Sensitivity of Gauss quadrature to changes in distribution function - an issue that has been missed?

Zdeněk Strakoš
Charles University, Prague
Jindřich Nečas Center for Mathematical Modelling

Edinburgh Mathematical Society, Strathclyde, January 2020

- Methods used in the 21st century scientific computations can have deep mathematical foundations. This can lead to unexpected and important results. An example:

- Methods used in the 21st century scientific computations can have deep mathematical foundations. This can lead to unexpected and important results. An example:

- The work of Stieltjes (Gauss, Jacobi, ...) helps in understanding the crucial problem of rounding error propagation in the Lanczos and conjugate gradient methods based on short recurrences.

Unexpected mathematical links

- Methods used in the 21st century scientific computations can have deep mathematical foundations. This can lead to unexpected and important results. An example:

- The work of Stieltjes (Gauss, Jacobi, ...) helps in understanding the crucial problem of rounding error propagation in the Lanczos and conjugate gradient methods based on short recurrences.

- Vice-versa, rounding error propagation in the Lanczos and conjugate gradient methods reveals the sensitivity of Gauss quadrature to particular modifications of the distribution function.

Unexpected mathematical links

- Methods used in the 21st century scientific computations can have deep mathematical foundations. This can lead to unexpected and important results. An example:
- The work of Stieltjes (Gauss, Jacobi, ...) helps in understanding the crucial problem of rounding error propagation in the Lanczos and conjugate gradient methods based on short recurrences.
- Vice-versa, rounding error propagation in the Lanczos and conjugate gradient methods reveals the sensitivity of Gauss quadrature to particular modifications of the distribution function.
- We will go through a large and interdisciplinary territory, therefore some important ideas and results will be mentioned only very briefly.

Outline

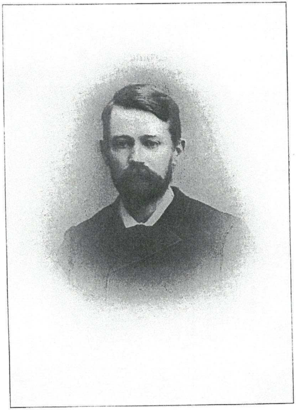
1. Problem of moments: Stieltjes (1894), Vorobyev (1958)

2. Gauss(1814), Lanczos (1950-52), Hestenes and Stiefel (1952):
   Lanczos and conjugate gradient methods for solving large systems of linear (algebraic) equations and approximating eigenvalues

3. Replacing individual points of increase in the associated distribution functions by tight clusters

4. Krylov subspace methods - mathematical challenge across disciplines

1 Thomas Jan Stieltjes (1856 - 1894)



Thomas Jan Stieltjes
1856–1894

Investigations on Continued Fractions

T. J. Stieltjes

Ann. Fac. Sci. Toulouse **8** (1894) J.1–122; **9** (1895) A.1–47 (translation)

Introduction

The object of this work is the study of the continued fraction

$$\cfrac{1}{a_1 z + \cfrac{1}{a_2 + \cfrac{1}{a_3 z + \cdots + \cfrac{1}{a_{2n} + \cfrac{1}{a_{2n+1} z + \cdots,}}}}} \qquad \text{(I)}$$

in which the a_i are positive real numbers, while z is a variable which can take all real or complex values.

Denoting by $\frac{P_n(z)}{Q_n(z)}$ the nth convergent[1], which depends only on the first n coefficients a_i, we shall determine in which cases this convergent tends to a limit for $n \to \infty$ and we shall investigate more closely the nature of this limit regarded as a function of z.

We shall summarize the principal result of this study. There are two distinct cases.

First case. – The series $\sum_1^\infty a_n$ is convergent.

In this case we have for each finite value of z,

$$\lim P_{2n}(z) = p(z),$$
$$\lim Q_{2n}(z) = q(z),$$
$$\lim P_{2n+1}(z) = p_1(z),$$
$$\lim Q_{2n+1}(z) = q_1(z),$$

$p(z), q(z), p_1(z), q_1(z)$ being holomorphic functions in the whole plane which satisfy the relation

$$q(z)p_1(z) - q_1(z)p(z) = +1.$$

These functions are of genus zero and admit only simple zeros which are

1 Continued fractions - approximation of (not only) irrationals

$$1+\frac{1}{\boxed{2}} \qquad 1+\frac{1}{\boxed{2+\frac{1}{2}}} \qquad 1+\frac{1}{2+\frac{1}{2+\frac{1}{2}}} \qquad 1+\frac{1}{2+\frac{1}{2+\frac{1}{2+\frac{1}{2+\ddots}}}}$$

$$=1.5 \qquad =1.4 \qquad =1.4166\bar{6} \qquad \longrightarrow \sqrt{2}$$

The nth convergent

$$\mathcal{F}_n(\lambda) \equiv \cfrac{1}{\lambda - \gamma_1 - \cfrac{\delta_2^2}{\lambda - \gamma_2 - \cfrac{\delta_3^2}{\lambda - \gamma_3 - \dots \cfrac{\ddots}{\lambda - \gamma_{n-1} - \cfrac{\delta_n^2}{\lambda - \gamma_n}}}}} = \frac{\mathcal{R}_n(\lambda)}{\mathcal{P}_n(\lambda)} .$$

Stieltjes (1894): "*we shall determine in which cases this convergent tends to a limit for* $n \to \infty$ *and we shall investigate more closely the nature of this limit regarded as a function of* λ*.*"

Here we use notation different from Stieltjes (1894), in particular $\lambda \equiv -z$.

