How accurate is Gauss quadrature with indefinite or complex weight function?

Martin H. Gutknecht

Seminar for Applied Mathematics
ETH Zurich

The Third International Workshop on Matrix Computations
Online, organized by Lanzhou University, PR China
April 16, 2022
Version of April 16, 2022

Can there be anything new on Gauss quadrature?

Recently, a group around Zdenek Strakoš made an effort to generalize **Gauss quadrature**. Included were in particular:

- Jörg Liesen
- Stefano Pozza
- Miroslav Pranić,
- Zdenek Strakoš

They were particularly interested in connections to related topics: the conjugate gradient method, the three-term Lanczos algorithm, orthogonal polynomials, the partial realisation problem of systems and control,...

Can there be anything new on Gauss quadrature? (cont'd)

Selected publications:

- J. Liesen, Z. Strakoš, Krylov Subspace Methods: Principles and Analysis, Oxford University Press, Oxford, 2013.
- S. Pozza, M.S. Pranić, Z. Strakoš, Gauss quadrature for quasi-definite linear functionals, IMA J. Numerical Analysis **37**, 1468–1495 (2017).
- S. Pozza, M.S. Pranić, Z. Strakoš, The Lanczos algorithm and complex Gauss quadrature, ETNA **50**, 1–19 (2018).
- S. Pozza, M.S. Pranić, The Gauss quadrature for general linear functionals, Lanczos algorithm, and minimal partial realization, arXiv:1903.11395 [math.NA] v1 (2019), v2 (2020); Numer. Algor. **88**, 647–678 (2021).

Gragg's circle of ideas

It is well known that there is a close relationship between the theories of the following 12 problems

- (formal) orthogonal polynomials (OPs and FOPs)
- Padé approximation
- continued Jacobi fractions (J-fractions)
- (formal) continued Stieltjes fractions (S-fractions)
- the qd algorithm
- the LR algorithm
- the 3-term Lanczos algorithm
- Lanczos' 2-term (bi)conjugate gradient algorithm
- fast Hankel solvers
- the partial realisation problem of systems and control
- the (modfied) Chebyshev algorithm
- Gauss quadrature

W. B. Gragg called that the **circle of ideas**.

Gragg's circle of ideas (cont'd)

Numerous mathematicians, in particular numerical analysts contributed algorithms for specific problems and specific cases of this list.

For some of the results on any of the 12 problems there are equivalent results on some or even all of the other problems.

But there are also aspects and results that are typical for a particular problem.

Example: classical Gauss quadrature

Given a positive weight function $\mu(x)$ defined on an interval $[a, b]$, and any positive integer n, one has to find n nodes x_j ($j = 1, ..., n$) and n weights ω_j such that the integral

$$\int_a^b f(x)\mu(x)dx$$

is approximated as good as possible by a **Gauss quadrature formula**

$$\sum_{j=1}^{n} f(x_j)\,\omega_j,$$

where “as good as possible” means “with zero error for all real polynomials of degree at most N”, where N is maximum.

Gauss quadrature achieves $N = 2n - 1$.

Example: classical Gauss quadrature (cont'd)

More generally one could handle Riemann-Stieltjes integrals with a nonnegative measure $d\mu(x)$ defined on $\mathbb{R}$ and satisfying suitable boundedness conditions:

$$\int_{\mathbb{R}} f(x) d\mu(x).$$

But I stick to the simpler formulation.

It turns out that the nodes x_j must be chosen as the zeros of the **orthogonal polynomial** P_n of degree n for the interval $[a, b]$ and the positive weight function $\mu(x)$.

If we let $\mathcal{P}_n$ be the set of (real or complex, respectively) polynomials of degree at most n, then P_n is characterized by

$$\int_a^b p(x) P_n(x) \mu(x) dx = 0 \qquad (\forall p \in \mathcal{P}_{n-1}).$$

The orthogonal polynomials P_n play also a role in the other problems.

An aspect typical for Gauss quadrature: need weights ω_j.

