Gauss quadrature, still a problem?

Zdeněk Strakoš
Charles University, Prague and Czech Academy of Sciences

http://www.karlin.mff.cuni.cz/˜strakos

Constanta, October 12 - 16, 2016

Gauss quadrature

I. Newton, *Philosophiæ Naturalis Principia Mathematica,* 1723;
R. Cotes, *Harmonia Mensurarum,* 1722

Interpolation

C. F. Gauss, *Methodus nova integralium valores per approximationem inveniendi,* 1814

Interpolatory quadrature with maximal algebraic degree of exactness

C. G. J. Jacobi, *Uber Gauss' neue Methode, die Werthe der Integrale näherungsweise zu finden,* 1826

Orthogonality and three term recurrence

Work recalled in this talk

G. H. Golub, Z. S., *Estimates in quadratic formulas,* 1994

Numerical stability analysis of the error estimates in iterative methods, Gauss and Gauss-Radau, sliding window

Z. S., P. Tichý, *On error estimation in the conjugate gradients and why it works in finite precision computation,* 2002

Based on formulas from the Hestenes and Stiefel 1952 paper; numerical stability analysis

D. P. O'Leary, Z. S. and P. Tichý, *On sensitivity of Gauss-Christoffel quadrature, 2007*

Not of computation of the Gauss-Christoffel quadrature

S. Pozza, M. S. Pranic, Z. S., *Gauss quadrature for quasi-definite linear functionals,* 2016

Generalization of Gauss quadrature to complex plane?

Outline

1. Stieltjes moment problem

2. Gauss Quadrature in complex plane

3. Sensitivity of Gauss quadrature

4. Guaranteed upper bounds for error norms in CG computations?

1 Stieltjes moment problem (1894) of order n

Consider $2n$ real numbers $m_0, m_1, \dots, m_{2n-1}$.
Solve the $2n$ equations

$$\sum_{j=1}^{n} \omega_j^{(n)} \{\theta_j^{(n)}\}^{\ell} \;=\; m_\ell\,, \qquad \ell = 0, 1, \dots, 2n-1\,,$$

for the $2n$ real unknowns $\omega_j^{(n)} > 0,\ \theta_j^{(n)}$.

Is this problem linear?
Does it look easy?
When does it have a solution?

1 Positive definite linear functionals on polynomials

Linear functional $\mathcal{L}(x)$ is positive definite on the space of polynomials $\mathcal{P}_n$ of degree at most n if its first $2n+1$ moments

$$\mathcal{L}(x^\ell) = m_\ell, \quad \ell = 0, 1, \dots, 2n$$

are real and the Hankel matrix M_n of moments is positive definite, i.e., $\Delta_n > 0\,,$ where

$$\Delta_n = |M_n| = \begin{vmatrix} m_0 & m_1 & \cdots & m_n \\ m_1 & m_2 & \cdots & m_{n+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_n & m_{n+1} & \cdots & m_{2n} \end{vmatrix}.$$

1 Related Riemann-Stieltjes integral representation

With the positive definite $\mathcal{L}(x)$ we can restrict ourselves to real polynomials of a real variable and write, using a non-decreasing positive distribution function μ defined on the real axis having finite limits at $\pm\infty$,

$$\mathcal{L}(f) = \int f(x)\,\mathrm{d}\mu(x)\,,$$

with the inner product

$$(p,q) := \mathcal{L}(p(x)q(x)) = \int p(x)q(x)\,\mathrm{d}\mu(x)\,.$$

Solution of the Stieltjes moment problem of order n exists
and it is unique if and only if (with some $m_{2n} > 0$) we have $\Delta_n > 0$.

1 The unknown $\omega_j^{(n)}, \ \theta_j^{(n)}$?

- Cholesky decomposition of the matrix of moments $M_n = L_n L_n^T$

- The entries of the ℓth row of the the inverse L_n^{-1} give the coefficients of the ℓth orthonormal polynomial determined by the positive definite linear functional $\mathcal{L}(x)$ associated with the matrix of moments M_n .

- Roots of the ℓth orthogonal polynomial give the quadrature nodes $\theta_j^{(\ell)}$. The weights $\omega_j^{(\ell)}$ are given by the formula for the interpolatory quadrature.

- Computations are done differently (Gragg and Harrod, Gautschi, Laurie, ...)

Outline

1. Stieltjes moment problem
2. Gauss Quadrature in complex plane
3. Sensitivity of Gauss quadrature
4. Guaranteed upper bounds for error norms in CG computations?