1 Remarkable history

- Euclid (300BC), Hippassus from Metapontum (before 400BC), ... ,
- Bhascara II (around 1150), Brouncker and Wallis (1655-56):
  Three term recurrences (for numbers)
- Euler (1737, 1748), ... , Brezinski (1991), Khrushchev (2008)
- Gauss (1814), Jacobi (1826), Christoffel (1858, 1857), ... ,
  Chebyshev (1855, 1859), Markov (1884), Stieltjes (1884, 1893-94):
  Orthogonal polynomials, quadrature, analytic theory of continued fractions, problem of moments, minimal partial realization, Riemann-Stieltjes integral
  Gautschi (1981, 2004), Brezinski (1991), Van Assche (1993), Kjeldsen (1993)
- Hilbert (1906, 1912), ... , Von Neumann (1927, 1932), Wintner (1929):
  resolution of unity, integral representation of operator functions, mathematical foundation of quantum mechanics

1 Matrix computation and control theory context

- Krylov (1931), Lanczos (1950, 1952, 1952c), Hestenes and Stiefel (1952), Rutishauser (1953), Henrici (1958), Stiefel (1958), Rutishauser (1959), ... , Vorobyev (1954, 1958, 1965), Golub and Welsch (1968), ... , Laurie (1991 - 2001), ...
- Gordon (1968), Schlesinger and Schwartz (1966), Steen (1973), Reinhard (1979), ... , Horáček (1983 - ...), Simon (2007 - ...)
- Paige (1971, 1972, 1976, 1980), Reid (1971), Greenbaum (1989), ...
- Magnus (1962a,b), Gragg (1974), Kalman (1979), Gragg, Lindquist (1983), Gallivan, Grimme, Van Dooren (1994), ...

1 Moment problem considered by Stieltjes (1894)

Consider an infinite sequence of real numbers $m_0, m_1, m_2, \ldots$.

Find the necessary and sufficient conditions for the existence of the Riemann-Stieltjes integral with the (positive nondecreasing) distribution function $\omega(\lambda)$ such that

$$\int_0^\infty \lambda^\ell \, d\omega(\lambda) \; = \; m_\ell \, , \qquad \ell = 0, 1, 2, \ldots$$

and determine $\omega(\lambda)$.

1 Moment problem considered by Stieltjes (1894)

Consider an infinite sequence of real numbers $m_0, m_1, m_2, \ldots$.

Find the necessary and sufficient conditions for the existence of the Riemann-Stieltjes integral with the (positive nondecreasing) distribution function $\omega(\lambda)$ such that

$$\int_0^\infty \lambda^\ell \, d\omega(\lambda) \; = \; m_\ell \,, \qquad \ell = 0, 1, 2, \ldots$$

and determine $\omega(\lambda)$.

Related moment problem can also be formulated while approximating bounded linear (positive definite self-adjoint) operators in Hilbert spaces; see Vorobyev (1958, 1965).

Let $\mathcal{B}$ be a bounded linear operator on Hilbert space V. Choosing an initial element z_0, we first form a sequence of elements $z_1, z_2, \ldots, z_n, \ldots$ such that

$$z_0,\ z_1 = \mathcal{B}z_0,\ z_2 = \mathcal{B}z_1 = \mathcal{B}^2 z_0,\ \ldots\ , z_n = \mathcal{B}z_{n-1} = \mathcal{B}^n z_{n-1},\ \ldots$$

At the the present time, $z_1, \ldots, z_n$ are assumed to be linearly independent. Determine a sequence of operators $\mathcal{B}_n$ defined on the sequence of nested subspaces V_n generated by $z_0, z_1, z_2, \ldots, z_{n-1}$, $n = 1, 2, \ldots$ such that

$$\begin{aligned} z_1 &= \mathcal{B}z_0 = \mathcal{B}_n z_0, \\ z_2 &= \mathcal{B}^2 z_0 = (\mathcal{B}_n)^2 z_0, \\ &\vdots \\ z_{n-1} &= \mathcal{B}^{n-1} z_0 = (\mathcal{B}_n)^{n-1} z_0, \\ E_n z_n &= E_n \mathcal{B}^n z_0 = (\mathcal{B}_n)^n z_0. \end{aligned}$$

1 Interpretation as model reduction using Krylov subspaces

Using the projection E_n onto V_n we can write for the operators constructed above (here we need the linearity of $\mathcal{B}$)

$$\mathcal{B}_n = E_n \mathcal{B} E_n .$$

The finite dimensional operators $\mathcal{B}_n$ can be used to obtain approximate solutions to various linear problems. The choice of the elements $z_0, \ldots, z_n, \ldots$ as above gives Krylov subspaces that are determined by:

- the operator (given by, e.g., a partial differential equation)
- and the initial element z_0 (given by, e.g., boundary conditions and outer forces).

Two key ingrediences:

I. Krylov subspaces,
II. Projections that can lead to optimality.

See the method of conjugate gradients using orthogonal projections (to follow).

Outline

1. Problem of moments: Stieltjes (1894), Vorobyev (1958)
2. Gauss(1814), Lanczos (1950-52), Hestenes and Stiefel (1952): Lanczos and conjugate gradient methods for solving large systems of linear (algebraic) equations and approximating eigenvalues
3. Replacing individual points of increase in the associated distribution function by tight clusters
4. Krylov subspace methods - mathematical challenge across disciplines

2 Conjugate Gradient (CG) method for $Ax = b$ with A SPD (1952)

$r_0 = b - Ax_0,\ p_0 = r_0$. For $n = 1, \ldots, n_{\max}$:

$$\begin{aligned} \alpha_{n-1} &= \frac{r_{n-1}^* r_{n-1}}{p_{n-1}^* A p_{n-1}} \\ x_n &= x_{n-1} + \alpha_{n-1} p_{n-1}, \quad \text{stop when the stopping criterion is satisfied} \\ r_n &= r_{n-1} - \alpha_{n-1} A p_{n-1} \\ \beta_n &= \frac{r_n^* r_n}{r_{n-1}^* r_{n-1}} \\ p_n &= r_n + \beta_n p_{n-1} \end{aligned}$$

