THE MECHANISM OF THE MULTIVARIATE PADE PROCESS

Annie Cuyt
Department of Mathematics UIA
Universiteitsplein 1
2610 Wilrijk (Antwerp) / Belgium

Abstract

In [3] (n,m) multivariate Padé approximants were introduced by means of a shift of the degrees in numerator and denominator over nm This definition is repeated here in section 3. In various papers many properties of those Padé approximants were proved; the analogy with the univariate case is remarkable. Here we show that the shift of the degrees over nm also arises in a natural way if we want to preserve some numerical algorithms or some geometrical pictures. Thus the paper provides new insights into the mechanism of the multivariate Padé process, and also some compact formulas for the multivariate Padé approximant itself.

1. The ε-algorithm and the qd-algorithm

Consider a series $\sum_{i=0}^{\infty} t_i$ in $\mathbb{R}$ and also the sequence $(s_i)_{i\in\mathbb{N}}$ of its partial sums; so $s_i = t_0 + \ldots + t_i$.

Input of the ε-algorithm are the elements s_i . We perform the following computations:

a) $$\left.\begin{aligned} \varepsilon_{-1}^{(i)} &= 0 \\ \varepsilon_{0}^{(i)} &= s_i \end{aligned}\right. \qquad i = 0,1,\ldots$$

b) $$\varepsilon_{2j}^{(-j-1)} = 0 \qquad j = 0,1,\ldots$$

c) $\varepsilon_{j+1}^{(i)} = \varepsilon_{j-1}^{(i+1)} + \dfrac{1}{\varepsilon_j^{(i+1)} - \varepsilon_j^{(i)}}$ $\qquad j = 0,1,\ldots \quad i = -j,-j+1,\ldots$

The index j refers to a column while i refers to a diagonal in the ε-table. If the algorithm does not break down the following property can be proved for the ε-algorithm. The proof is very technical and can be found in [1 pp. 44-46]. We denote by $\Delta s_k = s_{k+1} - s_k$.

Theorem 1.1.:

With $s_i = 0$ for $i < o$

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

The relation of the ε-algorithm with the multivariate Padé process and the geometrical picture that we will set up is explained in the following sections.

Input of the qd-algorithm are the terms t_i . One performs the following calculations:

a) $e_o^{(i)} = 0$, $q_1^{(i)} = \dfrac{t_{i+1}}{t_i}$ $\qquad i = 0,1,\ldots$

b) $e_j^{(i)} = q_j^{(i+1)} + e_{j-1}^{(i+1)} - q_j^{(i)}$ $\qquad i = 0,1,2,\ldots \qquad j = 1,2,\ldots$

c) $q_{j+1}^{(i)} = q_j^{(i+1)} \cdot e_j^{(i+1)} / e_j^{(i)}$ $\qquad i = 0,1,2,\ldots \qquad j = 1,2,\ldots$

Again the index j refers to a column while i refers to a diagonal. If all the $q_j^{(i)}$ and $e_j^{(i)}$ exist, one can prove the following property [2].

Theorem 1.2.:

For $s_i = t_o + \ldots + t_i$,

$$\varepsilon_{2j}^{(i)} = s_i + \frac{t_{i+1}|}{|1} - \frac{q_1^{(i+1)}|}{|1} - \frac{e_1^{(i+1)}|}{|1} - \frac{q_2^{(i+1)}|}{|1} - \frac{e_2^{(i+1)}|}{|1} - \ldots - \frac{q_j^{(i+1)}|}{|1}$$

The qd-algorithm will also be used to set up a certain geometrical picture and to provide the multivariate Padé approximants defined in [3]. Let us denote the partial numerators of the continued fraction given above by

$a_{ki}(k = 0,\ldots,2j)$; so $a_{oi} = s_i$, $a_{1i} = t_{i+1}$, $a_{ki} = -q_{\frac{k}{2}}^{(i+1)}$ if k is even and $a_{ki} = -e_{\frac{k-1}{2}}^{(i+1)}$ if k is odd.

2. Geometrical picture

Let us now construct with the $2m + 1$ numbers $s_{n-m},\ldots,s_{n+m}$ the vectors

$$s^{(k)} = (s_{n-m+k}, \Delta s_{n-m+k}, \ldots, \Delta s_{n+k-1})^t \quad \text{in } \mathbb{R}^{m+1}$$

for $k = o,\ldots,m$. With the partial numerators $a_{o,n-m},\ldots,a_{2m,n-m}$ we construct the vectors

$$t^{(k)} = (\underbrace{0,\ldots,0}_{k\ \text{times}},a_{k,n-m},1,-1,0,\ldots,0) \quad \text{in } \mathbb{R}^{2m+2}$$

for $k = o,\ldots,2m-1$, and the vector

$$t^{(2m)} = (\underbrace{0,\ldots,0}_{2m\ \text{times}},a_{2m,n-m},1)$$

A) We can draw an m-dimensional hyperplane through the points $s^{(k)}$ $(k=o,\ldots,m)$ in $\mathbb{R}^{m+1}$. Suppose that the vector normal to that hyperplane, is given by $u = (u_o,\ldots,u_m)^t$.

