Bi-orthogonality in rational approximation

A. ISERLES
D.A.M.T.P., University of Cambridge, Cambridge CB3 9EW, United Kingdom

E.B. SAFF *
Department of Mathematics, University of South Florida, Tampa, FL 33620, U.S.A.

Abstract: The well-known connection between Padé approximants to Stieltjes functions and orthogonal polynomials is crucial in locating zeros and poles and in convergence theorems. In the present paper we extend similar types of analysis to more elaborate forms of approximation. It transpires that the link with orthogonal polynomials remains valid with regard to rational interpolants, whereas simultaneous Padé and Levin–Sidi approximants yield themselves analysis with bi-orthogonal polynomials.

1. Introduction

The theme of the present paper is a generalisation of the familiar Padé theory of Stieltjes functions to more elaborate forms of approximation—rational interpolation, simultaneous Padé approximation with German polynomials and Levin–Sidi approximation.

Let f be a Stieltjes function that is analytic at the origin,

$$f(z) = \int_0^\infty \frac{d\psi(\tau)}{1 + \tau z}, \quad z \in \mathbb{C} \setminus (-\infty, 0),$$

where ψ is a distribution. Moreover, let P_m/Q_n be the m/n Padé approximant to f, $m \geqslant n - 1$, normalised so that $Q_n(0) = 1$. It is well known [2] that the inverted polynomial

$$\tilde{Q}_n(x) := x^n Q_n(-1/x),$$

is the degree n monic orthogonal polynomial with respect to the weight function $\tau^{m-n+1} \, d\psi(\tau)$. Consequently, all the poles of Q_n lie in $(-\infty, 0)$ and are simple.

The case $m = n - 1$ is of particular interest, since now the inverted polynomial

$$\tilde{P}_{n-1}(x) := x^{n-1} P_{n-1}(-1/x)$$

is the $(n-1)$st numerator polynomial [7] with respect to the orthogonal polynomial system $\{\tilde{Q}_k\}_{k=0}^\infty$. Thus, the zeros of $\tilde{Q}_n$ and $\tilde{P}_{n-1}$ interlace, implying that the residua at the poles of the approximant are positive. Hence P_{n-1}/Q_n remains uniformly bounded in any compact subset of $\mathbb{C} \setminus (-\infty, 0)$. Convergence now follows easily by the Vitali theorem.

It is of interest to extend this analysis to other forms of approximation. It transpires that the connection between an underlying approximant and orthogonality is helpful in one application. However, for our remaining two forms of approximation we need to use, instead, *bi-orthogonality*. The theory of bi-orthogonal polynomials has been systematically introduced in [9]. Here it suffices to mention that, given a function $\phi(x, \mu)$, $x \in (a, b)$, $\mu \in \Omega \subseteq \mathbb{R}$, which is a distribution in x, and distinct points $\mu_1, \ldots, \mu_m \in \Omega$, the monic degree m polynomial p_m is *bi-orthogonal* if

$$\int_a^b p_m(x; \mu_1, \ldots, \mu_m) \, d\phi(x, \mu_l) = 0, \quad 1 \leqslant l \leqslant m.$$

It is necessary and sufficient for a unique p_m to exist for all $m \geqslant 1$ and distinct $\mu_1, \ldots, \mu_m \in \Omega$ that the matrices

$$\left(\int_a^b x^k \, d\phi(x, \mu_l) \right)_{l=1,\ldots,m}^{k=0,\ldots,m-1}$$

are non-singular for all $m \geqslant 1$ and distinct $\mu_1, \ldots, \mu_m$. In that case we say that ϕ is *regular*. Finally, let $d\phi(x, \mu) = w(x, \mu) \, d\alpha(x)$, where α is a distribution which is independent of μ and w is in $C^1(\Omega)$ as a function of μ. We say that ϕ possesses the *interpolation property* if for all $m \geqslant 1$, distinct $x_1, \ldots, x_m \in (a, b)$ and distinct $\mu_1, \ldots, \mu_m \in \Omega$, the matrix $(w(x_k, \mu_l))_{k,l=1,\ldots,m}$ is non-singular. If ϕ possesses the interpolation property, then each p_m has m distinct zeros within the support (a, b).