Here α_{n-1} ensures the minimization of the energy norm $\|x - x_n\|_A$ along the line

$$z(\alpha) = x_{n-1} + \alpha p_{n-1}\,.$$

Provided that

$$p_i \perp_A p_j, \quad i \neq j,$$

the one-dimensional line minimizations at the individual steps 1 to n result in the n-dimensional minimization over the whole shifted Krylov subspace

$$x_0 + \mathcal{K}_n(A, r_0) \; = \; x_0 + \operatorname{span}\{p_0, p_1, \dots, p_{n-1}\}\,.$$

Indeed,

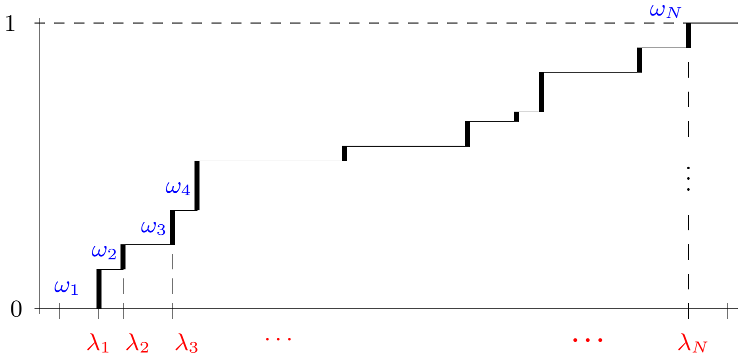
$$x - x_0 \; = \; \sum_{\ell=0}^{N-1} \alpha_\ell p_\ell \; = \; \sum_{\ell=0}^{n-1} \alpha_\ell p_\ell \;\; + \;\; x - x_n\,,$$

where

$$x - x_n \;\perp_A\; K_n(A, r_0)\,, \quad \text{or, equivalently,} \quad r_n \;\perp\; K_n(A, r_0)\,.$$

Distribution function $\omega(\lambda)$ associated with $Ax = b,\ r_0 = b - Ax_0,\ A$ SPD,

$$\lambda_i, y_i \text{ are the eigenpairs of } A\,,\ \omega_i = |(y_i, w_1)|^2\,, \quad (w_1 = r_0/\|r_0\|)$$



First moment

$$
\begin{aligned}
w_1^* A \, w_1 &= w_1^* \left(\sum_{\ell=1}^{N} \lambda_\ell \, y_\ell y_\ell^* \right) w_1 \equiv w_1^* \left(\int \lambda \, dE(\lambda) \right) w_1 \\
&= \sum_{\ell=1}^{N} \lambda_\ell \, |(y_\ell, w_1)|^2 = \sum_{\ell=1}^{N} \lambda_\ell \, \omega_\ell = \int \lambda \, d\omega(\lambda) \, ,
\end{aligned}
$$

where the spectral function $E(\lambda)$ of A is understood to be a nondecreasing family of projections with increasing λ, symbolically $dE(\lambda_\ell) \equiv y_\ell y_\ell^*$ and

$$
I = \sum_{\ell=1}^{N} y_\ell y_\ell^* \equiv \int dE(\lambda) \, .
$$

Hilbert (1906, 1912, 1928), Von Neumann (1927, 1932), Wintner (1929) .

2 First moment of the resolvent using continued/partial fractions

$$
\begin{aligned}
w_1^* \, (zI - A)^{-1} \, w_1 \quad &= \quad \int_0^\infty \frac{d\omega(\lambda)}{z-\lambda} \; = \; \sum_{j=1}^{N} \frac{\omega_j}{z-\lambda_j} \; = \; \frac{\mathcal{R}_N(z)}{\mathcal{P}_N(z)} \; = \; \mathcal{F}_N(z) \\
&\approx \quad \sum_{j=1}^{n} \frac{\omega_j^{(n)}}{z-\lambda_j^{(n)}} \; = \; \frac{\mathcal{R}_n(z)}{\mathcal{P}_n(z)} \; = \; \mathcal{F}_n(z)\,,
\end{aligned}
$$

The denominator $\mathcal{P}_n(z)$ corresponding to the nth convergent $\mathcal{F}_n(z)$ of $\mathcal{F}_N(z)$, $n = 1, 2, \ldots$ is the nth monic orthogonal polynomial in the sequence determined by the distribution function ω and the numerator $\mathcal{R}_n(z)$ is determined by the same recurrence started instead of 1 and z with 0 and 1, see Chebyshev (1855).

- With $\omega(\lambda)$ determined by the SPD A and r_0, solve the finite Stieltjes moment problem, i.e., determine the distribution function $\omega^{(n)}(\lambda)$ with the n points of increase such that the first $2n$ moments are matched, i.e.,

$$m_\ell = \int_0^\infty \lambda^\ell \, d\omega(\lambda) = \int_0^\infty \lambda^\ell \, \omega^{(n)}(\lambda) \,, \qquad \ell = 0, 1, 2, \ldots, 2n-1 \,, \quad n < N \,.$$

- With $\omega(\lambda)$ determined by the SPD A and r_0, solve the finite Stieltjes moment problem, i.e., determine the distribution function $\omega^{(n)}(\lambda)$ with the n points of increase such that the first $2n$ moments are matched, i.e.,

$$m_\ell = \int_0^\infty \lambda^\ell \, d\omega(\lambda) = \int_0^\infty \lambda^\ell \, \omega^{(n)}(\lambda) \,, \qquad \ell = 0, 1, 2, \ldots, 2n-1 \,, \quad n < N \,.$$

- Equivalently, compute the nodes and weights of the n-point Gauss quadrature for the Riemann-Stieltjes integral with the distribution function $\omega(\lambda)$.

