From Hankel, Aitken and Wynn to de Prony, Rutishauser and Henrici

Annie Cuyt
Department of Mathematics and Computer Science
Universiteit Antwerpen (CMI)
Middelheimlaan 1, B-2020 Antwerpen (Belgium)
annie.cuytuantwerpen.be

October 14, 2019

Abstract

What do the epsilon-algorithm, Padé approximants, Gaussian quadrature rules, exponential analysis and tensor decomposition have in common? Hankel matrices! While Padé approximants belong to rational approximation theory, Gaussian quadrature rules are high order numerical integration rules derived from orthogonal polynomial families, exponential analysis is a parametric spectral analysis tool and tensor decomposition is situated in mathematical data analysis.

Between all these applications, computational tools (the qd-algorithm, polynomial interpolation, generalized eigenvalue solvers, ...) and mathematical theorems (convergence results for Padé approximants, Froissart doublet behaviour, computational complexity analysis, ...) can now be exchanged because it is possible to rewrite one problem statement in the form of another. This opens up a whole new world to explore.

The formally orthogonal Hadamard polynomials introduced in Section 1 are related to Padé denominator polynomials in Section 2. The Hadamard polynomial zeroes form the nodes of Gaussian quadrature rules in Section 3 and in Section 4 the Hadamard polynomial equals the so-called Prony polynomial of exponential analysis. The zeroes are also the nodes of the Vandermonde structured factor matrices in Section 5 in the decomposition of a higher order tensor with Hankel structured slices.

All of the above have been generalized to higher dimensions, with the preservation of several connections, as briefly discussed in Section 6.

Introduction

Research in numerical approximation theory introduces one to many great mathematicians, such as the already mentioned Hankel, Sylvester, Schweins, Aitken, Wynn, Rutishauser, de Prony, Padé, Henrici, Brezinski, to name just a few. They are the giants on whose shoulders we can stand.

The rich world of determinant identities leads to Wynn's epsilon-algorithm and gives rise to several generalizations when moving to more general concepts than complex numbers, such as vectors, multivariate functions, matrices, tensors, nonlinear operators, etc.

Less well-known but equally important are the connections between several seemingly disjoint topics. For instance, the connection between sparse interpolation studied in the computer algebra community and Padé approximation, was already pointed out in [23] by connecting to the method of de Prony [18], which is at the basis of the parametric spectral analysis method called exponential analysis. The connection with formal orthogonal polynomials, Padé-type approximation and Gaussian quadrature rules are further elaborated by Brezinski in [4]. More recently, we completed the circle by relating all of the above to tensor decomposition methods [11] and establishing several multivariate generalizations, thereby preserving as many connections as possible.

Let us now discuss the different problem statements one by one, pointing out the relationships as we go along. Only at the end we briefly touch upon their multivariate versions.

1 Classical orthogonal polynomials

We denote $\mathbb{C}[z]$ for the linear space of polynomials in the variable z with complex coefficients and define the linear functional $\lambda : \mathbb{C}[z] \to \mathbb{C}$ which associates the number e_i with the monomial t^i, so

$$\lambda(t^i) = e_i, \qquad i = 0, 1, \ldots$$

Formally we can write

$$\sum_{i=0}^{\infty} e_i z^i = \lambda\left(\frac{1}{1-tz}\right).$$

The sequence of polynomials $V_m(z), m = 0, 1, \ldots$ given by

$$V_m(z) = \sum_{j=0}^{m} b_{m-j} z^j = b_m + b_{m-1}z + \ldots + b_0 z^m, \qquad b_0 \neq 0$$

and satisfying the conditions

$$\lambda(t^i V_m(t)) = 0, \qquad i = 0, \ldots, m-1, \tag{1}$$

is a sequence of polynomials orthogonal with respect to the linear functional λ [4, p. 40] since (1) implies that

$$\lambda(V_k(t)V_m(t)) = 0, \qquad k \neq m.$$

Let us further define the Hankel determinants

$$H_m^{(s)} = \begin{vmatrix} e_s & \dots & e_{s+m-1} \\ \vdots & & \\ & & \vdots \\ e_{s+m-1} & \dots & e_{s+2m-2} \end{vmatrix}, \qquad H_0^{(s)} = 1, \qquad s = 0, 1, \dots$$