2 Generalization to complex Gauss quadrature?

Linear functional $\mathcal{L}(x)$ is quasi-definite on the space of polynomials $\mathcal{P}_n$ of degree at most n if the Hankel matrix M_n of moments

$$\mathcal{L}(x^\ell) = m_\ell, \quad \ell = 0, 1, \ldots, 2n$$

is strongly regular, i.e., $\Delta_j \neq 0\,, \quad j = 0, 1, \ldots, n\,,$ where

$$\Delta_j = \begin{vmatrix} m_0 & m_1 & \cdots & m_j \\ m_1 & m_2 & \cdots & m_{j+1} \\ \vdots & \vdots & \ddots & \vdots \\ m_j & m_{j+1} & \cdots & m_{2j} \end{vmatrix} .$$

2 Three conditions for generalization

- G1: The n-node Gauss quadrature attains the maximal algebraic degree of exactness $2n-1$.

- G2: The n-node Gauss quadrature is well-defined and it is unique. Moreover, the Gauss quadratures with a smaller number of nodes also exist and they are unique.

- G3: The Gauss quadrature of a function f can be written in the form $m_0\, \mathbf{e}_1^T f(J_n)\mathbf{e}_1$, where J_n is the Jacobi matrix containing the coefficients from the three-term recurrence relation for orthonormal polynomials associated with $\mathcal{L}$; $m_0 = \mathcal{L}(x^0)$.

2 Principal difficulty

Theorem (e.g. Chihara 1978, Lorentzen and Waadeland 1992)

A sequence $\{\pi_j\}_{j=0}^{k}$ of orthogonal polynomials with respect to the linear functional $\mathcal{L}$ exists if and only if $\mathcal{L}$ is quasi-definite on $\mathcal{P}_k$.

Unlike in the positive-definite case, for $\mathcal{L}$ quasi-definite the coefficients of the associated orthogonal polynomials are not necessarily real, the coefficients in the three-term recurrence relation are, in general, complex, and zeros of the orthogonal polynomials can be complex and multiple.

2 n-weight quadrature

Instead of the usual form of an n-node quadrature

$$\mathcal{L}(f) = \sum_{i=1}^{n} \omega_i f(\lambda_i) + R_n(f),$$

where the nodes $\lambda_1, \ldots, \lambda_n$ are complex and distinct (and the last term stands for the quadrature error), we will consider (see also Milovanovic and Cvetkovic 2003) the n-weight quadrature formula

$$\mathcal{L}(f) = \sum_{i=1}^{k} \sum_{j=0}^{s_i-1} \omega_{i,j} \, f^{(j)}(\lambda_i) + R_n(f)\,,$$

where $n = s_1 + \ldots + s_k\,.$

2 Main result

Theorem

Quasi-definitness of the linear functional $\mathcal{L}$ is the necessary and sufficient condition for the n-weight quadrature

$$\mathcal{L}(f) = \sum_{i=1}^{k} \sum_{j=0}^{s_i - 1} \omega_{i,j} \, f^{(j)}(\lambda_i) + R_n(f) \, .$$

to have all three properties G1, G2 and G3. For non-definite linear functionals all three properties cannot hold.

Outline

1. Stieltjes moment problem
2. Gauss Quadrature in complex plane
3. Sensitivity of Gauss quadrature
4. Guaranteed upper bounds for error norms in CG computations?

3 Matching moments

Consider a non-decreasing distribution function $\omega(\lambda)$, $\lambda \geq 0$ with the moments