2 Gauss quadrature ???

- With $\omega(\lambda)$ determined by the SPD A and r_0, solve the finite Stieltjes moment problem, i.e., determine the distribution function $\omega^{(n)}(\lambda)$ with the n points of increase such that the first $2n$ moments are matched, i.e.,

$$m_\ell = \int_0^\infty \lambda^\ell \, d\omega(\lambda) = \int_0^\infty \lambda^\ell \, \omega^{(n)}(\lambda) \,, \qquad \ell = 0, 1, 2, \dots, 2n-1 \,, \quad n < N \,.$$

- Equivalently, compute the nodes and weights of the n-point Gauss quadrature for the Riemann-Stieltjes integral with the distribution function $\omega(\lambda)$.

- Apply the Vorobyev method of moments to the vectors $z_0, z_1, \dots, z_n$ given by $r_0, Ar_0, \dots, A^n r_0$, which corresponds to matching the first $2n$ moments

$$w_1^* A^\ell w_1 = w_1^* A_n^\ell w_1 = e_1^* T_n^\ell e_1 \,, \qquad \ell = 0, 1, \dots, 2n-1 \,.$$

- With $\omega(\lambda)$ determined by the SPD A and r_0, solve the finite Stieltjes moment problem, i.e., determine the distribution function $\omega^{(n)}(\lambda)$ with the n points of increase such that the first $2n$ moments are matched, i.e.,

$$m_\ell = \int_0^\infty \lambda^\ell \, d\omega(\lambda) = \int_0^\infty \lambda^\ell \, \omega^{(n)}(\lambda), \qquad \ell = 0, 1, 2, \dots, 2n-1, \quad n < N.$$

- Equivalently, compute the nodes and weights of the n-point Gauss quadrature for the Riemann-Stieltjes integral with the distribution function $\omega(\lambda)$.

- Apply the Vorobyev method of moments to the vectors $z_0, z_1, \dots, z_n$ given by $r_0, Ar_0, \dots, A^n r_0$, which corresponds to matching the first $2n$ moments

$$w_1^* A^\ell w_1 = w_1^* A_n^\ell w_1 = e_1^* T_n^\ell e_1, \qquad \ell = 0, 1, \dots, 2n-1.$$

- Equivalently, consider the system $A\,x = b$ with the initial approximation x_0 and compute n iterations of the conjugate gradient/Lanczos method.

Let $W_n = [w_1, \dots, w_n]$, $\quad A\,W_n \;=\; W_n\,T_n \;+\; \delta_{n+1} w_{n+1} e_n^T$, form the Lanczos orthonormal basis of the Krylov subspace $K_n(A, r_0)$. Here the Jacobi matrix of the orthonormalization coefficients

$$T_n = \begin{pmatrix} \gamma_1 & \delta_2 & & & \\ \delta_2 & \ddots & \ddots & & \\ & \ddots & \ddots & \ddots & \\ & & \ddots & \ddots & \delta_n \\ & & & \delta_n & \gamma_n \end{pmatrix}$$

represents, at the same time, the matrix of the restricted and orthogonally projected operator $A_n = W_n W_n^* A$ on $K_n(A, r_0)$ in the basis W_n.
The CG approximation is determined by

$$T_n\, t_n \;=\; \|r_0\| e_1\,, \quad x_n \;=\; x_0 + W_n\, t_n\,.$$

$$A,\ v_1 = r_0/\|r_0\| \quad \longleftrightarrow \quad \omega(\lambda),\ \int f(\lambda)\, d\omega(\lambda)$$

$$\downarrow \qquad\qquad\qquad\qquad \downarrow$$

Hermitian Lanczos/CG — Gauss-Christoffel quadrature

$$\downarrow \qquad\qquad\qquad\qquad \downarrow$$

$$T_n,\ e_1 \quad \longleftrightarrow \quad \omega^{(n)}(\lambda),\quad \sum_{j=1}^{n} \omega_j^{(n)} f(\theta_j^{(n)})$$

The quadrature nodes $\lambda_j^{(n)}$ are the eigenvalues $\theta_j^{(n)}$ of T_n and the weights $\omega_j^{(n)}$ are the squared first components of the associated normalized eigenvectors.

2 Matching moments equations

$$\sum_{i=1}^{n} \omega_i^{(n)} \{\theta_i^{(n)}\}^\ell \;=\; m_\ell \,, \qquad \ell \;= 0, 1, 2, \ldots, 2n-1, \quad n < N.$$

System of $2n$ nonlinear equations for $2n$ unknowns

$$\omega_i^{(n)} \quad \text{and} \quad \theta_i^{(n)} \,, \quad i = 1, \ldots, n \,.$$

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

$$\int_0^\infty f(\lambda)\, d\omega(\lambda) \;=\; \sum_{i=1}^{n} \omega_i^{(n)} f(\theta_i^{(n)}) \;+\; R_n(f)\,.$$

For $f(\lambda) \equiv \lambda^{-1}$,

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

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

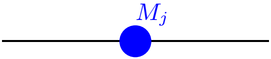
$$\int_0^\infty f(\lambda)\, d\omega(\lambda) \;=\; \sum_{i=1}^{n} \omega_i^{(n)} f(\theta_i^{(n)}) \;+\; R_n(f)\,.$$

For $f(\lambda) \equiv \lambda^{-1}$,

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

CG was linked with Gauss quadrature and continued fractions in the founding paper Hestenes and Stiefel (1952)!

2 Clustering of eigenvalues can make a difference!



$\longrightarrow$

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

Replacing a single eigenvalue by a tight cluster can make a substantial difference; Greenbaum (1989); Greenbaum, S (1992); Golub, S (1994). This was revealed due to the investigation of the propagation of rounding errors; see Part 3 to follow.