The linear functional λ is called definite if

$$H_m^{(0)} \neq 0, \qquad m = 0, 1, \dots$$

From (1), the linear system

$$\sum_{j=0}^{m} e_{i+j} b_{m-j} = 0, \qquad i = 0, \dots, m-1 \tag{2}$$

is obtained directly, which allows to compute the orthogonal polynomial $V_m(z)$ up to a normalization. If λ is definite then the linear system of equations (2) has maximal rank. With $b_0 = 1$, the monic orthogonal polynomial $V_m(z)$ is then given by

$$V_m(z) = \frac{1}{H_m^{(0)}} \begin{vmatrix} e_0 & \dots & e_{m-1} & e_m \\ \vdots & & & \vdots \\ e_{m-1} & & \dots & e_{2m-1} \\ 1 & z & \dots & z^m \end{vmatrix}, \qquad V_0(z) = 1.$$

In [19] the orthogonal polynomial $V_m(z)$ is called the formally orthogonal Hadamard polynomial.

2 Connection with Padé approximation

Let at first $V_m(z)$ be general and not necessarily satisfy (1). For a given $V_m(z)$ we define the associated polynomials $W_{m-1}(z)$ of degree $m-1$ by [4, p. 10]

$$\begin{aligned} W_{m-1}(z) :&= \lambda \left(\frac{V_m(z) - V_m(t)}{z - t} \right) \\ &= \lambda \left(b_{m-1} + b_{m-2}(z+t) + \dots + b_0(z^{m-1} + z^{m-2}t + \dots + zt^{m-2} + t^{m-1}) \right). \end{aligned}$$

We also define the reverse polynomials

$$\begin{aligned} \tilde{V}_m(z) &:= z^m V_m(1/z) \\ \tilde{W}_{m-1}(z) &:= z^{m-1} W_{m-1}(1/z) \end{aligned}$$

which jointly satisfy [4, p. 11]

$$\sum_{i=0}^{\infty} e_i z^i - \frac{\tilde{W}_{m-1}(z) i}{\tilde{V}_m(z)} = \sum_{i=m}^{\infty} d_i z^i. \tag{3}$$

The rational function $\tilde{W}_{m-1}(z)/\tilde{V}_m(z)$ is called a Padé-type approximant of degree $m-1$ in the numerator and m in the denominator to the formal power series

$$F(z) := \sum_{i=0}^{\infty} e_i z^i. \tag{4}$$

Now let $V_m(z)$ satisfy the orthogonality conditions (1). Then (3) improves to [4, pp. 34–35]

$$\sum_{i=0}^{\infty} e_i z^i - \frac{\tilde{W}_{m-1}(z)}{\tilde{V}_m(z)} = \sum_{i=2m}^{\infty} d_i z^i. \tag{5}$$

The rational function $\tilde{W}_{m-1}(z)/\tilde{V}_m(z)$ is then called a Padé approximant of degree $m-1$ in the numerator and m in the denominator to the series $F(z)$ and often denoted by $[m-1/m]_F(z)$. So the reverse of the Padé denominator

$$\tilde{V}_m(z) = \sum_{j=0}^{m} b_{m-j} z^{m-j},$$

which can be computed from the linear system (2), is the orthogonal polynomial $V_m(z)$. After computing its coefficients b_j, the Padé numerator

$$\tilde{W}_m(z) = \sum_{j=0}^{m-1} a_j z^j$$

can be obtained from the linear system

$$\sum_{j=0}^{m} e_{i-j} b_j = a_i, \qquad i = 0, \dots, m-1,$$

where $e_i := 0$ for $i < 0$.