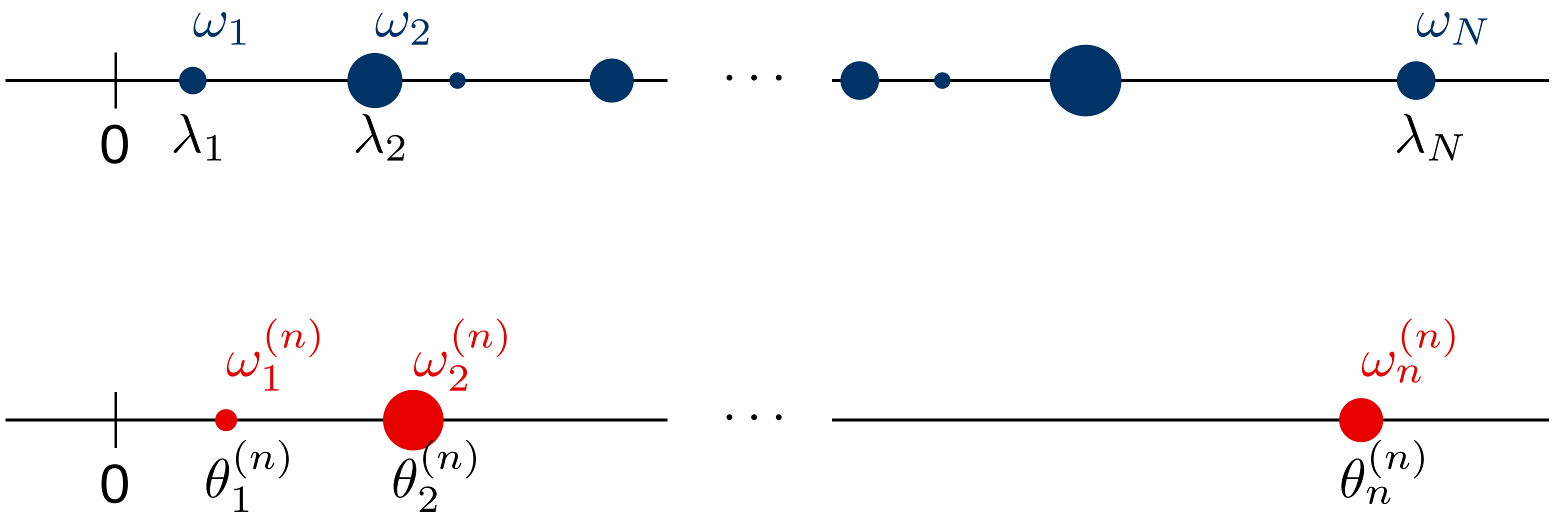
$$m_k = \int_0^\infty \lambda^k \, d\omega(\lambda) \,, \quad k = 0, 1, \ldots \,.$$

Find the distribution function $\omega^{(n)}(\lambda)$ with n points of increase $\theta_i^{(n)}$ which matches the first $2n$ moments for the distribution function $\omega(\lambda)$,

$$\int_0^\infty \lambda^k \, d\omega^{(n)}(\lambda) \equiv \sum_{i=1}^{n} \omega_i^{(n)} (\theta_i^{(n)})^k = m_k, \quad k = 0, 1, \ldots, 2n-1 \,.$$

3 Moment problem illustration (finite case)

For a given n find a distribution function with n mass points in such a way that it in a best way captures the properties of the original distribution function



3 CG is matrix formulation of the Gauss Q

$$Ax = b\,,\ x_0 \quad \longleftrightarrow \quad \int (\lambda)^{-1}\, d\omega(\lambda)$$

$$\uparrow \qquad\qquad\qquad \uparrow$$

$$T_n\, y_n \;=\; \|r_0\|\, e_1 \quad \longleftrightarrow \quad \sum_{i=1}^{n} \omega_i^{(n)} \left(\theta_i^{(n)}\right)^{-1}$$

$$x_n \;=\; x_0 + W_n\, y_n$$

$$\omega^{(n)}(\lambda) \longrightarrow \omega(\lambda)$$

3 CG and Gauss quadrature errors

At any iteration step n, CG represents the matrix formulation of the n-point Gauss quadrature of the R-S integral determined by A and r_0,

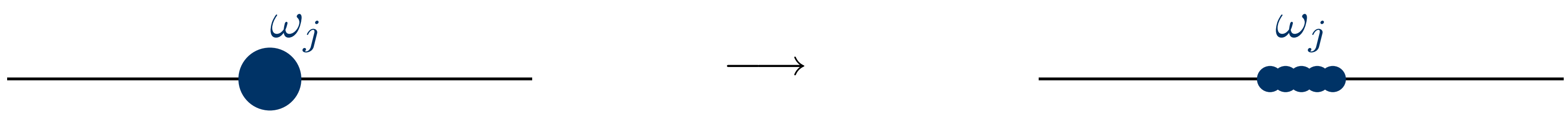
$$\int f(\lambda)\, d\omega(\lambda) \;=\; \sum_{i=1}^{n} \omega_i^{(n)} f(\theta_i^{(n)}) \;+\; R_n(f)\,.$$

For $f(\lambda) \equiv \lambda^{-1}$ the formula takes the form

$$\frac{\|x - x_0\|_{\mathbf{A}}^2}{\|r_0\|^2} \;=\; n\text{-th Gauss quadrature} \;+\; \frac{\|x - x_n\|_{\mathbf{A}}^2}{\|r_0\|^2}\,.$$

This has became a base for the CG error estimation (see above); see the surveys in S and Tichý, 2002; Meurant and S, 2006; Liesen and S, 2013.