Consider distribution functions $\omega(x)$ and $\tilde{\omega}(x)$. Let

$$p_n(x) = (x - x_1)\dots(x - x_n) \quad \text{and} \quad \tilde{p}_n(x) = (x - \tilde{x}_1)\dots(x - \tilde{x}_n)$$

be the nth orthogonal polynomials corresponding to ω and $\tilde{\omega}$ respectively, with their least common multiple

$$\hat{p}_c(x) = (x - \xi_1)\dots(x - \xi_c)$$

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

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

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

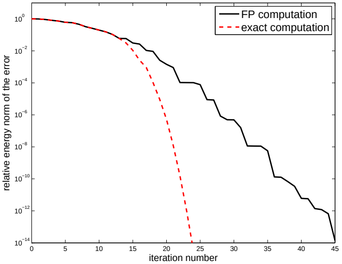
1. Gauss-Christoffel quadrature (for a fixed number of quadrature nodes) can be highly sensitive to changes in the distribution function enlarging its support.

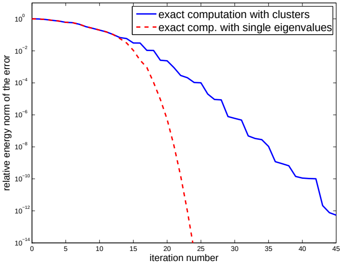
2. This sensitivity in Gauss-Christoffel quadrature can be observed for discontinuous, continuous, and even analytic distribution functions, and for analytic integrands uncorrelated with changes in the distribution functions, and with no singularity close to the interval of integration.

Outline

1. Problem of moments: Stieltjes (1894), Vorobyev (1958)
2. Gauss(1814), Lanczos (1950-52), Hestenes and Stiefel (1952): Lanczos and conjugate gradient methods for solving large systems of linear (algebraic) equations and approximating eigenvalues
3. Replacing individual points of increase in the associated distribution function by tight clusters
4. Krylov subspace methods - mathematical challenge across disciplines

3 FP CG and clustering of eigenvalues in EXACT CG





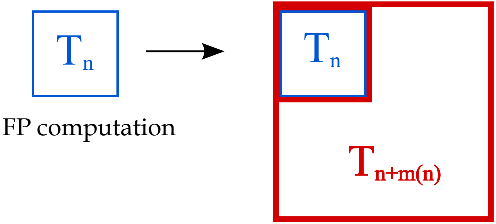
Rounding errors in finite precision CG computations can cause a large delay of convergence.

Exact CG computation for a matrix, where each eigenvalue is replaced by a tight cluster.

Understanding is based on the spectral information in the sequence of the computed (nested) Jacobi matrices $\mathbf{T}_n$, $n = 1, 2, \ldots$.
Seminal contribution of C.C. Paige (1971–80).

3 Beautiful idea of A. Greenbaum (1989)

- Consider the Jacobi matrix $\mathbf{T}_n$ computed in n steps of CG in FP arithmetic. This matrix can be extended to a larger Jacobi matrix $\mathbf{T}_{n+m(n)}$ having all its eigenvalues close to the eigenvalues of the matrix $\mathbf{A}$.
- Then the EXACT CG (Lanczos) applied to this extended Jacobi matrix and the initial residual e_1 gives in the first n steps $\mathbf{T}_n$.



In this way, finite precision computation is viewed and analyzed as exact computation for the problem having clusters of eigenvalues.

1. CG should be used when it has a chance to accelerate its convergence due to adaptation to the information (hidden) in data, e.g., when the eigenvalues are far from being uniformly spread throughout the spectral interval. Presence of large outlying eigenvalues may seem as the most favourable case.

3 Why is consideration of rounding errors so fundamental?

1. CG should be used when it has a chance to accelerate its convergence due to adaptation to the information (hidden) in data, e.g., when the eigenvalues are far from being uniformly spread throughout the spectral interval. Presence of large outlying eigenvalues may seem as the most favourable case.

2. Hovewer, apart from the trivial situation mentioned next, in such cases CG convergence behaviour is typically substantially affected by rounding errors due to the loss of orthogonality among the direction vectors (and residuals).

3 Why is consideration of rounding errors so fundamental?

1. CG should be used when it has a chance to accelerate its convergence due to adaptation to the information (hidden) in data, e.g., when the eigenvalues are far from being uniformly spread throughout the spectral interval. Presence of large outlying eigenvalues may seem as the most favourable case.

2. Hovever, apart from the trivial situation mentioned next, in such cases CG convergence behaviour is typically substantially affected by rounding errors due to the loss of orthogonality among the direction vectors (and residuals).

3. If CG behaviour is not affected by rounding errors, then either we are lucky because convergence is so fast that rounding errors have not enough iterations to amplify (trivial cases), or CG convergence is hopelessly linear with no chance to accelerate. In the latter case linear methods would with high probability be more efficient in terms of computing time (energy consumption).

3 Why is consideration of rounding errors so fundamental?

1. CG should be used when it has a chance to accelerate its convergence due to adaptation to the information (hidden) in data, e.g., when the eigenvalues are far from being uniformly spread throughout the spectral interval. Presence of large outlying eigenvalues may seem as the most favourable case.

2. Hovever, apart from the trivial situation mentioned next, in such cases CG convergence behaviour is typically substantially affected by rounding errors due to the loss of orthogonality among the direction vectors (and residuals).

3. If CG behaviour is not affected by rounding errors, then either we are lucky because convergence is so fast that rounding errors have not enough iterations to amplify (trivial cases), or CG convergence is hopelessly linear with no chance to accelerate. In the latter case linear methods would with high probability be more efficient in terms of computing time (energy consumption).