Elsewhere in this issue [10], the theory of bi-orthogonal polynomials has been applied to the problem of mapping zeros of polynomials under various transformations, whereas in [8] it has been used in investigating numerical methods for ordinary differential equations.

2. Rational interpolation

Let m and n be two non-negative integers, $m \geqslant n - 1$, and $\omega_0, \omega_1, \ldots, \omega_{n+m}$ be complex numbers which are either real and positive or appear in conjugate pairs. We consider the interpolation of a Stieltjes function f by a rational function P_m/Q_n of type $m - 1/n$, $Q_n(0) = 1$, at the points ω_k, $0 \leqslant k \leqslant n + m$.

Let, again, $\tilde{Q}_n$ be the 'inverted' denominator Q_n. It has been proved in [4] that $\tilde{Q}_n$ is the monic orthogonal polynomial with respect to the distribution

$$\tau^{m-n+1} \, d\psi(\tau) \bigg/ \prod_{k=0}^{n+m} (1 + \omega_k \tau).$$

Note that, as $\omega_k \to 0$, $0 \leqslant k \leqslant n + m$, we recover, as expected, the Padé approximant.

Consequently, all the poles of the interpolant are negative and distinct. Alas, the zeros can no longer be analysed by the familiar Padé techniques—e.g., for $m = n - 1$ the 'inverted' P_{n-1} is no longer the numerator polynomial (in the sense of [7]) of $\tilde{Q}_n$. Instead, we can use a dynamic proof to demonstrate that, for $n - 1 \leqslant m \leqslant n$, poles and zeros interlace. Herewith the proof for $m = n$: let P_n^*/Q_n^* be the n/n Padé approximant of f, with zeros $\{\zeta_k^*\}_{k=1}^n$ and poles $\{\eta_k^*\}_{k=1}^n$. Since interlace is true in the Padé case, we may assume without loss of generality that

$$\zeta_1^* < \eta_1^* < \zeta_2^* < \cdots < \zeta_n^* < \eta_n^* < 0. \tag{1}$$

We set $\sigma_k(\tau) := \tau\omega_k$, $0 \leqslant k \leqslant 2n$, $\tau \in [0, 1]$ and consider the n/n functions $\hat{P}_n(\cdot, \tau)/\hat{Q}_n(\cdot, \tau)$ that interpolate f at $\sigma_k(\tau)$, $0 \leqslant k \leqslant 2n$. Clearly,

$$\hat{P}_n(z, 0)/\hat{Q}_n(z, 0) \equiv P_n^*(z)/Q_n^*(z); \qquad \hat{P}_n(z, 1)/\hat{Q}_n(z, 1) \equiv P_n(z)/Q_n(z).$$

Moreover, it follows from [4] that no m/n function can interpolate f at more than $m + n + 1$ positive or complex conjugate points. Hence $\hat{P}(\cdot, \tau)/\hat{Q}(\cdot, \tau)$ exists and is unique for all $0 \leqslant \tau \leqslant 1$.

We denote the zeros of $\hat{P}(\cdot, \tau)$ and $\hat{Q}(\cdot, \tau)$ by $\{\hat{\zeta}_k(\tau)\}_{k=1}^n$ and $\{\hat{\eta}_k(\tau)\}_{k=1}^n$ respectively. Thus, it follows from (1) that for $0 \leqslant \tau \ll 1$, without loss of generality

$$\hat{\zeta}_1(\tau) < \hat{\eta}_1(\tau) < \hat{\zeta}_2(\tau) < \cdots \hat{\zeta}_n(\tau) < \hat{\eta}_n(\tau) < 0. \qquad (2)$$

Furthermore, as a consequence of our analysis,

$$\hat{\eta}_1(\tau) < \hat{\eta}_2(\tau) < \cdots \hat{\eta}_n(\tau) < 0$$

for all $0 \leqslant \tau$.

