Krylov subspace methods and non-self-adjoint operators

Zdeněk Strakoš[1]

[1] Charles University and Czech Academy of Sciences, Prague

Mathematical Aspects of Physics with Non-Self-Adjoint Operators,
CIRM, Luminy, June 5–9, 2017

Historical and other perspectives

- Brouncker and Wallis
- Euler, Gauss, Jacobi
- Chebyshev, Markov, Stieltjes
- Hilbert, von Neumann
- Krylov, Gantmakher
- Lanczos, Hestenes, Stiefel

Cornelius Lanczos, *Why Mathematics*, 1966

" Personally I would not hesitate not only to graduate with first class honors, but to give the Ph.D. (and with summa cum laude) without asking any further questions, to anybody who knew only one quarter of what Euler knew, provided that he knew it in the way in which Euler knew it. "

Analytic $\times$ algorithmic, application and computational perspectives.

Adaptation as the main principle

Cornelius Lanczos, March 9, 1947

"To obtain a solution in very few steps
means nearly always that one has found a way
that does justice to the inner nature of the problem."

Albert Einstein, March 18, 1947

"Your remark on the importance of
adapted approximation methods makes very
good sense to me, and I am convinced
that this is a fruitful mathematical aspect,
and not just a utilitarian one."

- Nonlinear adaptation of the iterations to linear problems?
- Inner nature of the problem?

Outline

1. Infinite dimensional problem and finite dimensional computations
2. Linear projections onto (nonlinear) Krylov subspaces
3. Matching moments model reduction
4. Convergence behavior and spectral information
5. Mathematical structure preserved at the presence of numerical errors
6. Discretization and preconditioning
7. Mathematical myths
8. Conclusions

1. Infinite dimensional problem and finite dimensional computations

Consider a numerical solution of equations

$$\mathcal{G}u = f, \quad f \in V \ ,$$

on an infinite dimensional Hilbert space V , where $\mathcal{G}$ is a linear invertible operator, $\mathcal{G} : V \to V$.

Krylov subspace methods at the step n implicitly construct a finite dimensional approximation $\mathcal{G}_n$ of $\mathcal{G}$ with the desired approximate solution u_n defined by $(u_0 = 0)$

$$u_n := p_{n-1}(\mathcal{G}_n)\, f \approx u = \mathcal{G}^{-1} f \ ,$$

where $p_{n-1}(\lambda)$ is a particular polynomial of degree at most $n-1$ and $\mathcal{G}_n$ is obtained by restricting and projecting $\mathcal{G}$ onto the nth Krylov subspace

$$K_n(\mathcal{G}, f) := \mathrm{span}\left\{ f, \mathcal{G}f, \ldots, \mathcal{G}^{n-1} f \right\} .$$

1 Four basic questions

1. How fast $u_n,\ n = 1, 2, \ldots$ approximate the desired solution u ? Nonlinear adaptation.

2. Which characteristics of $\mathcal{G}$ and f can be used in investigating the previous question? Inner nature of the problem.

3. How to handle efficiently discretization and computational issues? Provided that $K_n(\mathcal{G}, f)$ can be computed, the projection provides discretization of the infinite dimensional problem $\mathcal{G}u = f$.

4. How to handle transformation of $\mathcal{G}u = f$ into an easier-to-solve problem? Preconditioning.

1 Fundamental theorem of discretization of $\mathcal{G}, f$

Consistency deals with the question how closely $\mathcal{G}_h u_h = f_h$ approximates $\mathcal{G}u = f$.
The residual measure

$$\mathcal{G}_h \pi_h u - f_h$$

gives

$$\pi_h u - u_h \;=\; \mathcal{G}_h^{-1}(\mathcal{G}_h \pi_h u - f_h).$$

If $\|\mathcal{G}_h^{-1}\|_h$ is bounded from above uniformly in h (the discretization is stable), then consistency

$$\|\mathcal{G}_h \pi_h u - f_h\|_h \to 0 \quad \text{as} \quad h \to 0$$

implies convergence of the discretization scheme

$$\|\pi_h u - u_h\|_h \to 0 \quad \text{as} \quad h \to 0\ .$$

1 Approximation of bounded invertible operators?

What do we mean by an approximation of $\mathcal{G}$ by a finite dimensional operator?