4. There seems to be no escape from the implications of the presented facts.

3 In which sense is the CG polynomial $\varphi_n^{\mathrm{CG}}(\lambda)$ optimal?

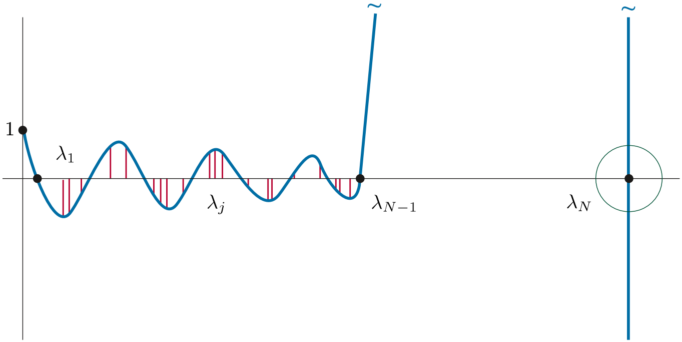
$$\begin{aligned}\|\mathbf{x}-\mathbf{x}_n\|_{\mathbf{A}}^2 &= \min_{\varphi\in\Pi_n} \|\varphi(\mathbf{A})(\mathbf{x}-\mathbf{x}_0)\|_{\mathbf{A}}^2 \\ &= \sum_{j=1}^{N} \lambda_j \, \zeta_j^2 \, \varphi_n^{\mathrm{CG}}(\lambda_j)^2 , \quad j = 1, 2, \ldots\end{aligned}$$

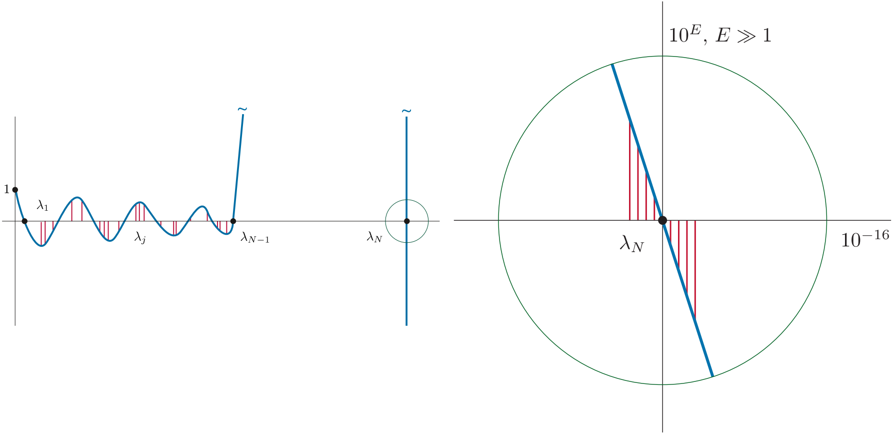
Here

$$\varphi_n^{\mathrm{CG}}(\lambda) = \frac{(\lambda-\theta_1^{(n)})\cdots(\lambda-\theta_n^{(n)})}{(-1)^n\,\theta_1^{(n)}\cdots\theta_n^{(n)}}$$

is determined by the eigenvalues of the orthogonally restricted operator, i.e., by the eigenvalues $\theta_1^{(n)}, \ldots, \theta_n^{(n)}$ of $\mathbf{T}_n$ (Ritz values) .

3 Illustration of $\varphi_n^{\mathrm{CG}}(\lambda)$, $\mathbf{A}\mathbf{x} = \mathbf{b}$, a single large outlier λ_N





Since E in the illustration of the slope is enormous, many roots close to λ_N are needed.

Outline

1. Problem of moments: Stieltjes (1894), Vorobyev (1958)
2. Gauss(1814), Lanczos (1950-52), Hestenes and Stiefel (1952): Lanczos and conjugate gradient methods for solving large systems of linear (algebraic) equations and approximating eigenvalues
3. Replacing individual points of increase in the associated distribution function by tight clusters
4. Krylov subspace methods - mathematical challenge across disciplines

4 Recall the CG polynomial $\varphi_n^{\mathrm{CG}}(\lambda)$ is optimal

$$
\begin{aligned}
\|\mathbf{x}-\mathbf{x}_n\|_{\mathbf{A}}^2 &= \min_{\varphi\in\Pi_n} \|\varphi(\mathbf{A})(\mathbf{x}-\mathbf{x}_0)\|_{\mathbf{A}}^2 \\
&= \sum_{j=1}^{N} \lambda_j\, \zeta_j^2\, \varphi_n^{\mathrm{CG}}(\lambda_j)^2\,, \quad j=1,2,\ldots
\end{aligned}
$$

Here

$$
\varphi_n^{\mathrm{CG}}(\lambda) = \frac{(\lambda-\theta_1^{(n)})\cdots(\lambda-\theta_n^{(n)})}{(-1)^n\,\theta_1^{(n)}\cdots\theta_n^{(n)}}
$$

is determined by the eigenvalues of the orthogonally restricted operator, i.e., by the eigenvalues $\theta_1^{(n)},\ldots,\theta_n^{(n)}$ of $\mathbf{T}_n$ (Ritz values).

4 In contrast, linear "description" of the nonlinear CG method

- The CG optimality property

$$\|x - x_n\|_A = \min_{z \in x_0 + \mathcal{K}_n(A, r_0)} \|x - z\|_A \;=\; \min_{\varphi \in \Pi_n} \|\varphi(A)(x - x_0)\|_A$$

yields in two derivation steps the (worst case) linear convergence bound valid and relevant for the Chebyshev method

$$\frac{\|x - x_n\|_A}{\|x - x_0\|_A} \le \min_{\varphi \in \Pi_n} \max_{1 \le j \le N} |\varphi(\lambda_j)| \le \min_{p \in \Pi} \max_{\lambda \in [\lambda_1, \lambda_N]} |p(\lambda)|$$

$$\le 2 \left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right)^n, \quad \kappa(A) = \frac{\lambda_N}{\lambda_1}.$$

- The worst-case nonlinear bound is completely determined by the distribution of the eigenvalues of A.