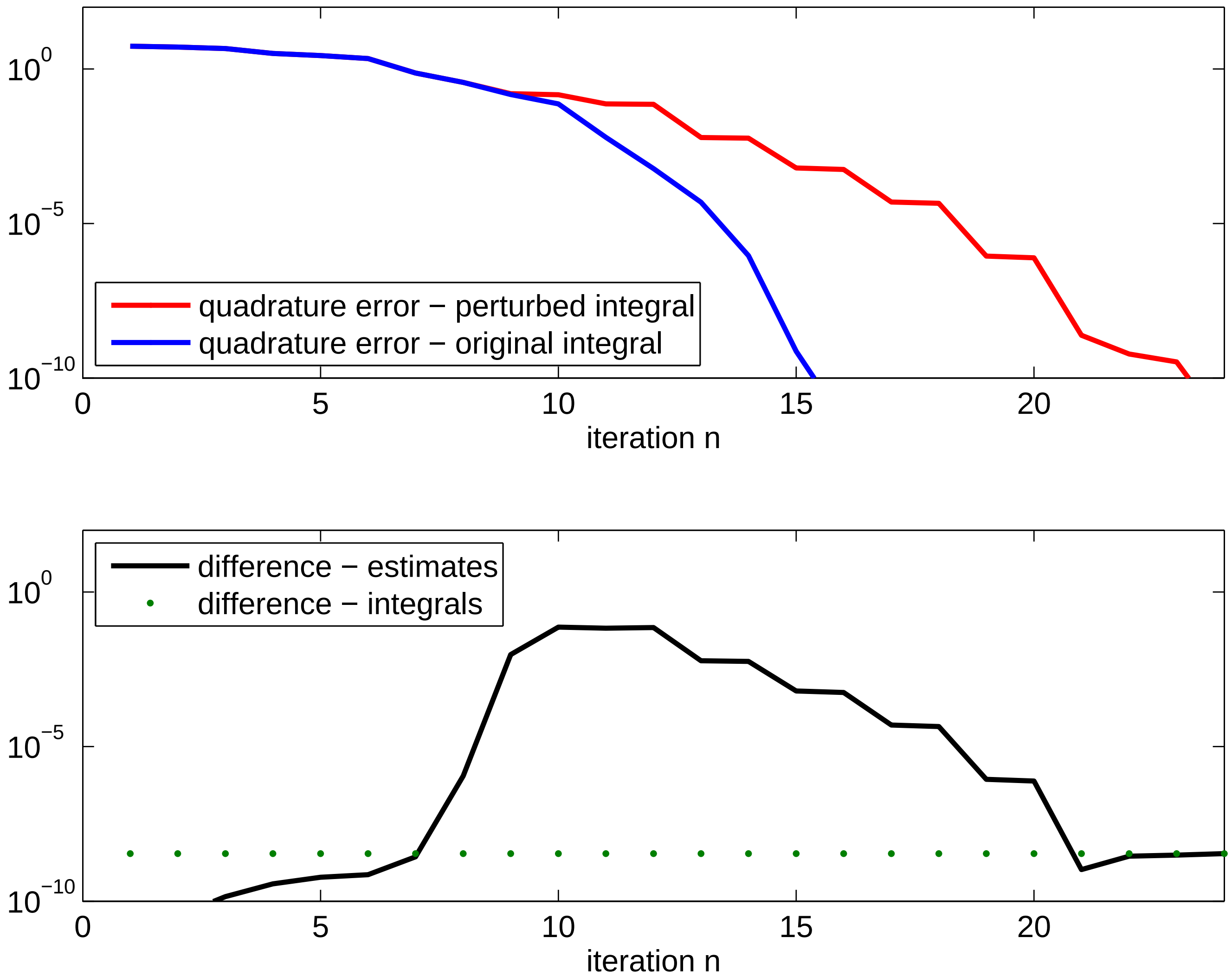
3 Gauss quadrature (CG) with tight clusters



