1x_2}$$

and

$$R_{11}(x)=\frac{1+10x_1-10.1x_2}{1-10.1x_2}.$$

For $l=3$ and $m=1$ we get

$$\frac{p(x)}{q(x)}=\frac{\begin{vmatrix}1+10x_1+101x_1x_2+10^3x_1x_2^2 & 1+10x_1+101x_1x_2\\ 10^4x_1x_2^3 & 10^3x_1x_2^2\end{vmatrix}}{\begin{vmatrix}1 & 1\\ 10^4x_1x_2^3 & 10^3x_1x_2^2\end{vmatrix}}$$

$$=\frac{10^3x_1x_2^2(1+10x_1-10x_2+x_1x_2-10x_1x_2^2)}{10^3x_1x_2^2(1-10x_2)}$$

and

$$R_{31}(x)=\frac{1+10x_1-10x_2+x_1x_2-10x_1x_2^2}{1-10x_2}.$$

When $l=1$ and $m=2$ we have

$$\frac{p(x)}{q(x)}=\frac{\begin{vmatrix}1+10x_1 & 1 & 0\\ 101x_1x_2 & 10x_1 & 1\\ 10^3x_1x_2^2 & 101x_1x_2 & 10x_1\end{vmatrix}}{\begin{vmatrix}1 & 1 & 1\\ 101x_1x_2 & 10x_1 & 1\\ 10^3x_1x_2^2 & 101x_1x_2 & 10x_1\end{vmatrix}}$$

$$=\frac{100x_1(x_1-1.01x_2+10x_2^2+10x_1^2-20.2x_1x_2)}{100x_1(x_1-1.01x_2+10x_2^2-10.1x_1x_2+2.01x_1x_2^2)}$$

and

$$R_{12}(x)=\frac{x_1-1.01x_2+10x_2^2+10x_1^2-20.2x_1x_2}{x_1-1.01x_2+10x_2^2-10.1x_1x_2+2.01x_1x_2^2}.$$

If we apply the ε-algorithm now to the row $\left\{s_i=\sum_{k=0}^{i} c_k x^k\ i=0,1,\ldots\right\}$ we get for $\varepsilon_{2m}^{(l-m)}$ as a result of Theorem 2.1, precisely the quotient of determinants given in Theorem 3.1. So the multivariate Padé-approximants given by Definition 3.1 can be calculated recursively. To conclude we fill up a part of the ε-table and find back the multivariate Padé-approximants for $f(x_1,x_2)=1+\dfrac{x_1}{0.1-x_2}+\sin(x_1\cdot x_2)$.

Table 1.

	0		0		0
0		1		$2-10x_1$	
	1		$\frac{1}{1-10x_1}$		$\frac{1}{1-10x_1-101x_1x_2+100x_1^2}$
0		$\frac{1}{10x_1}$		$\frac{1-20x_1+101x_1x_2}{10x_1(10\cdot 1x_2-10x_1)}$	
	$1+10x_1$		$\frac{1-10.1x_2+10x_1}{1-10.1x_2}$		$\frac{x_1-1.01x_2+10x_2^2+10x_1^2-20.2x_1x_2}{x_1-1.01x_2+10x_2^2-10.1x_1x_2+2.01x_1x_2^2}$
0		$\frac{1}{101x_1x_2}$		$\frac{1-20.2x_2+100x_2^2}{-20.1x_1x_2^2}$	
	$1+10x_1+101x_1x_2$		$\frac{1+10x_1+101x_1x_2-\frac{10^3}{101}x_2-\frac{10^4}{101}x_1x_2}{1-\frac{10^3}{101}x_2}$		
0		$\frac{1}{10^3x_1x_2^2}$			
	$1+10x_1+101x_1x_2+10^3x_1x_2^2$		$\frac{1+10x_1-10x_2+x_1x_2-10x_1x_2^2}{1-10x_2}$		
0		$\frac{1}{10^4x_1x_2}$			
	$1+10x_1+101x_1x_2+10^3x_1x_2^2$ $+10^4x_1x_2^3$				

References

1. Brezinski, C.: Algoritmes d'accélération de la convergence. Editions Technip, Paris 1978
2. Brezinski, C.: Accélération de la convergence en analyse numérique. LNM 584, Berlin: Springer 1977
3. Cuyt, A.: Abstract Padé Approximants in operator theory, pp-61-87. LNM 765, Berlin: Springer 1979
4. Perron, O.: Die Lehre von den Kettenbrüchen II. Stuttgart: Teubner 1977

Received March 19, 1982