4 In contrast, linear "description" of the nonlinear CG method

- The CG optimality property

$$\|x - x_n\|_A = \min_{z \in x_0 + \mathcal{K}_n(A, r_0)} \|x - z\|_A = \min_{\varphi \in \Pi_n} \|\varphi(A)(x - x_0)\|_A$$

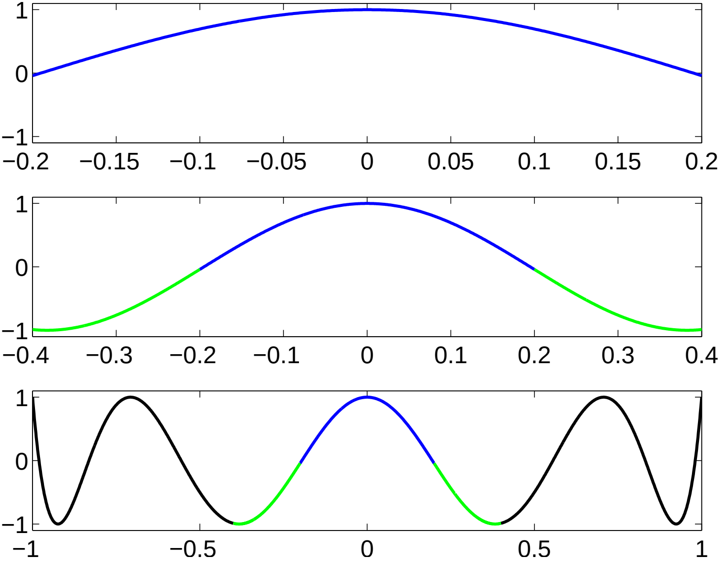
yields in two derivation steps the (worst case) linear convergence bound valid and relevant for the Chebyshev method

$$\frac{\|x - x_n\|_A}{\|x - x_0\|_A} \leq \min_{\varphi \in \Pi_n} \max_{1 \leq j \leq N} |\varphi(\lambda_j)| \leq \min_{p \in \Pi} \max_{\lambda \in [\lambda_1, \lambda_N]} |p(\lambda)|$$

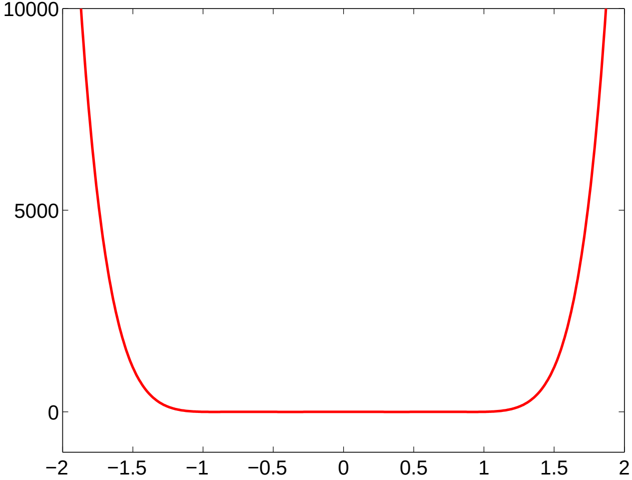
$$\leq 2 \left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right)^n, \quad \kappa(A) = \frac{\lambda_N}{\lambda_1}.$$

- The worst-case nonlinear bound is completely determined by the distribution of the eigenvalues of A.

4 Tschebyshev polynomials of the first kind



4 It grows very fast outside the interval $[-1, 1]$



Definition:

$$T_n(x) = cos(n\,cos^{-1}(x)),\ x \in [-1,1];\ T_n(x) = cosh(n\,cosh^{-1}(x)),\ x \notin [-1,1]\,.$$

Recurrence:

$$T_0(x) = 1\,,\ T_1(x) = x\,,\quad T_{n+1}(x) \;=\; 2\,T_n(x) \;-\; T_{n-1}(x)\,.$$

Orthogonality:

$$\int_{-1}^{1} T_n(x)\,T_m(x)\,(1-x^2)^{-1/2}dx \;=\; 0 \quad \text{for} \quad m \neq n\,.$$

Optimality:

Tschebyshev polynomial is the fastest growing polynomial outside the interval $[-1,1]$ from all polynomials of the given degree that are on the interval $[-1,1]$ in the absolute value less or equal to one.

4 Remarkable history, rarely quoted

- Markov (1890)
- Flanders and Shortley (1950)
- Lanczos (1952–53), Kincaid (1947), Young (1954, ...)
- Stiefel (1958), Rutishauser (1959)
- Meinardus (1963), Kaniel (1966)
- Daniel (1967a, 1967b)
- Luenberger (1969)

Derivations are repeated in recent textbooks and monographs and the resulting bound is identified with the convergence of CG without noticing severe limitations.

4 Statement published in 1952 is right to the point.

C. Lanczos, *Solution of systems of linear equations by minimized iterations*, J. of Research of the National Burreau of Standards, 49 (1952), pp. 33–53:

"The principle by which this process [the conjugate gradient method] *gives good attenuation, is quite different from the previous one.* ['Purification' based on Tschebyshev polynomials.] *Here we take heed of the specific nature of the matrix A and operate in a selective way. The polynomials of this process ... have the peculiarity that they attenuate due to the nearness of their zeros to those λ-values* [eigenvalues] *which are present in A. These polynomials take advantage of the fact that the spectrum to be attenuated is a line spectrum and not a continuous spectrum. They work efficiently in the neighbourhood of the λ_i of the matrix but not for the intermediate values."*

The condition-number-based bound should be used with a great care in connection with the behaviour of CG unless $\kappa(A) = \lambda_N/\lambda_1$ is really small or unless the (very special) distribution of eigenvalues makes the bound tight.