Let us suppose that there exists $\tau^* \in (0, 1]$ for which (2) is no longer true. Then there must exist $\tau_1 \in (0, \tau^*]$ for which either $\lim_{\tau \to \tau_1} \hat{\zeta}_1(\tau) = -\infty$ or $\hat{\zeta}_k(\tau_1) = \hat{\eta}_k(\tau_1)$ for some $k \in \{1, \ldots, n\}$ or $\hat{\zeta}_{k+1}(\tau_1) = \hat{\eta}_k(\tau_1)$ for some $k \in \{1, \ldots, n-1\}$. Each of these implies degeneracy of the rational interpolant, i.e. the existence of $0 \leqslant n_1, n_2, n_1 + n_2 \leqslant 2n - 1$, such that $\deg \hat{P}_n(\cdot, \tau_1) = n_1$, $\deg \hat{Q}_n(\cdot, \tau_1) = n_2$. Since this is impossible, it transpires that (2) is valid for all $\tau \in [0, 1]$, hence the desired interlace.

Since interlace implies uniform boundedness in compact subsets of $\mathbb{C} \setminus (-\infty, 0]$, it is now elementary to use a normal families argument (Vitali's theorem) to prove convergence for various configurations of interpolation points—cf. [4] for the case of best L_∞ approximants.

Rational interpolation of Stieltjes functions has already been explored by Barnsley [3], who used continued S-fractions to prove convergence of multipoint Padé approximants. Alas, since S-fractions, unlike J-fractions, are not linked to orthogonal polynomials, the present theory provides, in our view, a more natural extension of the classical Padé theory.

3. German polynomials

Let f_1 and f_2 be two Stieltjes functions, that correspond to distributions ψ_1 and ψ_2 respectively. In the present section we consider simultaneous Padé approximation (German polynomials) to f_1 and f_2.

For simplicity we assume that each f_k is approximated to order $n + m$ by $P_{k,m}/Q_{2n}$, where $\deg P_{k,m} = m$, $\deg Q_{2n} = 2n$ and $m \geqslant 2n - 1$.

In the case $m = 2n - 1$, Angelesco [1] has already proved that convergence occurs if the supports of ψ_1 and ψ_2 are two mutually exclusive intervals. More interesting analysis is due to Nikišin [12]. It is based, essentially, on bi-orthogonal polynomials and herewith we re-formulate it in this formalism:

Let

$$f_k(z) = \sum_{l=0}^\infty f_{k,l} x^l, \quad f_{k,l} = (-1)^l \int_0^\infty \tau^l \, d\psi_k(\tau), \qquad l \geqslant 0, \quad k = 1, 2,$$

and

$$Q_{2n}(z) = \sum_{k=0}^{2n} q_k x^k, \quad q_0 = 1.$$

The order conditions lead to

$$\sum_{l=0}^{\min\{2n, j\}} f_{k,j-l} q_l = 0, \qquad m + 1 \leqslant j \leqslant n + m, \quad k = 1, 2 \tag{3}$$

(see [6]). Let $d\hat{\psi}_k(\tau) := \tau^{m+1-2n} \, d\psi_k(\tau)$, $k = 1, 2$. Then it follows easily that (3) is equivalent to

$$\int_0^\infty \tau^r \tilde{Q}_{2n}(\tau) \, d\hat{\psi}_k(\tau) = 0, \qquad k = 1, 2, \quad 0 \leqslant r \leqslant n - 1, \tag{4}$$

where $\tilde{Q}_{2n}$ is the 'inverted' denominator,

$$\tilde{Q}_{2n}(x) := x^{2n} Q_{2n}(-1/x).$$

Let $\Omega = \{1, 2, 3, \ldots\}$. We define

$$d\phi(x, r) := x^{r-1} \, d\hat{\psi}_1(x), \quad 1 \leqslant r \leqslant n,$$

$$d\phi(x, r) := x^{r-n-1} \, d\hat{\psi}_2(x), \quad n + 1 \leqslant r \leqslant 2n.$$

It now follows at once from (4) and $q_0 = 1$ that $\tilde{Q}_{2n}(x) \equiv p_{2n}(x; 1, 2, \ldots, 2n)$, the corresponding bi-orthogonal polynomial. Consequently (and with obvious modifications for a discrete Ω), the satisfaction of interpolation property by ϕ implies that Q_{2n} has $2n$ distinct zeros in $(-\infty, 0)$.

Nikišin goes to prove, by quadrature techniques, that, in the case $m = 2n - 1$, for certain ψ_1 and ψ_2, each $P_{k,m}$ has m negative zeros that interlace with the zeros of Q_{2n}. There is a short step from this to convergence in compact subsets of $\mathbb{C} \setminus (-\infty, 0]$.