Then we have

$$u.s^{(k)} = u_o.s_{n-m+k} + \sum_{i=1}^{m} u_i \ \Delta s_{n-m+k+i-1} = 0 \quad k = o,\ldots,m \ . \tag{1}$$

We call $(v_m,0,\ldots,0)$ the point where that hyperplane intersects the first axis. Then also

$$u_o \ . \ v_m = u.s^{(k)} \ . \tag{2}$$

From (1) and (2) we obtain

$$s_{n-m+k} = -\sum_{i=1}^{m} \frac{u_i}{u_o} \Delta s_{n-m+k+i-1} + v_m \quad k = o,\ldots,m$$

which we can write as a linear system of equations

$$A \begin{pmatrix} v_o \\ \cdot \\ \cdot \\ \cdot \\ v_m \end{pmatrix} = \begin{pmatrix} s_{n-m} \\ \cdot \\ \cdot \\ \cdot \\ s_n \end{pmatrix} \tag{3}$$

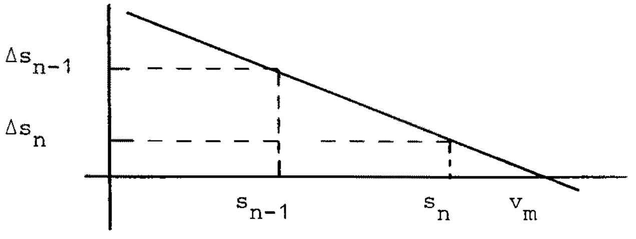
where $v_i = \dfrac{-u_{i+1}}{u_o}$ $(i = o,\ldots,m-1)$ and the matrix A is given by

$$A = \begin{pmatrix} \Delta s_{n-m} & \Delta s_{n-m+1} & \cdots & \Delta s_{n-1} & 1 \\ \vdots & & & \vdots & \vdots \\ \Delta s_n & \cdots & & \Delta s_{n+m-1} & 1 \end{pmatrix}$$

Cramer's rule for the solution of such a system then gives:

$$v_m = \frac{\begin{vmatrix} \Delta s_{n-m} & \cdots & \Delta s_{n-1} & s_{n-m} \\ \vdots & & & \\ \Delta s_n & \cdots & \Delta s_{n+m-1} & s_n \end{vmatrix}}{\begin{vmatrix} \Delta s_{n-m} & \cdots & \Delta s_{n-1} & 1 \\ \vdots & & & \vdots \\ \Delta s_n & \cdots & \Delta s_{n+m-1} & 1 \end{vmatrix}}$$

So clearly $\varepsilon_{2m}^{(n-m)}$ is the last unknown of the system of equations (3). Let us draw a picture in the case $m = 1$. Then $s^{(o)}$ and $s^{(1)}$ are two vectors in $\mathbb{R}^2$ through which a straight line is drawn and v_1 is the abscis of the intersection with the first axis.



So v_m results from extrapolating certain differences to zero. That is why we expect v_m to be more efficient the larger n is and to be an estimate of the limit of the sequence $(s_i)_{i \in \mathbb{N}}$ if it exists.

B) If the vector $u = (u_o,\ldots,u_{2m+1})$ in $\mathbb{R}^{2m+2}$ is in the 1-dimensional subspace orthogonal on $\{t^{(k)} | k = 0,\ldots,2m\}$, then

$$u.t^{(k)} = 0 \qquad k = o,\dots,2m \tag{4}$$

which we can write as a linear system

$$Bu = 0 \tag{5}$$

where the $(2m+1)\times(2m+2)$ matrix B is given by

$$B = \begin{pmatrix} a_{o,n-m} & 1 & -1 & & 0 \\ & a_{1,n-m} & 1 & -1 & \\ 0 & & \ddots & \ddots & -1 \\ & & & a_{2m,n-m} & 1 \end{pmatrix}$$

In [8] Miklosko proved that

$$t_o + \dots + t_{n-m} + \frac{t_{n-m+1}|}{|1} - \frac{q_1^{(n-m+1)}|}{|1} - \frac{e_1^{(n-m+1)}|}{|1} - \dots - \frac{q_m^{(n-m+1)}|}{|1}$$

is equal to $\frac{-u_1}{u_o}$, or in other words that it is the first unknown v_1 of the linear system of equations

$$C \begin{pmatrix} v_1 \\ \vdots \\ v_{2m+1} \end{pmatrix} = \begin{pmatrix} a_{o,n-m} \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{6}$$

where C is given by

$$C = \begin{pmatrix} 1 & -1 & & 0 \\ a_{1,n-m} & \ddots & \ddots & -1 \\ 0 & \ddots & a_{2m,n-m} & 1 \end{pmatrix}$$

So $\varepsilon_{2m}^{(n-m)}$ is also the first unknown of the system of equations (6).