3 Connection with Gaussian quadrature

In the sequel of this section we consider z as a parameter and t as the variable. Also the numbers e_i are classical moments, given on the standard interval $[-1, 1]$ for the weight function $w(z)$,

$$e_i := \int_{-1}^{1} w(t) t^i \, dt, \qquad \int_{-1}^{1} w(t) \, dt > 0.$$

The formal power series $F(z)$ defined in (4) then equals

$$F(z) = \lambda\left(\frac{1}{1-tz}\right) = \int_{-1}^{1} w(t) \frac{1}{1-tz} \, dt.$$

Let $z_i^{(m)}, i = 1, \ldots, m$ denote the zeroes of $V_m(z)$, so with $b_0 = 1$ we have

$$V_m(z) = (z - z_1^{(m)}) \cdots (z - z_m^{(m)}).$$

Consider the Hermite interpolating polynomial $p_{m-1}(t; z)$ for $1/(1-tz)$ through the interpolation points $z_1^{(m)}, \ldots, z_m^{(m)}$, so

$$\begin{aligned} p_{m-1}(t; z) &= \sum_{i=1}^{m} \left(\frac{1}{1 - z_i^{(m)} z} \right) \frac{V_m(t)}{(t - z_i^{(m)}) V_m'(z_i^{(m)})} \\ &= \sum_{i=1}^{m} \left(\frac{1}{1 - z_i^{(m)} z} \right) \frac{V_m(t) - V_m(z_i^{(m)})}{(t - z_i^{(m)}) V_m'(z_i^{(m)})} . \end{aligned}$$

Then the approximation

$$\begin{aligned} \lambda \left(\frac{1}{1 - tz} \right) &\approx \lambda \left(p_{m-1}(t; z) \right) \\ &= \sum_{i=1}^{m} \lambda \left(\frac{V_m(t) - V_m(z_i^{(m)})}{(t - z_i^{(m)}) V_m'(z_i^{(m)})} \right) \frac{1}{1 - z_i^{(m)} z} \end{aligned}$$

is a quadrature rule for the integration of $1/(1 - tz)$, with nodes $z_1^{(m)}, \ldots, z_m^{(m)}$ and weights

$$\begin{aligned} A_i^{(m)} &= \lambda \left(\frac{V_m(t) - V_m(z_i^{(m)})}{(t - z_i^{(m)}) V_m'(z_i^{(m)})} \right) \\ &= \frac{W_{m-1}(z_i^{(m)})}{V_m'(z_i^{(m)})}. \end{aligned}$$

In other words

$$\lambda \left(\frac{1}{1 - tz} \right) \approx \sum_{i=1}^{m} A_i^{(m)} \frac{1}{1 - z_i^{(m)} z}.$$

When the $z_i^{(m)}, i = 1, \ldots, m$ are all distinct, then the quadrature rule is a Gaussian rule guaranteeing correctness up to an including polynomials of degree $2m - 1$: for $q(z) \in \mathbb{C}[z]$ of degree $2m - 1$ we have

$$\lambda \left(q(t) \right) = \int_{-1}^{1} q(t) \, dt = \sum_{i=1}^{m} A_i^{(m)} q(z_i^{(m)}).$$

So the nodes and weights of a formal m-point Gaussian quadrature rule are closely connected with the orthogonal and associated polynomials $V_m(z)$ and $W_{m-1}(z)$.

4 Connection with exponential analysis

Let us consider a different situation, where the e_i follow a very structured rule,

$$e_i = \lambda(t_i) = \sum_{j=1}^{m} \alpha_j \Phi_j^i, \qquad \alpha_j, \Phi_j \in \mathbb{C}. \tag{6}$$

Here the $\Phi_j, j = 1, \ldots, m$ are (for simplicity) assumed to be mutually distinct and given by $\Phi_j = \exp(2\pi\phi_j/M)$ with $\phi_j, j = 1, \ldots, m \in \mathbb{C}$ and M satisfying $|\Im(\phi_j)| < M/2$ to avoid periodicity problems [22, 14].