- A uniform (norm) limit of finite dimensional operators $\mathcal{G}_n$ can only be a compact operator. Moreover, every compact operator on a Hilbert space is a uniform limit of a sequence of finite dimensional operators.
- A uniform limit of a sequence of compact operators can only be a compact operator.

Consequence:

Bounded invertible operators in Hilbert spaces can be approximated by compact or finite dimensional operators only in the sense of strong convergence (pointwise limit)

$$\|\mathcal{G}_h\, w - \mathcal{G}\, w\| \to 0 \quad \text{as} \quad h \to 0 \quad \text{for all } w \in V\,;$$

The convergence $\mathcal{G}_h\, w \to \mathcal{G}\, w$ is not uniform; the role of right hand sides.

- Finite dimensional self-adjoint operators (finite Hermitian matrices)

$$\mathcal{A} = \frac{1}{2\pi\iota}\int_{\Gamma}\lambda\,(\lambda I_N - \mathcal{A})^{-1}\,d\lambda \;=\; \frac{1}{2\pi\iota}\sum_{j=1}^{N}\int_{\Gamma_j}\lambda\,(\lambda I_N - \mathcal{A})^{-1}\,d\lambda$$

$$= \sum_{j=1}^{N} Y\,\mathrm{diag}\left(\frac{1}{2\pi\iota}\int_{\Gamma_j}\frac{1}{\lambda-\lambda_\ell}\,d\lambda\right)\,Y^* \;=\; \sum_{j=1}^{N}\lambda_j\,y_j y_j^*$$

$$= \int_{m(\mathcal{A})}^{M(\mathcal{A})}\lambda\,dE(\lambda)\,.$$

- Compact infinite dimensional self-adjoint operators
- Bounded infinite dimensional self-adjoint operators
- Generalization to bounded normal and non-normal operators

Let Γ_h surrounds the spectra of $\mathcal{G}$ and of its discrete approximation $\mathcal{G}_h$ with a distance related to h. For any polynomial

$$p(\mathcal{G}) - p(\mathcal{G}_h) = \frac{1}{2\pi i} \int_{\Gamma_h} p(\lambda) \left[(\lambda - \mathcal{G})^{-1} - (\lambda - \mathcal{G}_h)^{-1} \right] d\lambda \,.$$

If $\|\mathcal{G} - \mathcal{G}_h\| \to 0$ as $h \to 0$ and, consequently,

$$\|(\lambda - \mathcal{G})^{-1} - (\lambda - \mathcal{G}_h)^{-1}\| \to 0 \quad \text{as} \quad h \to 0 \quad \text{uniformly on} \quad \Gamma_h \,,$$

then it seems that one can investigate $p(\mathcal{G})$ instead of $p(\mathcal{G}_h)$.

But the *assumption* $\|\mathcal{G} - \mathcal{G}_h\| \to 0$ as $h \to 0$ does not hold for any bounded infinite dimensional operator $\mathcal{G}$ having a bounded inverse.

2. Linear projections onto (nonlinear) Krylov subspaces

For more details and references:

Y. Saad, *Iterative Methods for Sparse Linear Systems*, SIAM, Philadelphia, PA, second ed. (2003)

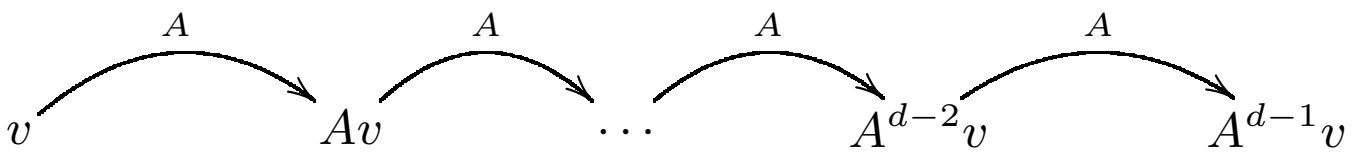
J. Liesen. and Z.S., *Krylov Subspace Methods, Principles and Analysis.* Oxford University Press (2013), Chapter 2