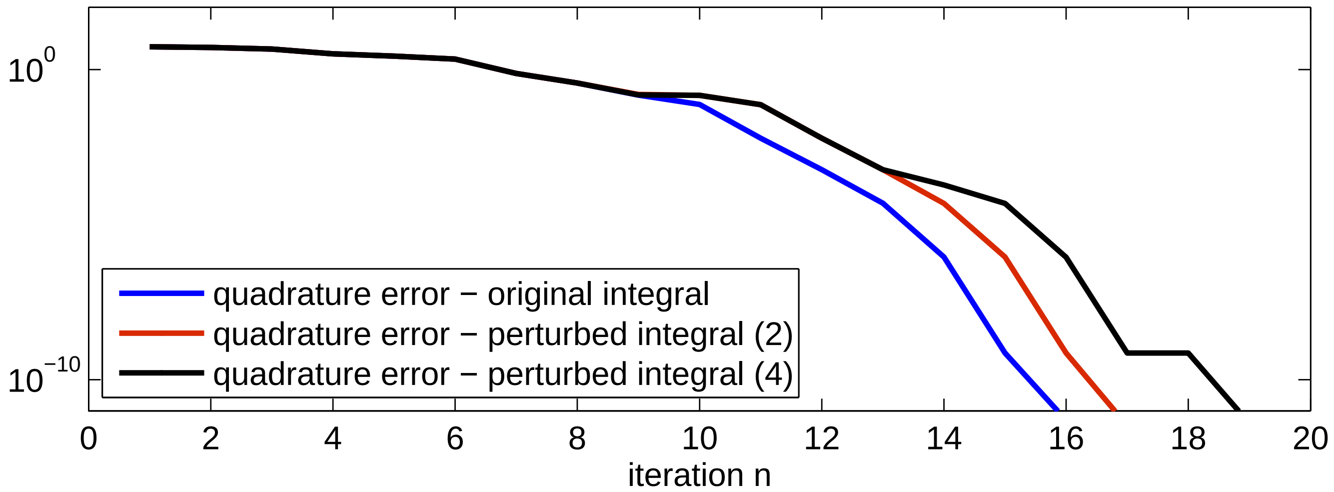
single eigenvalue λ_j $\longrightarrow$ many close eigenvalues $\widehat{\lambda}_{j_1}, \widehat{\lambda}_{j_2}, \ldots, \widehat{\lambda}_{j_\ell}$

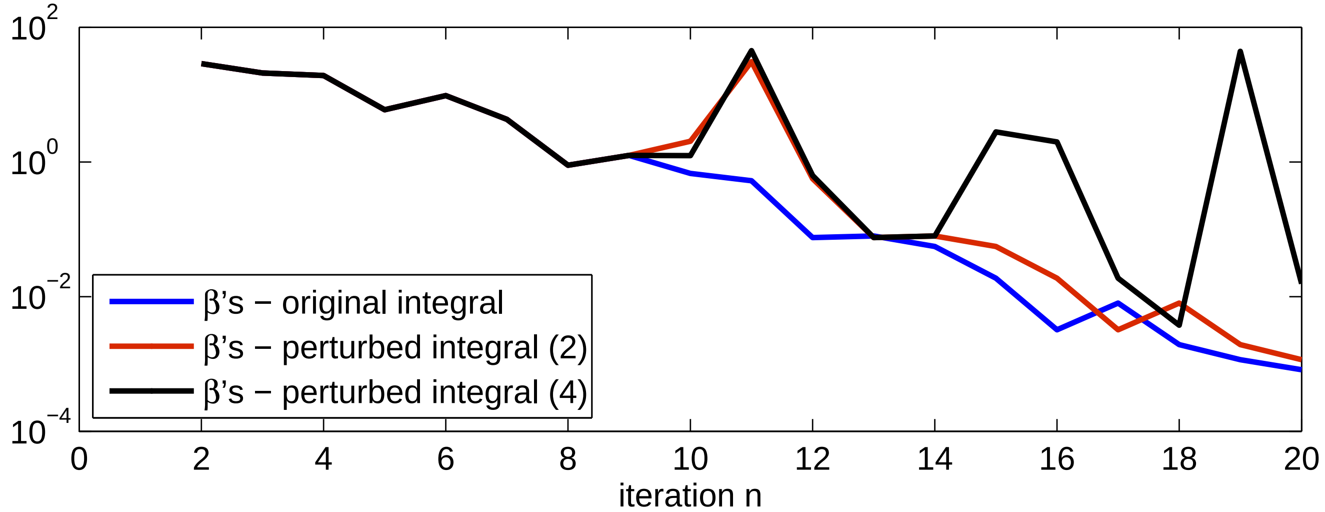
Replacing single eigenvalues by tight clusters can make a difference; see Greenbaum (1989); Greenbaum, S (1992); Golub, S (1994).
Otherwise CG behaves almost linearly and it can be described by contraction. In such case - is it worth using?