The orthogonal polynomials P_n satisfy a three-term relation

$$\gamma_i P_i(x) = (x - \alpha_{i-1})P_{i-1}(x) - \beta_{i-1}P_{i-2}(x)\,,$$

which can be summed up as

$$x\mathbf{p}_n(x) = \mathbf{T}_n\mathbf{p}_n(x) + \beta_n P_n(x)\mathbf{e}_n$$

with

$$\mathbf{T}_n :\equiv \begin{bmatrix} \alpha_0 & \beta_1 & & & \\ \gamma_0 & \alpha_1 & \beta_2 & & \\ & \gamma_1 & \alpha_2 & \ddots & \\ & & \ddots & \ddots & \beta_{n-1} \\ & & & \gamma_{n-2} & \alpha_{n-1} \end{bmatrix}, \quad \mathbf{p}_n :\equiv \begin{bmatrix} P_0 \\ P_1 \\ \vdots \\ P_{n-1} \end{bmatrix} \quad \mathbf{e}_n :\equiv \begin{bmatrix} 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix}.$$

If we choose **orthonormal polynomials** P_n such that

$$\int_a^b P_m(x)P_n(x)\mu(x)dx = \delta_{m,n}\,,$$

then $\mathbf{T}_n$ is a symmetric positive definite (spd) **Jacobi matrix**.

If $\mathbf{q}_1, \ldots, \mathbf{q}_n$ are the normalized eigenvectors of $\mathbf{T}_n$, and $q_{1,1}, \ldots, q_{1,n}$ are their first components, then the weights ω_j of the Gauss quadrature rule are the squares of these first components:

$$\omega_j = q_{1,j}^2 \qquad (j = 1, \ldots, n).$$

Moreover,

$$\sum_{j=1}^{n} f(x_j)\,\omega_j = \phi_0\,\mathbf{e}_1^T\,f(\mathbf{T}_n)\,\mathbf{e}_1\,, \qquad \phi_i = \phi_0\,\mathbf{e}_1^T\,(\mathbf{T}_n)^i\,\mathbf{e}_1\,.$$

Example: classical Gauss quadrature (cont'd)

There may be a finite or an infinite sequence of orthogonal polynomials — depending on the support of the weight function.

Correspondingly, we get a finite or an infinite tridiagonal matrix with the recurrence coefficients.

Gauss quadrature is mainly of importance in case of a positive real (or non-negative real) weight function or measure, that is, in the real symmetric positive definite (spd) case.

What will happen if the weight function is real indefinite or even complex?

Before 2010 there have been some publications, but not very many, and some of them only short ones.

- Gragg (1974),
- Draux (1983),
- Freund and Hochbruck (1993),
- Stroud and Secrest (1996),
- Saylor and Smolarski (2001),
- Milovanović and Cvetković (2003),
- Gautschi (2004).

Gauss quadrature with real indefinite or complex measure

However, most people seem to have missed

André Draux (1983), *Polynômes Orthogonaux Formels — Application*, LNM vol. 974, Springer-Verlag, Berlin.

This is André Draux's PhD thesis advised by Claude Brezinski.
It has 625 pages, but no equation numbers!

It covers everything on formal orthogonal polynomials (FOPs) for real linear functionals and their application.

Chapter 5 (35 pages) is on Gauss quadrature.

Example: linear functionals

We consider more generally a real or complex **linear functional**

$$\Phi : \mathcal{P} \to \mathbb{E}, \qquad \text{where } \mathbb{E} :\equiv \mathbb{R} \text{ or } \mathbb{C}$$

defined on the space $\mathcal{P}$ of real or complex polynomials, respectively, and determined by the values it takes on the monomials, the **moments** (or **Markov parameters**) ϕ_k:

$$\Phi(z^k) := \phi_k \qquad (k \in \mathbb{N}).$$

Now an nth **(true) formal orthogonal polynomial (FOP)** $P_n \in \mathcal{P}_n$ is required to have exact degree n and to satisfy (Draux(1983))

$$\Phi(pP_n) = 0 \qquad (\forall p \in \mathcal{P}_{n-1}).$$

Note that this is not an inner product if $\mathbb{E} = \mathbb{C}$.