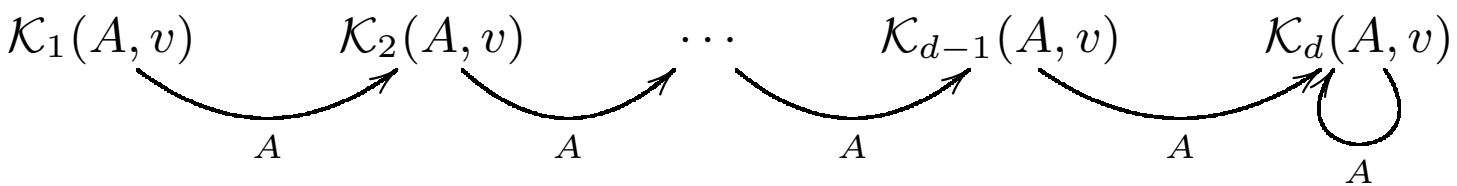
- The *n*th Krylov subspace generated by $A \in \mathbb{C}^{N\times N}$ and $v \in \mathbb{C}^N$

$$\mathcal{K}_n(A,v) := \operatorname{span}\{v, Av, \ldots, A^{n-1}v\}, \quad n = 1, 2, \ldots$$

- By construction,

$$\mathcal{K}_1(A,v) \subset \mathcal{K}_2(A,v) \subset \cdots \subset \mathcal{K}_d(A,v) = \mathcal{K}_{d+k}(A,v) \quad \text{for all } k \geq 1.$$





Krylov subspace methods are based on a sequence of projections onto the nested Krylov subspaces that form the search spaces.

- Linear algebraic system $Ax = b$: x_0 (possibly zero), $r_0 = b - Ax_0$.

  $x_n \in x_0 + \mathcal{S}_n = x_0 + \mathcal{K}_n(A, r_0)$ such that $r_n = b - Ax_n \perp \mathcal{C}_n$, $n = 1, 2, \ldots$

  n-dimensional constraints space $\mathcal{C}_n$ determines the different methods.

- Eigenvalue problem $Ax = \lambda x$: given nonzero v, find (λ_n, x_n) such that

  $$x_n \in \mathcal{K}_n(A, v) \quad \text{and} \quad r_n = Ax_n - \lambda_n x_n \perp \mathcal{C}_n.$$

  Examples: Lanczos and Arnoldi methods, where $\mathcal{C}_n = \mathcal{K}_n(A, v)$.

- Method is well defined when x_n is uniquely determined for $n = 1, 2, \ldots, d-1$, and $x_d = x$ (in exact arithmetic).

- Conjugate gradient (CG) method: $\mathcal{S}_n = \mathcal{C}_n = \mathcal{K}_n(A, r_0)$.
  - Well defined for HPD matrices A; short recurrences.
  - Orthogonality $r_n \perp \mathcal{K}_n(A, v)$ is equivalent to optimality:

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

- GMRES method: $\mathcal{S}_n = \mathcal{K}_n(A, r_0)$, $\mathcal{C}_n = A\mathcal{K}_n(A, r_0)$.
  - Well defined for nonsingular matrices A; full recurrences.
  - Orthogonality $r_n \perp A\mathcal{K}_n(A, v)$ is equivalent to optimality:

$$\|b - Ax_n\|_2 = \min_{z \in x_0 + \mathcal{K}_n(A, r_0)} \|b - Az\|_2.$$

- Closely related FOM method with $\mathcal{S}_n = \mathcal{C}_n = \mathcal{K}_n(A, r_0)$ is not well defined.

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

with the formulation via the Lanczos process, $w_1 = r_0/\|r_0\|$,

$$A\,W_n \;=\; W_n\,T_n \;+\; \delta_{n+1} w_{n+1} e_n^T\,, \quad T_n \;=\; W_n^*(A, r_0)\,A\,W_n(A, r_0)\,,$$

and the CG approximation given by

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

$$A_n \;=\; Q_n\,A\,Q_n \;=\; W_n W_n^*\,A\,W_n W_n^* \;=\; W_n\,T_n\,W_n^*\,,$$

Clearly, the projection process is very highly nonlinear in both A and r_0. This allows for the adaptation to the data apparent from the model reduction and moment matching properties.