An important instance when the interpolation property holds for $d\phi(x, r)$ is $d\psi_2(x) = x^\alpha \, d\psi_1(x)$, α non-integer [9]. Moreover, in the absence of interpolation property things may well go wrong: consider

$$f_1(z) = \frac{\log(1 + z)}{z}, \qquad f_2(z) = \tfrac{13}{20} + \tfrac{3}{20} \frac{1}{1 + z} + \tfrac{2}{5} \frac{1}{2 + z}.$$

Both are, clearly, Stieltjes functions. Let $n = m = 1$. Since $\hat{\psi}_1(x) = x$ for $x \in [0, 1]$, whereas $\hat{\psi}_2(x)$ is a step function with jumps of $\tfrac{13}{20}$, $\tfrac{1}{5}$, and $\tfrac{3}{20}$ at 0, $\tfrac{1}{2}$ and 1 respectively; the interpolation property is invalid. Moreover, the simultaneous Padé approximants are

$$\frac{1 + \tfrac{1}{30} z}{1 + \tfrac{8}{15} z - \tfrac{1}{15} z^2} \quad \text{and} \quad \frac{1 + \tfrac{17}{60} z}{1 + \tfrac{8}{15} z - \tfrac{1}{15} z^2}$$

respectively, with poles at $4 \pm \sqrt{31}$ —one outside $(-\infty, 0)$.

4. Levin–Sidi approximants

Levin [11] introduced a powerful algorithm to accelerate convergence of sequences. If that algorithm is used to sum up power series it generates a sequence of rational approximants, in parallel with the familiar ϵ-algorithm that generates Padé approximants. The algorithm has been

analysed by Sidi [14], who also showed that it produced very efficient quadrature schemes which compete well with Gaussian quadrature.

Let

$$g(z) = \sum_{k=0}^{\infty} c_k z^k = \int_0^{\infty} \frac{d\psi(\tau)}{1 - \tau z}$$

be a Stieltjes function. Then, for $m \geqslant n - 1$, the Levin–Sidi approximant can be given explicitly by P_m/Q_n, where

$$P_m(z) = \sum_{l=1}^{m+1} \left(\sum_{j=\max\{0, l-m+n-1\}}^{n} (-1)^j \binom{n}{j} (m - n + 2 + j)^{n-1} \frac{c_{m-n+j+1-l}}{c_{m-n+j+1}} \right) z^{m+1-l},$$

$$\tag{5}$$

$$Q_n(z) = \sum_{l=0}^{n} (-1)^l \binom{n}{l} (m - n + 2 + l)^{n-1} \frac{z^{n-l}}{c_{m-n+l+1}}.$$

This can be obtained, after obvious modifications, from Sidi [14]. It is a Padé-type approximant [5] of order $m + 1$.

Let $\Omega = (a, b) = (0, \infty)$ and

$$d\phi(x, \mu) = (x/\mu)^{m-n+1} d\psi(x/\mu), \quad \mu \in \Omega. \tag{6}$$

It is easy to verify that for every distinct $\mu_1, \ldots, \mu_n \in \Omega$ the nth bi-orthogonal polynomial reads explicitly

$$p_n(x; \mu_1, \ldots, \mu_n) = c_{m+1} \sum_{k=0}^{n} \frac{q_k}{c_{m-n+1+k}} x^k, \tag{7}$$

where $\sum_{k=0}^{n} q_k x^k = \prod_{k=1}^{n} (x - \mu_k)$ [9]. Sidi proves that the polynomial

$$-r_{q,n}(z) := \sum_{k=0}^{n} (-1)^k \binom{n}{k} (q + k)^{n-1} z^k, \quad q > 0,$$

has n distinct zeros in $(0, 1]$. Consequently, if these zeros are $\mu_1, \ldots, \mu_n$ then it follows from (5) and (7) that

$$Q_n(z) = (-1)^n \frac{(m + 2)^{n-1}}{c_{m+1}} z^n p_n(1/z; \mu_1, \ldots, \mu_n).$$

Therefore, it follows that:

Theorem 1. *If ϕ, as given in (6), possesses the interpolation property then the Levin–Sidi approximation has n distinct positive poles.*

Several interesting choices of ψ give rise, via (6), to interpolation property. The following lemmata are in the spirit of [10] and follow readily by identical reasoning.