3 Sensitivity of the Gauss Quadrature



3 Simplified problem





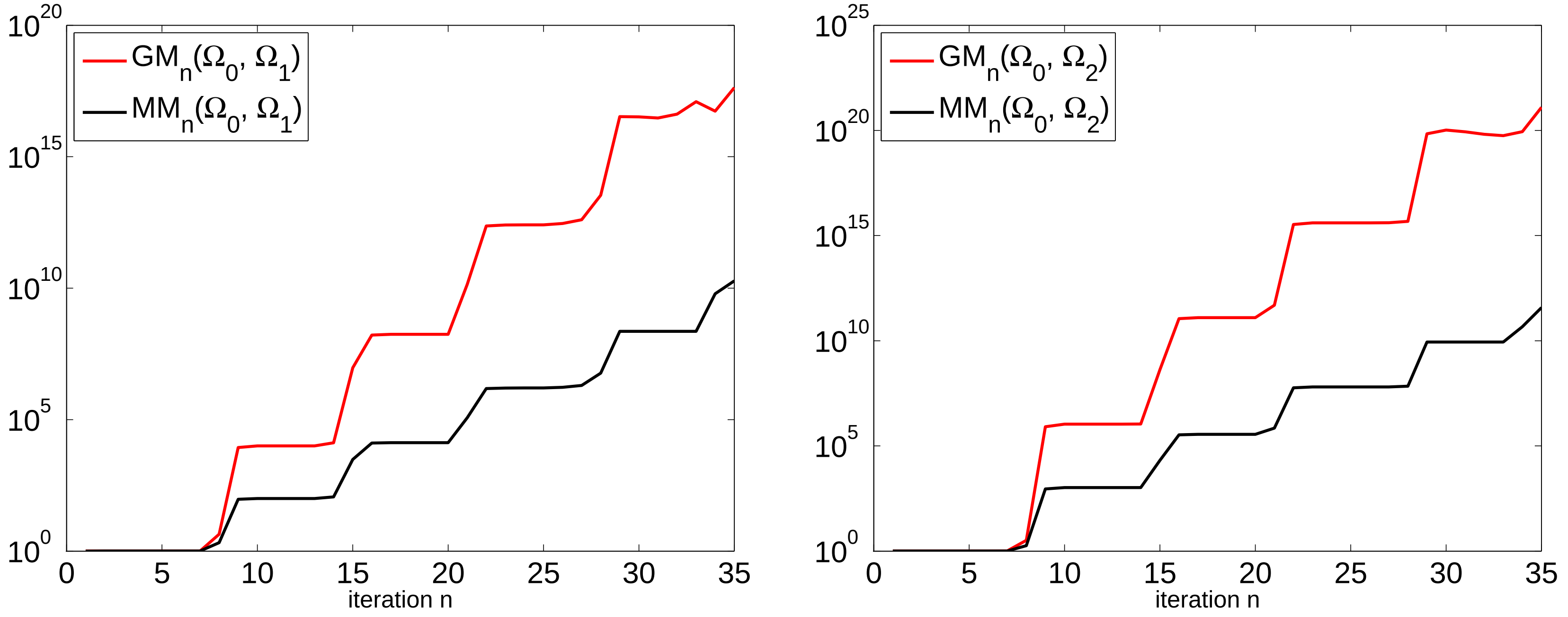
3 : Theorem - O'Leary, S, Tichý (2007)

Consider distribution functions $\omega(x)$ and $\tilde{\omega}(x)$. Let $p_n(x) = (x - x_1) \dots (x - x_n)$ and $\tilde{p}_n(x) = (x - \tilde{x}_1) \dots (x - \tilde{x}_n)$ be the nth orthogonal polynomials corresponding to ω and $\tilde{\omega}$ respectively, with $\hat{p}_c(x) = (x - \xi_1) \dots (x - \xi_c)$ their least common multiple.

If f'' is continuous, then the difference $\Delta^n_{\omega,\tilde{\omega}} = |I^n_\omega - I^n_{\tilde{\omega}}|$ between the approximations I^n_ω to I_ω and $I^n_{\tilde{\omega}}$ to $I_{\tilde{\omega}}$, obtained from the n-point Gauss quadrature, is bounded as

$$
\begin{aligned}
|\Delta^n_{\omega,\tilde{\omega}}| \;&\leq\; \left| \int \hat{p}_c(x) f[\xi_1, \dots, \xi_c, x] \, d\omega(x) - \int \hat{p}_c(x) f[\xi_1, \dots, \xi_c, x] \, d\tilde{\omega}(x) \right| \\
&+\; \left| \int f(x) \, d\omega(x) - \int f(x) \, d\tilde{\omega}(x) \right| .
\end{aligned}
$$

3 Modified moments do not tell the story



Condition numbers of the matrix of the modified moments (GM) and the matrix of the mixed moments (MM). Left - enlarged supports, right - shifted supports.