3. Model reduction and moment matching

For more details and references:

J. Liesen. and Z.S., *Krylov Subspace Methods, Principles and Analysis.* Oxford University Press (2013), Chapter 3

- Let A be HPD with spectral decomposition $A = Y\Lambda Y^*$, where $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_N$ (distinct eigenvalues for simplicity).
- Suppose $\omega_k = |(v_1, y_k)|^2 > 0$, $k = 1, \ldots, N$, and define the distribution function

$$\omega(\lambda) = \begin{cases} 0, & \text{if } \lambda < \lambda_1, \\ \sum_{k=1}^{\ell} \omega_k, & \text{if } \lambda_\ell \le \lambda < \lambda_{\ell+1}, \text{ for } \ell = 1, \ldots, N-1, \\ 1, & \text{if } \lambda_N \le \lambda. \end{cases}$$

- The moments of $\omega(\lambda)$ are given by

$$\int \lambda^k d\omega(\lambda) = \sum_{\ell=1}^{N} \omega_\ell \{\lambda_\ell\}^k = v_1^* A^k v_1, \quad k = 0, 1, 2, \ldots$$

- Analogous construction applied to $T_n = V_n^* A V_n$ yields a distribution function $\omega^{(n)}(\lambda)$ with moments given by

$$\int \lambda^k d\omega^{(n)}(\lambda) = \sum_{\ell=1}^{n} \omega_\ell^{(n)} \{\lambda_\ell^{(n)}\}^k = e_1^T T_n^k e_1, \quad k = 0, 1, 2, \ldots$$

- Let A be HPD with spectral decomposition $A = Y\Lambda Y^*$, where $0 < \lambda_1 < \lambda_2 < \cdots < \lambda_N$ (distinct eigenvalues for simplicity).
- Suppose $\omega_k = |(v_1, y_k)|^2 > 0$, $k = 1, \ldots, N$, and define the distribution function

$$\omega(\lambda) = \begin{cases} 0, & \text{if } \lambda < \lambda_1, \\ \sum_{k=1}^{\ell} \omega_k, & \text{if } \lambda_\ell \le \lambda < \lambda_{\ell+1}, \text{ for } \ell = 1, \ldots, N-1, \\ 1, & \text{if } \lambda_N \le \lambda. \end{cases}$$

- The moments of $\omega(\lambda)$ are given by

$$\int \lambda^k d\omega(\lambda) = \sum_{\ell=1}^{N} \omega_\ell \{\lambda_\ell\}^k = v_1^* A^k v_1, \quad k = 0, 1, 2, \ldots$$

- Analogous construction applied to $T_n = V_n^* A V_n$ yields a distribution function $\omega^{(n)}(\lambda)$ with moments given by

$$\int \lambda^k d\omega^{(n)}(\lambda) = \sum_{\ell=1}^{n} \omega_\ell^{(n)} \{\lambda_\ell^{(n)}\}^k = e_1^T T_n^k e_1, \quad k = 0, 1, 2, \ldots$$

Let $\phi_0(\lambda) \equiv 1, \phi_1(\lambda), \ldots, \phi_n(\lambda)$ be the first $n+1$ orthonormal polynomials corresponding to the distribution function $\omega(\lambda)$.
Then, writing $\Phi_n(\lambda) = [\phi_0(\lambda), \ldots, \phi_{n-1}(\lambda)]^*$,

$$\lambda \, \Phi_n(\lambda) \; = \; T_n \, \Phi_n(\lambda) \; + \; \delta_{n+1} \, \phi_n(\lambda) \, e_n$$