Because of the nature of the e_i, the Hankel matrices

$$\mathcal{H}_m^{(s)} = \begin{pmatrix} e_s & \cdots & e_{s+m-1} \\ \vdots & & \\ & & \vdots \\ e_{s+m-1} & \cdots & e_{s+2m-2} \end{pmatrix}, \qquad \mathcal{H}_0^{(s)} = 1, \qquad s = 0, 1, \ldots$$

can be factorized as [13]

$$\mathcal{H}_m^{(s)} = \mathcal{V}_m \mathcal{D}_\alpha \mathcal{D}_\Phi^s \mathcal{V}_m^T \tag{7}$$

where $\mathcal{V}_m$ is the Vandermonde matrix

$$\mathcal{V}_m = \left(\Phi_j^{i-1}\right)_{i,j=1}^m$$

and $\mathcal{D}_\alpha$ and $\mathcal{D}_\Phi$ are $m \times m$ diagonal matrices filled with the vectors $(\alpha_1, \ldots, \alpha_m)$ and $(\Phi_1, \ldots, \Phi_m)$. In addition, the power series $F(z)$ given by (4) reduces to the rational function

$$\begin{aligned} F(z) &= \sum_{i=0}^{\infty} \lambda(t^i) z^i \\ &= \sum_{i=0}^{\infty} \left(\sum_{j=1}^{m} \alpha_j \Phi_j^i \right) z^i = \sum_{j=1}^{m} \frac{\alpha_j}{1 - \Phi_j z} . \end{aligned}$$

Because of the consistency property for Padé approximants [4, p. 36], the denominator polynomials of $F(z)$ and $[m-1/m]_F(z)$ must equal

$$(1 - \Phi_1 z) \cdots (1 - \Phi_m z)$$

which in its turn equals $\tilde{V}_m(z)$. The zeroes $z_i^{(m)}, i = 1, \ldots, m$ of $V_m(z)$, which are the Gaussian quadrature nodes in the previous section, are in this exponentially structured special case given by $z_i^{(m)} = \Phi_i$. In the literature, the polynomial $V_m(z)$ is called the Prony polynomial [20]: its zeroes Φ_i are the atoms in the exponential sum (6).

In case the ϕ_j and α_j are given, then the e_i can be computed. In the inverse case, where the e_i are somehow known or measured, the ϕ_j and α_j can be extracted from the $e_0, \ldots, e_{2m-1}$ as follows. Because of (7), the Φ_j are the generalized eigenvalues [21] of

$$\mathcal{H}_m^{(1)} v_j = \Phi_j \mathcal{H}_m^{(0)} v_j$$

where the $v_j, j = 1, \ldots, m$ are the right generalized eigenvectors. The constraint $|\Im(\phi_j)| < M/2$ now allows to extract the complex value ϕ_j unambiguously from the generalized eigenvalue Φ_j. With the Φ_j known, the linear coefficients α_j are obtained from the Vandermonde system of interpolation conditions

$$\sum_{j=1}^{m} \alpha_j \Phi_j^i = e_i, \qquad i = s, \ldots, s+m-1, \qquad 0 \le s \le m.$$

The extraction of the nonlinear parameters ϕ_j and the linear coefficients α_j from the equidistant samples

$$e_i = \sum_{j=1}^{m} \alpha_j \exp(2\pi(i/M)\phi_j), \qquad i = 0, \ldots, 2m-1$$

is a frequent problem statement in signal processing.

5 Connection with tensor decomposition

With the e_i given by (6), we fill an order n tensor $T \in \mathbb{C}^{m_1 \times \cdots \times m_n}$ where

$$2 \le m_k \le m, \qquad 1 \le k \le n, \qquad 3 \le n \le 2m-1,$$
$$\sum_{k=1}^{n} m_k = 2m + n - 1$$

and

$$T_{i_1,\ldots,i_n} := e_{i_1+\cdots+i_n-n}, \qquad 1 \le i_k \le m_k. \tag{8}$$

The tensor of smallest order $n = 3$ is of size $m \times m \times 2$ and the one of largest order $n = 2m-1$ is symmetric and of size $2 \times \cdots \times 2$. For the sequel we generalize the definition of the square Hankel matrix above to cover rectangular Hankel structured matrices

$$\mathcal{H}_{m_1,m_2}^{(s)} = \begin{pmatrix} e_s & e_{s+1} & \cdots & e_{s+m_2-1} \\ e_{s+1} & e_{s+2} & \cdots & e_{s+m_2} \\ \vdots & \vdots & \ddots & \vdots \\ e_{s+m_1-1} & e_{s+m_1} & \cdots & e_{s+m_1+m_2-2} \end{pmatrix}.$$