Lemma 2. *Let $q(x) = \sum_{k=0}^{m} q_k x^k$ be a monic polynomial with m distinct positive zeros and let α, β be any positive numbers. Then the polynomial*

$$\sum_{k=0}^{m} \frac{q_k}{\Gamma\left(\dfrac{k + \beta}{\alpha}\right)} x^k$$

possesses m distinct positive zeros.

Proof. By considering $d\psi(x) = x^{\beta-1} e^{-x^\alpha} dx$, $m = n - 1$, (6) and (7). $\square$

The lemma generalises a result by Pólya and Szegö [13], which pertains to $\alpha = 1$, $\beta = 2$.

Lemma 3. *Let $q(x) = \sum_{k=0}^{m} q_k x^k$ be a monic polynomial with m distinct positive zeros and let α be a number in $(0, 1)$. Then the polynomial*

$$\tilde{p}_m(x) = \sum_{k=0}^{m} \alpha^{k^2} q_k x^k$$

possesses m distinct positive zeros.

Proof. We consider the distribution of the Stieltjes–Wigert polynomials.

$$d\psi(x) = \frac{\sigma}{\sqrt{\pi}} e^{-\sigma^2(\log x)^2} dx,$$

where $\sigma := 1/(2\sqrt{-\log \alpha}) > 0$. ϕ is, again, given by (6). It is known that

$$c_k = \exp\left(\left(\frac{k+1}{2\sigma}\right)^2\right) = \alpha^{-(k+1)^2}, \quad k \geqslant 0$$

(see [7]), hence, by (7),

$$p_m(x; \mu_1, \ldots, \mu_m) = \alpha^{-(m+1)^2} \sum_{k=0}^{m} \alpha^{(k+1)^2} q_k x^k.$$

where $\mu_1, \ldots, \mu_m > 0$ are the zeros of q. Thus

$$\tilde{p}_m(x) = \alpha^{m(m+2)} p_m(\alpha^{-2}x; \mu_1, \ldots, \mu_m).$$

Finally, we demonstrate the interpolation property: set

$$w(x, \mu) := \frac{1}{\mu} e^{-\sigma^2(\log x/\mu)^2}, \qquad \alpha(x) := \frac{\sigma}{\sqrt{\pi}} x.$$

Since

$$e^{-\sigma^2(\log(x/\mu))^2} = e^{-\sigma^2((\log x)^2 + (\log\mu)^2)} x^{2\sigma^2\log\mu},$$

it follows that

$$\det(w(x_k, \mu_l))_{k,l=1,\ldots,m}$$

$$= \frac{\sigma^m}{\pi^{1/2m}\mu_1, \mu_2, \ldots, \mu_m} \exp\left(-\sigma^2 \sum_{k=1}^{m} \left((\log x_k)^2 + (\log \mu_k)^2\right)\right) \det\left(x_k^{2\sigma^2\log\mu_l}\right)_{k,l=1,\ldots,m}.$$

Interpolation property follows since $\det(x_k^{y_l})_{k,l=1,\ldots,m} \neq 0$ for every distinct $x_1, \ldots, x_m$ and distinct $y_1, \ldots, y_m$ [9]. This concludes the proof. $\square$

The significance of the last two lemmata to Levin–Sidi approximation is clear although, of course, they are of interest on their own merit.

Unfortunately, the interpolation property fails in numerous interesting instances. In particular, it is invalid if the support of ψ is finite. Sometimes one may then resort to different techniques.

Lemma 4. *Given* $n \leqslant m \leqslant 2n - 2$, *let us suppose that the polynomial*

$$s_n(x) := \sum_{k=0}^{n} (-1)^k \binom{n}{k} \frac{x^k}{c_{m-n+k+1}}$$

has n distinct positive zeros. Then so has the denominator Q_n.