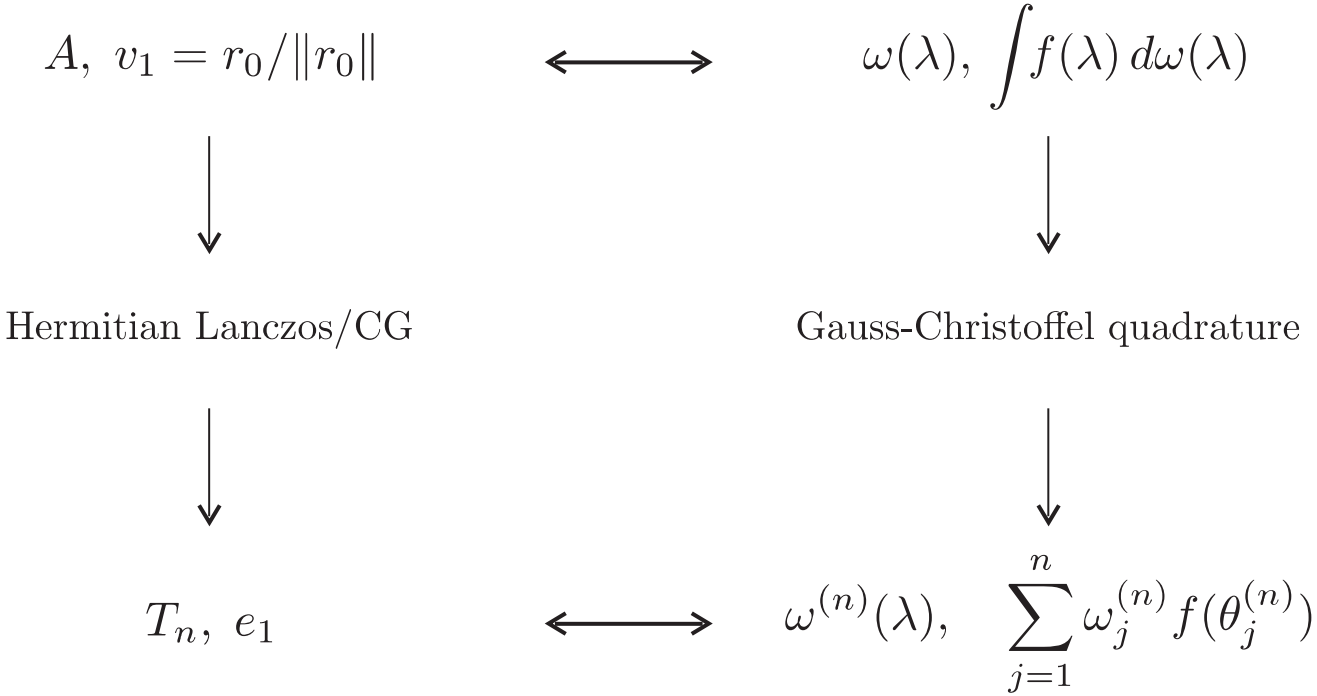
represents the Stieltjes recurrence (1893-4), see Chebyshev (1855), Brouncker (1655), Wallis (1656), Toeplitz and Hellinger (1914) with the Jacobi matrix

$$T_n \equiv \begin{pmatrix} \gamma_1 & \delta_2 & & \\ \delta_2 & \gamma_2 & \ddots & \\ & \ddots & \ddots & \delta_n \\ & & \delta_n & \gamma_n \end{pmatrix} , \quad \delta_l > 0 \, , \ell = 2, \ldots, n \, .$$

3 Fundamental relationship with Gauss quadrature

$$v_1^* A^k v_1 = e_1^T T_n^k e_1 \quad \text{for} \quad k = 0, 1, \ldots, 2n-1.$$

represents the n-node Gauss-Christoffel quadrature.



$$b^*(\lambda I - A)^{-1} b \;=\; \int_L^U \frac{d\omega(\mu)}{\lambda - \mu} \;=\; \sum_{j=1}^{N} \frac{\omega_j}{\lambda - \lambda_j} \;=\; \frac{\mathcal{R}_N(\lambda)}{\mathcal{P}_N(\lambda)}\,,$$

$$\frac{\mathcal{R}_N(\lambda)}{\mathcal{P}_N(\lambda)} \;=\; \mathcal{F}_N(\lambda) \;\equiv\; \cfrac{1}{\lambda - \gamma_1 - \cfrac{\delta_2^2}{\lambda - \gamma_2 - \cfrac{\delta_3^2}{\lambda - \gamma_3 - \ldots \cfrac{\ddots}{\lambda - \gamma_{N-1} - \cfrac{\delta_N^2}{\lambda - \gamma_N}}}}}$$

The denominator $\mathcal{P}_n(\lambda)$ corresponding to the nth convergent $\mathcal{F}_n(\lambda)$ of $\mathcal{F}_N(\lambda)$, $n = 1, 2, \ldots$ is the nth orthogonal polynomial in the sequence determined by $\omega(\lambda)$; see Chebyshev (1855).

Let $\mathcal{B}$ be a bounded linear operator on Hilbert space V. Given z_0, form a sequence $z_0,\ z_1 = \mathcal{B}z_0,\ z_2 = \mathcal{B}z_1 = \mathcal{B}^2 z_0,\ \ldots,\ z_n = \mathcal{B}z_{n-1} = \mathcal{B}^n z_{n-1},\ \ldots$

Let $z_1, \ldots, z_n$ be linearly independent. Determine a sequence of operators $\mathcal{B}_n$ defined on the nested subspaces V_n generated by $z_0, z_1, \ldots, z_{n-1}$,

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

where $E_n z_n$ is the projection (orthogonal or oblique) from V onto V_n.