Example: linear functionals

The formula

$$\Phi(pP_n) = 0 \qquad (\forall p \in \mathcal{P}_{n-1}).$$

yields a homogeneous linear system of n equations in $n+1$ unknowns (the coefficients of P_n):

(1.6)

$$\begin{bmatrix} \phi_0 & \phi_1 & \cdots & \phi_{n-1} & \phi_n \\ \phi_1 & \phi_2 & \cdots & \phi_n & \phi_{n+1} \\ \vdots & \vdots & \iddots & \vdots & \vdots \\ \phi_{n-1} & \phi_n & \cdots & \phi_{2n-2} & \phi_{2n-1} \end{bmatrix} \begin{bmatrix} \pi_{0,n} \\ \pi_{1,n} \\ \vdots \\ \pi_{n,n} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

If P_n is normalized to be monic, the last column of the matrix is moved to the right-hand side and the matrix becomes a square *Hankel matrix*, the **moment matrix $\mathbf{M}_n$**.

Note that $\mathbf{M}_{n-1}$ is the leading principal submatrix of $\mathbf{M}_n$.

Example: linear functionals

The following well-known result follows immediately:

Theorem

A FOP P_n exists if and only if (1.6) has a solution with $\pi_{n,n} = 1$. It is regular (i.e., unique) if and only if the $n \times n$ moment matrix

$$\mathbf{M}_n :\equiv \begin{bmatrix} \phi_0 & \phi_1 & \cdots & \phi_{n-1} \\ \phi_1 & \phi_2 & \cdots & \phi_n \\ \vdots & \vdots & & \vdots \\ \phi_{n-1} & \phi_n & \cdots & \phi_{2n-2} \end{bmatrix} \tag{1.7}$$

is nonsingular, i.e.,

$$\Delta_n :\equiv \det \mathbf{M}_n \neq 0 \,.$$

Example: linear functionals

In general, a (true) FOP need not exist for every n.

If it exists, we will assume that P_n (with $n \geq 1$) is **normalized** by either being monic of degree n or by being orthonormal in the sense that $\Phi(P_n^2) = 1$.

P_n is called a **regular FOP** if it is uniquely determined; otherwise it is a **singular FOP** (Draux(1983)).

Example: linear functionals (cont'd)

If there exists a sequence of (true) orthonormal FOPs $P_0, P_1, \ldots, P_\nu$, such that $\Phi(P_m, P_n) = \delta_{m,n}$ for $m = 0, 1, \ldots, \nu$, $m = 0, 1, \ldots, \nu$, then Φ is said to be **quasi-definite** on $\mathcal{P}_\nu$.

There is a corresponding tridiagonal $\nu \times \nu$ matrix with the coefficients of the three-term recurrences.

This is the situation that is assumed in all of the first three publications of Strakoš *et al.* that I mentioned.

We can extend the situation further by replacing the sequence $\{\phi_k\}_{k=0}^{\infty}$ by a doubly infinite sequence $\{\phi_k\}_{k=-\infty}^{\infty}$.

Then a linear functional $\Phi_l : \mathcal{P} \to \mathbb{C}$ can be defined for each $l \in \mathbb{Z}$ by

$$\Phi_l(z^k) := \phi_{k+l} \qquad (k \in \mathbb{N}). \tag{1.1}$$

An nth true FOP with respect to Φ_l is denoted by $P_{l;\,n}$. By definition, $P_{l;\,n}$ is monic of degree n and satisfies

$$\Phi_l(pP_{l;\,n}) = 0 \qquad (\forall p \in \mathcal{P}_{n-1}). \tag{1.2}$$

Example: Lanczos algorithms

In CG, BiCG and in the suitably normalized 3-term Lanczos algorithm the residual polynomials are FOPs normalized by $P_n(0) = 1$.