The tensor slices $T_{\cdot,\cdot,i_3,\ldots,i_n}$ of our tensor T then equal

$$T_{\cdot,\cdot,i_3,\ldots,i_n} = H_{m_1,m_2}^{(i_3+\cdots+i_n-n+2)}$$

and so are Hankel structured. The tensor T decomposes as

$$T = \sum_{j=1}^{m} \alpha_j \begin{pmatrix} 1 \\ \Phi_j \\ \vdots \\ \Phi_j^{m_1-1} \end{pmatrix} \circ \cdots \circ \begin{pmatrix} 1 \\ \Phi_j \\ \vdots \\ \Phi_j^{m_n-1} \end{pmatrix}, \qquad (9)$$

where $\circ$ denotes the outer product and the Φ_j are mutually distinct. The decomposition (9) is easily verified by checking the element at position $(i_1, \ldots, i_n)$ in the left hand side and the right hand side of (9). The factor matrices are the rectangular Vandermonde structured matrices

$$\mathcal{V}_{m_k,m} = \left(\Phi_j^{i-1}\right)_{i=1,j=1}^{m_k,m}, \qquad 1 \leq k \leq n.$$

Because of the Vandermonde structure of the factor matrices with $m_k \leq m, k = 1, \ldots, n$, their Kruskal rank equals m_k for all k. Since $m_1 + \cdots + m_n = 2m + n - 1$ we find that the sum of the Kruskal ranks of the n factor matrices of the rank m tensor T is bounded below by $2m + n - 1$. Hence the Kruskal condition is satisfied and the unicity of the decomposition is guaranteed.

6 Multidimensional generalizations

The concept of the formally orthogonal polynomial $V_m(z)$ is generalized in [10], for different radial weight functions, to so-called spherical orthogonal polynomials. The latter differ from several other definitions of multivariate orthogonal polynomials, in that they preserve the connections laid out here in the Sections 2 and 3. At the heart are again Hankel matrices and determinants, but now parameterized [9, 17].

Homogeneous multivariate Padé approximants, as defined in [6, 7], can be computed using the epsilon-algorithm [5] and can also be obtained from the spherical orthogonal polynomials in a similar way as described here in Section 2 [1, 2]. The homogeneous definition satisfies a very strong projection property, in the sense that this multivariate Padé approximant reduces to the univariate Padé approximant on every one-dimensional subspace.

A whole lot of Gaussian cubature rules on the disk can be united in a single approach when developing the rules from these spherical orthogonal polynomials [3]. What's more, the nodes and weights of such Gaussian cubature rules on the disk can be obtained as the solution of a multivariate Prony-like system of interpolation conditions [3]. And this brings us to the next connection.

The result that an m-term exponential analysis problem of the form

$$\sum_{j=1}^{m} \alpha_j \exp(2\pi(i/M)\phi_j) = e_i, \qquad i = 0, \ldots, 2m-1$$

can be solved uniquely for the $\alpha_j, \phi_j, j = 1, \ldots, m$ from only $2m$ samples e_i was only recently generalized to the d-dimensional setting [15] in its full flavour.

The equations

$$\sum_{j=1}^{m} \alpha_j \exp(\langle x_i, \phi_j \rangle) = e_i, \qquad \phi_j in \mathbb{C}^d, \qquad x_i \in \mathbb{R}^d,$$

can be solved for the vectors ϕ_j and the coefficients α_j from a mere $(d+1)m$ samples collected at vectors x_i. This number of samples is also the theoretical minimal number [15]. While the $2m$ samples in the one-dimensional problem are collected uniformly, the $(d+1)m$ samples in the d-dimensional problem are located on d parallel lines constructed from a basis in $\mathbb{R}^d$.

This multivariate Prony generalization is in its turn closely connected to both multivariate and multidimensional generalizations of the Padé approximant concept and to various tensor decompositions [16, 11]. Among other things, we also mention an algorithm to locate zeroes of the homogeneous Padé denominator [8], which, as we know from Section 4, are related to the atoms in the exponential analysis problem. The detailed interpretation of the output of this algorithm is through the convergence result for homogeneous Padé approximants given in [12].

References

[1] B. Benouahmane. Approximants de Padé “homogènes” et polynômes orthogonaux á deux variables. *Rend. Mat. (7)*, 11:673–689, 1991.