3 Model reduction using Krylov subspaces Vorobyev (1958, 1965)

Using the projection E_n onto V_n we can write for the operators constructed above (for the method of moments see Vorobyev (1958, 1965))

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

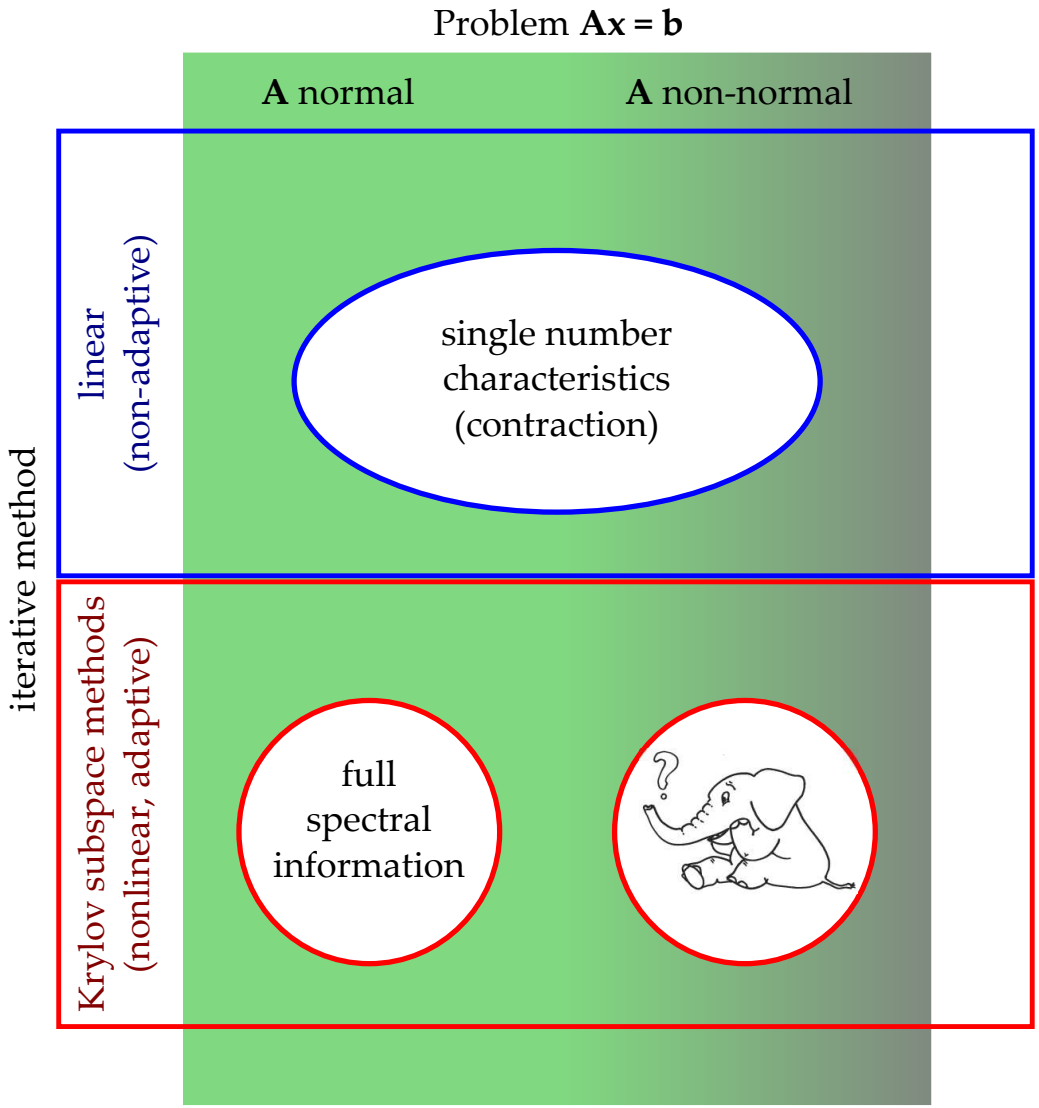
For finite non-Hermitian matrices:

- The first n steps of non-Hermitian Lanczos give complex generalization of the Gauss quadrature matching the first $2n$ moments. Auxiliary subspaces, oblique projections.
- The first n steps of Arnoldi match only the first n moments. No auxiliary subspaces, orthogonal projections.