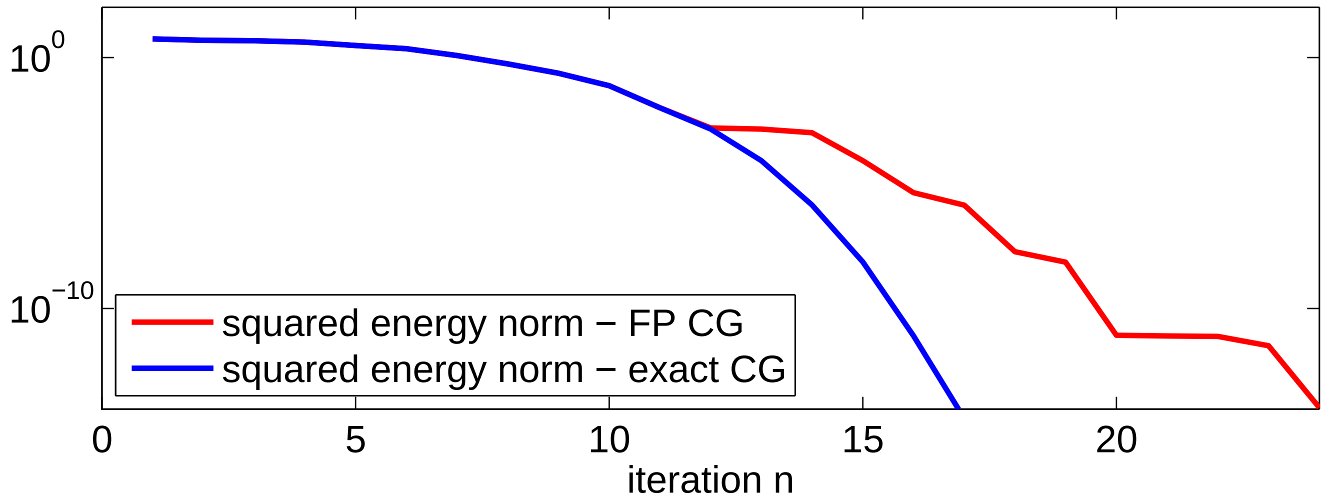
3 Summary

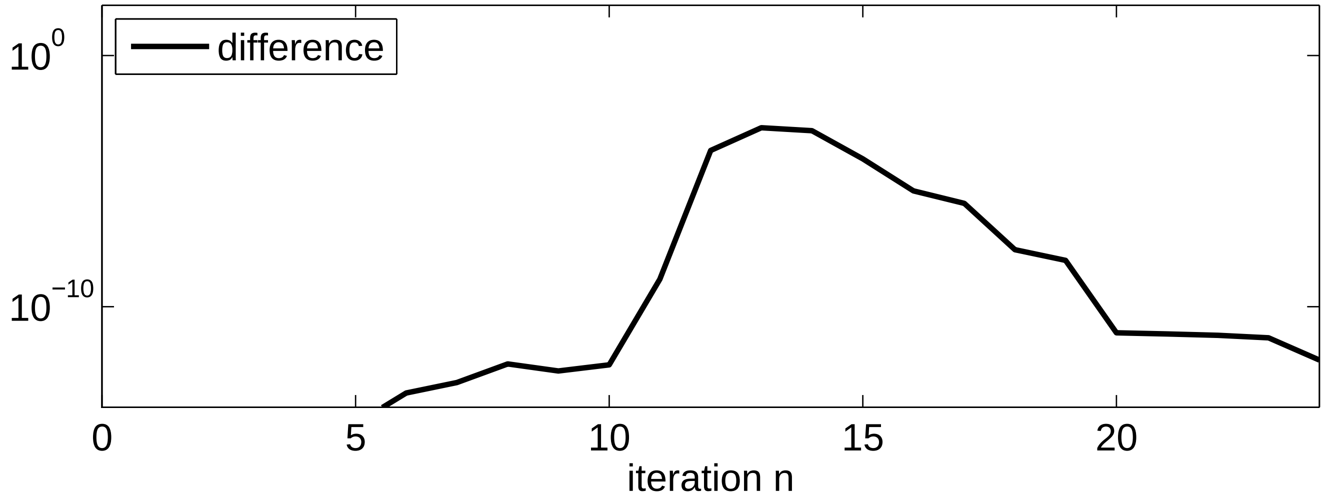
1. Gauss-Christoffel quadrature for a small number of quadrature nodes can be highly sensitive to small changes in the distribution function enlarging its support.

   In particular, the difference between the corresponding quadrature approximations (using the same number of quadrature nodes) can be many orders of magnitude larger than the difference between the integrals being approximated.

2. This sensitivity in Gauss-Christoffel quadrature can be observed for discontinuous, continuous, and even analytic distribution functions, and for analytic integrands uncorrelated with changes in the distribution functions and with no singularity close to the interval of integration.