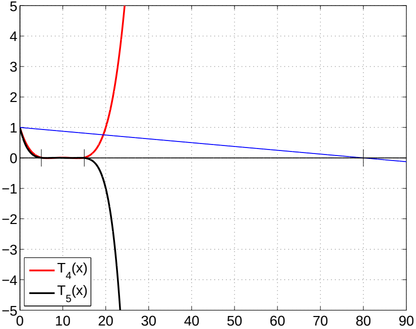
The condition-number-based bound should be used with a great care in connection with the behaviour of CG unless $\kappa(A) = \lambda_N/\lambda_1$ is really small or unless the (very special) distribution of eigenvalues makes the bound tight.

In particular, one should be very careful while using it as a part of a composite bound in the presence of large outlying eigenvalues

$$\min_{\substack{p(0)=1 \\ \deg(p)\le n-s}} \max_{1\le j\le N} \left| q_s(\lambda_j)\, p(\lambda_j) \right| \le \max_{1\le j\le N} |q_s(\lambda_j)| \left| \frac{T_{n-s}(\lambda_j)|}{T_{n-s}(0)} \right|$$

$$< \max_{1\le j\le N-s} \left| \frac{T_{n-s}(\lambda_j)}{T_{n-s}(0)} \right| .$$

This Chebyshev method bound for the spectral interval $[\lambda_1, \lambda_{N-s}]$ is then valid after s initial steps.



A single large outlying eigenvaue:

The shifted and scaled Chebyshev polynomials $T_4(\lambda)$, $T_5(\lambda)$, and the polynomial $q_1(\lambda)\,, q_1(0) = 1$ having the root at the large outlying eigenvalue.

Consider the desired accuracy ϵ, $\kappa_s(A) \equiv \lambda_{N-s}/\lambda_1$. Then, assuming exact arithmetic, n CG steps, where

$$n = \mathbf{s} + \left\lceil \frac{\ln(2/\epsilon)}{2} \sqrt{\kappa_s(A)} \right\rceil ,$$

will produce the approximate solution x_n satisfying

$$\|x - x_n\|_A \leq \epsilon \, \|x - x_0\|_A \, .$$

This statement has been used to explain superlinear convergence of CG at the presence of large outliers in the spectrum. Due to rounding errors, this concept can not be applied in a meaningful way to practical computations. Recall the mathematical model of FP computations that is based on replacing large outlying eigenvalues by large clusters.

- Purely algorithmic view ends the story with the finite termination property.

- Purely algorithmic view ends the story with the finite termination property.
- Oversimplified approximation view replaces nonlinear CG matching moment problem by Tschebyshev approximation relevant to the Tschebyshev method.

- Purely algorithmic view ends the story with the finite termination property.
- Oversimplified approximation view replaces nonlinear CG matching moment problem by Tschebyshev approximation relevant to the Tschebyshev method.
- Globally nonlinear CG for solving linear problem is being confused with CG for solving general nonlinear optimization problems which is based on local linearizations.

- Purely algorithmic view ends the story with the finite termination property.
- Oversimplified approximation view replaces nonlinear CG matching moment problem by Tschebyshev approximation relevant to the Tschebyshev method.
- Globally nonlinear CG for solving linear problem is being confused with CG for solving general nonlinear optimization problems which is based on local linearizations.
- Spectral description is used for interpreting convergence behaviour of Krylov subspace methods even for highly non-normal operators (matrices) without careful investigation. See next slide.

- When the operators (matrices arising from discretization) are far from normal and the spectral information is descriptive for convergence behavior of Krylov subspace methods, this points out to some fundamental mathematical, physical, ... properties of the problem.

- When the operators (matrices arising from discretization) are far from normal and the spectral information is descriptive for convergence behavior of Krylov subspace methods, this points out to some fundamental mathematical, physical, ... properties of the problem.

- In such cases there must be some special inner structure of invariant subspaces and/or special right hand side (in BVP that means boundary conditions and outer forces).

4 Positive results and unexplored research challenges

- When the operators (matrices arising from discretization) are far from normal and the spectral information is descriptive for convergence behavior of Krylov subspace methods, this points out to some fundamental mathematical, physical, ... properties of the problem.
- In such cases there must be some special inner structure of invariant subspaces and/or special right hand side (in BVP that means boundary conditions and outer forces).
- Very little has unfortunately been done in that much needed direction of research. Despite many results proving that information about the spectrum alone can not be descriptive, in general, for convergence behaviour of Krylov subspace methods such as GMRES, the challenge is rarely mentioned in literature.

See the works of Greenbaum, Pták, Arioli, S, Liesen, Eiermann, Ernst, Meurant, Tichý, Duintjer-Tebbens, ... with the first paper published in 1994.

4 Lanczos, Why Mathematics (1966)

“We will go on pondering and meditating, the great mysteries still ahead of us, we will err and stumble on the way, and if we win a little victory, we will be jubilant and thankful, without claiming, however, that we have done something that can eliminate the contribution of all the millenia before us.”

"We will go on pondering and meditating, the great mysteries still ahead of us, we will err and stumble on the way, and if we win a little victory, we will be jubilant and thankful, without claiming, however, that we have done something that can eliminate the contribution of all the millenia before us."

- It can be useful to look on what we do now through the eyes of the great women and men from the past.

4 Lanczos, Why Mathematics (1966)

"We will go on pondering and meditating, the great mysteries still ahead of us, we will err and stumble on the way, and if we win a little victory, we will be jubilant and thankful, without claiming, however, that we have done something that can eliminate the contribution of all the millenia before us."

- It can be useful to look on what we do now through the eyes of the great women and men from the past.
- It can be useful to look at the work of the great women and men from the past within the context of our work now.

4 Wallis, Arithmetica Infinitorium (1656) (see Khruschev 2008)

"There remains this: we beseech the skilled in these things, that what we thought worth showing, they will think openly receiving, an whatever it hides, worth imparting more properly by themselves to the wider mathematical community."

Thank you very much for your kind patience!