4. Convergence and spectral information

For more details and references:

- J. Liesen. and Z.S., *Krylov Subspace Methods, Principles and Analysis.* Oxford University Press (2013), Chapter 5, Sections 5.1 - 5.7
- T. Gergelits and Z.S., *Composite convergence bounds based on Chebyshev polynomials and finite precision conjugate gradient computations*, Numer. Alg. 65, 759-782 (2014)


Problem **Ax = b**
A normal
A non-normal
iterative method
linear
(non-adaptive)
single number
characteristics
(contraction)
Krylov subspace methods
(nonlinear, adaptive)
full
spectral
information

Stationary iterative methods, $A = K - L$

$$x - x_n = M^n (x - x_0), \quad M = K^{-1}L, \quad M = WJW^{-1},$$
$$W^{-1}(x - x_n) = J^n (W^{-1}(x - x_0)).$$

Stationary Richardson (assume A HPD)

$$x - x_n = (I - \omega^{-1} A)^n (x - x_0)$$

Chebyshev semiiterative method

$$x - x_n = \frac{1}{|\chi_n(0)|} \chi_n(A)(x - x_0), \quad \frac{1}{|\chi_n(0)|} \leq 2 \left(\frac{\sqrt{\kappa(A)} - 1}{\sqrt{\kappa(A)} + 1} \right)^n;$$

$$\|\chi_n(A)\| = \max_{\lambda_j} |\chi_k(\lambda_j)| \leq \max_{\lambda \in [\lambda_1, \lambda_N]} |\chi_k(\lambda)| = 1 .$$

Why can CG do substantially better?

4 Conjugate Gradient method (CG) for $Ax = b$ with A HPD (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 $\|x - x_n\|_A$ 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

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

- The orthogonality condition leads to short recurrences due to the relationship to the orthogonal polynomials that define the algebraic residuals and search vectors.

Inexact computations?

4 Lanczos (conjugate gradients) as the nonlinear moment problem

Let $\mathcal{G}$ be a linear bounded self-adjoint operator on a Hilbert space V, $f \in V$, $\|f\| = 1$. Consider the $2n$ real numbers

$$m_j = (\mathcal{G}^j f, f) = \int \lambda^j \, d\omega(\lambda), \quad j = 0, \ldots, 2n-1.$$

The method of conjugate gradients (assuming, in addition, the coercivity of $\mathcal{G}$), as well as the Lanczos method for approximating eigenvalues, solve the $2n$ equations

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

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

- Golub, Welsch (1968), Gordon (1968), ...
- An overview and generalization to quasi-definite linear functionals and complex Gauss quadrature is given in Pozza, Pranic, S (2016), relationship with the nonsymmetric Lanczos algorithm in Pozza, Pranic, S (2017).

4 Jacobi matrices gives the fundamental mathematical structure

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

is the Jacobi matrix of the Lanczos process coefficients at step n.

Whenever the bottom element of a normalized eigenvector of T_n vanishes, the associated eigenvalue of T_n closely approximates an eigenvalue of A and an analogous approximation must exist for T_{n+1}, T_{n+2} etc.

The notion of *"deflation"*.

4 CG, large outliers and "effective condition numbers"

Desired accuracy ϵ, $\kappa_s(A) \equiv \lambda_{N-s}/\lambda_1$. Then in exact arithmetic

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

CG steps will produce the approximate solution x_n satisfying

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

Assuming exact arithmetic, this statement qualitatively explains superlinear convergence of CG at the presence of large outliers in the spectrum.