Except for CG there may be breakdowns, most of which can be cured by **look-ahead**.

For the Lanczos algorithm, to get into the situation of a quasi-definite Φ on $\mathcal{P}_\nu$ means to apply the algorithm just till the first breakdown and stop. No look-ahead.

Example: Padé approximation

Given a **formal power series**

$$f(z) = \sum_{k=0}^{\infty} \phi_k z^k \,,$$

we want to match as many initial terms as possible by those of the Taylor series of a rational function $r_{m,n}$ of a given **type** (m,n), i.e. given **numerator degree** m and given **denominator degree** n.

We switch the notation, because we not only aim at formal orthogonal polynomials (FOPs), but at rational functions

$$r_{m,n}(z) :\equiv \frac{p_{m,n}(z)}{q_{m,n}(z)} \in \mathcal{R}_{m,n}$$

such that when $r_{m,n}(z)$ is expanded at 0 into a Taylor series, then

$$f(z) - r_{m,n}(z) = f(z) - \frac{p_{m,n}(z)}{q_{m,n}(z)} = \mathcal{O}(z^{\kappa}) \tag{1}$$

with κ as large as possible.

$r_{m,n}$ has $m+n+1$ free coefficients plus one that is normalized; so we can expect to match $m+n+1$ coefficients of the series.

To simplify matters we linearize the problem:

Given the formal power series f, find for every pair $(m,n) \in \mathbb{N} \times \mathbb{N}$ all **(m,n)-Padé forms** $(p_{m,n}, q_{m,n}) :\equiv (p,q) \in \mathcal{P}_m \times \mathcal{P}_n$ for which

$$f(z)q(z) - p(z) = \mathcal{O}(z^{m+n+1})\,. \tag{2}$$

The quotients $r_{m,n}(z) :\equiv \frac{p_{m,n}(z)}{q_{m,n}(z)} \in \mathcal{R}_{m,n}$ are the **(m,n)-Padé approximants** of f.

After cancelling common factors they are unique, while the (m,n)-Padé forms are never unique.

The denominators $q_{m,n}$ are FOPs.

More detail are on my old handwritten slides. [Link to pdf]

A special feature is the **Padé table** and its block structure.

Typical for the problem of Padé approximation is the need/use of numerators.

Gauss quadrature: the general case

I was a referee for a first version (submitted to SIMAX) of the the second paper by Pozza, Pranić, and Strakoš entitled *The Lanczos algorithm and complex Gauss quadrature*.

Though not at all an expert on Gauss quadrature I had the feeling that there must be a way to define Gauss quadrature for n beyond what Pozza *et al.* had analyzed. So I wrote:

A limitation of the theory of [21] and of the present manuscript is the restriction to indices $j \leq n$. In contrast, for the nonsymmetric Lanczos algorithm there is the look-ahead Lanczos theory that allows us to recover from a breakdown (or a near-breakdown) and to proceed beyond it. Admittedly, this theory and its application to stable algorithms is complicated and still not implemented well enough. But I could imagine that there is a corresponding theory for complex Gauss quadrature.

In the revised version of the paper that appeared in ETNA the three authors point out that there is indeed such a generalization. It is due to Draux and already part of his thesis (Draux (1983)).

This general definition of Gauss quadrature is reviewed in detail in Pozza/Pranić (2021).
Here is the final result;

Gauss quadrature: the general case (cont'd)

Theorem

The n–node Gauss quadrature $\mathcal{G}_n$ exists (and is unique) if and only if $\Delta_{n-1} \neq 0$.
Moreover, if $\Delta_n = \Delta_{n+1} = \cdots = \Delta_{n+j=0}$ then $\mathcal{G}_n$ has degree of exactness at least $2n + j$.
In particular if $n = \nu(t)$, then $\mathcal{G}_n$ has (maximal) degree of exactness $\nu(t) + \nu(t+1) - 2$, with $\nu(t+1) = +\infty$ when n is the last of the regular FOPs.