[2] B. Benouahmane and A. Cuyt. Multivariate orthogonal polynomials, homogeneous Padé approximants and Gaussian cubature. *Numer. Algorithms*, 24:1–15, 2000.

[3] B. Benouahmane, A. Cuyt, and I. Yaman. Near-minimal cubature formulae on the disk. *IMA J. Numer. Anal.*, 39(1):297–314, 2019.

[4] C. Brezinski. *Padé type approximation and general orthogonal polynomials.* ISNM 50, Birkhäuser Verlag, Basel, 1980.

[5] A. Cuyt. The epsilon-algorithm and multivariate Padé approximants. *Numer. Math.*, 40:39–46, 1982.

[6] A. Cuyt. Multivariate Padé approximants. *J. Math. Anal. Appl.*, 96:283–293, 1983.

[7] A. Cuyt. Multivariate Padé approximants revisited. *BIT*, 26:71–79, 1986.

[8] A. Cuyt. On the convergence of the multivariate “homogeneous” qd-algorithm. *BIT*, 34:535–545, 1994.

[9] A. Cuyt, B. Benouahmane, Hamsapriye, and I. Yaman. Symbolic-numeric Gaussian cubature rules. *Appl. Numer. Math.*, 61:929–945, 2011.

[10] A. Cuyt, B. Benouahmane, and B. Verdonk. Spherical orthogonal polynomials and symbolic-numeric Gaussian cubature formulas. In M. Bubak et al., editors, *LNCS 3037*, pages 557–560, Berlin, 2004. Springer-Verlag.

[11] A. Cuyt, F. Knaepkens, and W.-s. Lee. From exponential analysis to Padé approximation and Tensor decomposition, in one and more dimensions. In V.P. Gerdt et al., editors, *LNCS11077*, pages 116–130, 2018. Proceedings CASC 2018, Lille (France).

[12] A. Cuyt and D. Lubinsky. A de Montessus theorem for multivariate homogeneous Padé approximants. *Ann. Numer. Math.*, 4:217–228, 1997.

[13] Annie Cuyt and Wen-shin Lee. Sparse interpolation and rational approximation. volume 661 of *Contemporary Mathematics*, pages 229–242, Providence, RI, 2016. American Mathematical Society.

[14] Annie Cuyt and Wen-shin Lee. How to get high resolution results from sparse and coarsely sampled data. *Appl. Comput. Harmon. Anal.*, 2018. To appear.

[15] Annie Cuyt and Wen-shin Lee. Multivariate exponential analysis from the minimal number of samples. *Adv. Comput. Math.*, 44:987–1002, 2018.

[16] Annie Cuyt, Wen-shin Lee, and Xianglan Yang. On tensor decomposition, sparse interpolation and Padé approximation. *Jaén J. Approx.*, 8(1):33–58, 2016.

[17] Annie Cuyt, Irem Yaman, Ali Ibrahimoglu, and Brahim Benouahmane. Radial orthogonality and Lebesgue constants on the disk. *Numer. Algorithms*, 61:291–313, 2012.

[18] R. de Prony. Essai expérimental et analytique sur les lois de la dilatabilité des fluides élastiques et sur celles de la force expansive de la vapeur de l'eau et de la vapeur de l'alkool, à différentes températures. *J. Ec. Poly.*, 1:24–76, 1795.

[19] P. Henrici. *Applied and computational complex analysis I.* John Wiley & Sons, New York, 1974.

[20] F.B. Hildebrand. *Introduction to numerical analysis.* Mc Graw Hill, New York, 1956.

[21] Yingbo Hua and Tapan K. Sarkar. Matrix pencil method for estimating parameters of exponentially damped/undamped sinusoids in noise. *IEEE Trans. Acoust., Speech, Signal Process.*, 38:814–824, 1990.

[22] H. Nyquist. Certain topics in telegraph transmission theory. *Trans. Am. Inst. Electr. Eng.*, 47(2):617–644, April 1928.

[23] L. Weiss and R.N. McDonough. Prony's method, Z-transforms, and Padé approximation. *SIAM Rev.*, 5:145–149, 1963.