3. Multivariate Padé approximants

Let $f(x) = \sum_{k=o}^{\infty} c_k x^k$ where $x = (x_1,\ldots,x_\ell)$ and where

$$c_k x^k = \sum_{k_1+\ldots+k_\ell=k} c_{k_1\ldots k_\ell} x_1^{k_1}\ldots x_\ell^{k_\ell} .$$

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

$$a_i x^i = \sum_{i_1+\ldots+i_\ell=i} a_{i_1\ldots i_\ell} x_1^{i_1}\ldots x_\ell^{i_\ell}$$

$$b_j x^j = \sum_{j_1+\ldots+j_\ell=j} b_{j_1\ldots j_\ell} x_1^{j_1}\ldots x_\ell^{j_\ell} .$$

<u>Definition 3.1.:</u> If $p(x)$ and $q(x)$ satisfy

$$(f.q-p)(x_1,\ldots,x_\ell) = \sum_{k_1+\ldots+k_\ell\ge nm+n+m+1} d_{k_1\ldots k_\ell} x_1^{k_1}\ldots x_\ell^{k_\ell}$$

then the irreducible form $R_{n,m}(x)$ of $\frac{p(x)}{q(x)}$ is called the (n,m) multivariate Padé approximant for $f(x_1,\ldots,x_\ell)$.

The shift of the degrees in $p(x)$ and $q(x)$ by nm , has already been motivated in [4]. More about there multivariate Padé approximants can also be found in [3]. We shall now see that the shift of the degrees does also match the geometrical picture and that this geometrical picture provides some very compact formulas for $\frac{p(x)}{q(x)}$.

In [5] we proved that $\frac{p}{q}(x)$ was given by $\varepsilon_{2m}^{(n-m)}$ if $s_i = \sum_{k=o}^{i} c_k x^k$ was the multivariate partial sum of the multivariate Taylor series $f(x)$, i.e. if $t_i = c_i x^i$. Here we have seen that $\varepsilon_{2m}^{(n-m)}$ is also the last unknown of the system (3) which results from extrapolation to zero, since $\varepsilon_{2m}^{(n-m)}$ is the intersection-point of the interpolating hyperplane through the $s^{(k)}$ and the first axis. This enables us to write down the following compact expression for $\frac{p}{q}(x)$

$$\frac{p}{q}(x) = \left[A^{-1} \begin{pmatrix} s_{n-m} \\ \cdot \\ \cdot \\ \cdot \\ s_n \end{pmatrix} \right]_m \tag{7}$$

with $s_i = \sum_{k=o}^{i} c_k x^k$.

We have also shown that $\varepsilon_{2m}^{(n-m)}$ is the first unknown of the system of equations (6), so another compact formula for $\frac{p}{q}(x)$ is given by

$$\frac{p}{q}(x) = \left[C^{-1} \begin{pmatrix} s_{n-m} \\ 0 \\ \cdot \\ \cdot \\ \cdot \\ 0 \end{pmatrix} \right]_1 \tag{8}$$

For the univariate Padé approximants formula (7) can be found in [6] where s_i is the i^{th} partial sum of the univariate Taylor series, and formula (8) is a consequence of theorem 1.2. and Miklosko's result where t_i is the term of degree i in the univariate Taylor series. So if we want to preserve the univariate geometrical picture, we can for instance "define" the multivariate Padé approximant by means of (7) or (8). This automatically results in a shift of the degrees in $p(x_1,\ldots,x_\ell)$ and $q(x_1,\ldots,x_\ell)$ because we have proved here the validity of (7) and (8) for the multivariate Padé approximants given in definition 3.1.

As a consequence, the conclusion is now that the most natural way to generalize the concept of Padé approximant for multivariate functions is by means of definition 3.1.

References

[1] C. Brezinski:
Accélération de la convergence en analyse numérique.
LNM 584, Springer, Berlin (1977)

[2] C. Brezinski:
Padé-type approximation and general orthogonal polynomials.
ISNM 50, Birkhäuser Verlag, Basel (1980)

[3] A. Cuyt:
Multivariate Padé approximants.
Journ. Math. Anal. Applcs. 96 (1). 283-293 (1983)

[4] A. Cuyt:
Abstract Padé Approximants in Operator Theory: Theory and Applications.
LNM 1065, Springer Verlag, Berlin Heidelberg (1984)

[5] A. Cuyt:
The ε-algorithm and multivariate Padé approximants.
Numerische Mathematik 40, 39-46 (1982)

[6] R. Johnson:
Alternative approach to Padé approximants.
In [7], 53 - 67

[7] P. Graves-Morris:
Padé approximants and their applications.
Academic Press, New York (1973)

[8] J. Miklosko:
Investigation of algorithms for numerical computation of continued fractions.
USSR Comp. Math. and Math. Phys. 16(4), 1-12 (1976)