3 CG in finite precision arithmetic





Outline

1. Stieltjes moment problem

2. Gauss Quadrature in complex plane

3. Sensitivity of Gauss quadrature

4. Guaranteed upper bounds for error norms in CG computations?

4 When are exact precision results applicable?

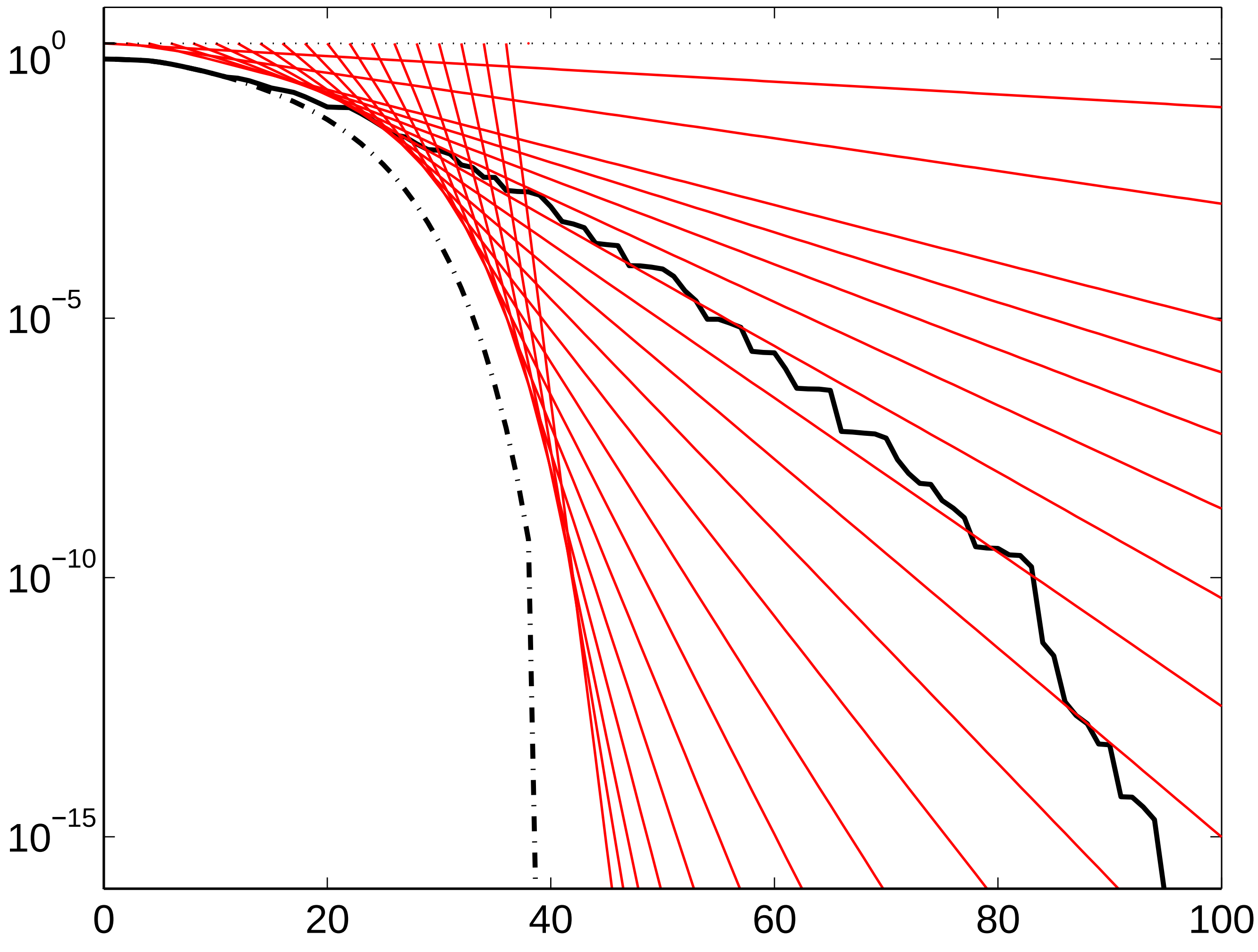
- Assuming exact arithmetic, Gauss quadrature gives a lover bound on the energy norm of the CG error.

- Assuming exact arithmetic, Gauss-Radau quadrature gives a reasonable upper bound on the energy norm of the CG error, providing that we have a very tight (positive) approximation of the smallest eigenvalue of the system matrix.

- CG behaviour in finite precision arithmetic can be very different from its exact arithmetic counterpart. Without numerical stability analysis, no claim on estimating its behavior using formulas derived under the exact arithmetic assumptions can be made.

- Golub, S, 1994 : Gauss quadrature estimates are relevant even in finite precision arithmetic; see also S, Tichý, 2002. But there is no way we can guarrantee that the Gauss-Radau-based estimates give in finite precision arithmetic an upper bound.

4 Delay of convergence due to inexactness



Thank you for your patience and help!