Proof. Let $\bar{Q}_n(z) := z^n Q_n(1/z)$ *and* $\nu := z(d/dz)$. It is straightforward that

$$\bar{Q}_n(z) = z^{-m+n-2} \nu^{n-1} \left\{ z^{m-n+2} s_n(z) \right\}.$$

Since $m \leqslant 2n - 2$, no zero of $x^{m-n+2} s_n(x)$ has multiplicity greater than n. The lemma now follows readily by repeated application of the Rolle theorem. □

Frequently s_n can be identified with an orthogonal polynomial with positive support: $c_k = \Gamma(m - n + \alpha + 1 + k)$, $k \geqslant 0$, leads to the Laguerre polynomial $L_n^{(m-n+\alpha)}$ although this choice of c_k's merely specialises Lemma 2. More interestingly, $c_k = (\alpha)_k / (\beta)_k$, $k \geqslant 0$, for $\alpha > n - m - 1$, $\beta > \alpha + n - 1$, yields the Jacobi polynomial $P_n^{(\alpha+m-n,\beta-\alpha-n)}$, shifted to $(0, 1)$. In both cases all zeros are positive and distinct.

Another approach is to consider

$$\bar{Q}_{n,r}(z) := \sum_{k=0}^{n} (-1)^k \binom{n}{k} (r + 1 + k)^{n-1} \frac{x^k}{c_{r+k}}.$$

Of course, $r = m - n + 1$ yields $\bar{Q}_n$. The polynomials $\bar{Q}_{n,r}$ obey the recurrence relation

$$\bar{Q}_{n,r}(z) := (r + 1)\bar{Q}_{n-1,r}(z) - (n + r + 1)z\bar{Q}_{n-1,r+1}(z), \qquad n \geqslant 1, \quad r \geqslant 0.$$

Hence, since $r = m - n + 1$ gives $n + r + 1 = m + 2 > 0$, it is easy to see that if the zeros of $\bar{Q}_{n-1,r}$ and $\bar{Q}_{n-1,r+1}$ are all positive and distinct and if they interlace then the zeros of $\bar{Q}_{n,r}$ are, likewise, positive and distinct. Unfortunately, no general conditions are presently available for the interlace of zeros of $\bar{Q}_{n-1,r}$ and $\bar{Q}_{n-1,r+1}$.

Acknowledgements

References

[1] A. Angelesco, Sur deux extensions des fractions continues algébriques, *C.R. Acad. Sci. Paris* 168 (1919) 262–265.

[2] G.A. Baker, *Essentials of Padé Approximants* (Academic Press, New York, 1975).

[3] M. Barnsley, The bounding properties of the multipoint Padé approximant to a series of Stieltjes on the real line, *J. Math. Phys.* 14 (1973) 299–313.

[4] H.-P. Blatt, A. Iserles and E.B. Saff, Remarks on the behavior of zeros of best approximating polynomials and rational functions, in: M.G. Cox and J.C. Mason, Ed., *Approximation of Functions and Data*, (Oxford University Press, Oxford, 1987).

[5] C. Brezinski, *Padé-Type Approximation and General Orthogonal Polynomials* (Birkhäuser, Basel, 1980).

[6] M.G. de Bruin, Some explicit formulae in simultaneous Padé approximation, *Lin. Algebra Appl.* 63 (1984) 271–281.

[7] T.S. Chichara, *An Introduction to Orthogonal Polynomials* (Gordon and Breach, New York, 1978).

[8] A. Iserles and S.P. Nørsett, Two-step methods and bi-orthogonality, Math. Comput., to appear.

[9] A. Iserles and S.P. Nørsett, On the theory of bi-orthogonal polynomials, Trans. AMS, to appear.

[10] A. Iserles and S.P. Nørsett, Bi-orthogonality and zeros of transformed polynomials, *J. Comput. Appl. Math.* 19 (1987) (this issue) 39–45.

[11] D. Levin, Development of non-linear transformations for improving convergence of sequences, *Intern. J. Comp. Math.* B3 (1973) 371–388.

[12] E.M. Nikišin, On simultaneous Padé approximants, *Math. USSR Sb.* 41 (1982) 409–425.

[13] G. Pólya and G. Szegö, *Aufgaben und Lehrsätze aus der Analysis, Vol. II* (Berlin, 1925).

[14] A. Sidi, Numerical quadrature and nonlinear sequence transformations; unified rules for efficient computation of integrals with algebraic and logarithmic endpoint singularities, *Math. Comput.* 35 (1980) 851–874.

Con

fixe(

Paul SA

UFR IEE

59655 Vill

Abstract:

For $x_0 >$

its conver

versions (

phenomei

present. I

series.

Keyword:

1. Intr

The
logarit,

Thoug
accele
Son
conve

where
conve