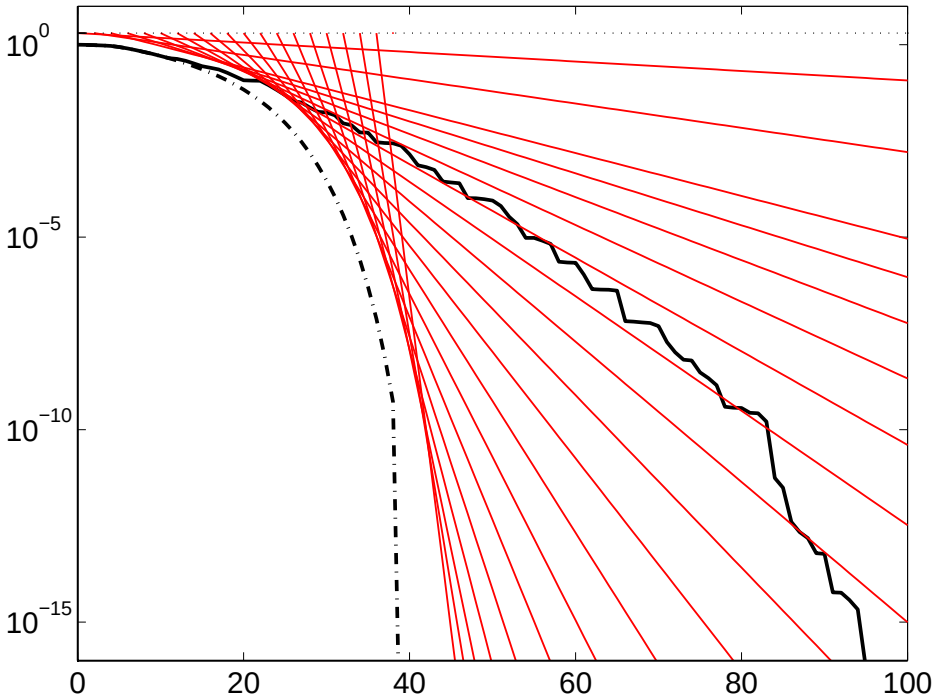
The assumption is crucial. Since CG uses short recurrences, the argument using "effective condition number" is completely void for non-trivial applications.

- We no longer have Krylov subspaces defined by the input data.
- Computed residuals are not orthogonal to the generated subspaces, i.e., the Galerkin orthogonality does not hold.

The structure of Krylov subspace methods as projection processes onto nested subspaces of increasing dimensionality seems to be completely lost.

Is anything preserved?

4 Adaptive Chebyshev bound? Liesen, S (2013)



The difference between the dash-dotted and the solid line?

$$\int \lambda^j \, d\omega(\lambda) \;\rightarrow\; \int \lambda^j \, d\omega_{1-n}(\lambda)$$

$$x - x_n \;=\; p_n(A)\,(x - x_0) \;=\; W p_n(J)\,W^{-1}\,(x - x_0)$$

does not offer an insight unless W is close to unitary and the spectrum of A has some particular structure. Separating the operator from the initial error and using Cauchy integral representation

$$p_n(A) \;=\; \frac{1}{2\pi i}\int_\Gamma p(\lambda)\,(\lambda - A)^{-1}\,d\lambda\,.$$

we get with denoting $\mathcal{L}(\Gamma)$ the length of the integrating curve

$$\|p_n(A)\| \;=\; \frac{\mathcal{L}(\Gamma)}{2\pi}\,\max_{\lambda\in\Gamma}\|(\lambda - A)^{-1}\|\,\max_{\lambda\in\Gamma}|p(\lambda)|\,.$$

Considering the curves Γ_ϵ on which the resolvent norm is constant $\|(\lambda - A)^{-1}\| = 1/\epsilon$, i.e., the boundaries of the ϵ-pseudospectra of A ,

$$\|p_n(A)\| \;=\; \frac{\mathcal{L}(\Gamma_\epsilon)}{2\pi\epsilon}\,\max_{\lambda\in\Gamma_\epsilon}|p(\lambda)|\,.$$

Assuming that A is diagonalizable, $A = W\Lambda W^{-1}$, $\kappa(W) = \|W\|\|W^{-1}\|$,

$$\frac{\|r_n\|}{\|r_0\|} \leq \kappa(W) \min_{p\in\mathcal{P}_n} \max_{i=1,\ldots,n} |p(\lambda_i)| .$$

The pseudospectra-based bound

$$\frac{\|r_n\|}{\|r_0\|} \leq \frac{\mathcal{L}(\Gamma_\epsilon)}{2\pi\epsilon} \min_{p\in\mathcal{P}_n} \max_{\lambda\in\Gamma_\epsilon} |p(\lambda)| .$$

None of the bounds is universally descriptive. Different eigenvalues with close invariant subspaces make big troubles.

4 Spectral information alone is not sufficient

Given any spectrum and any sequence of the nonincreasing residual norms, a complete parametrization is known of the set of all GMRES associated matrices and right-hand sides.

The set of problems for which the distribution of eigenvalues alone does not correspond to convergence behavior is not of measure zero and it is not pathological.

- Widespread eigenvalues alone can not be identified with poor convergence.
- Clustered eigenvalues alone can not be identified with fast convergence.

Here the link between the matrix and the right-hand side is crucial.

Theorem

1° The spectrum of A is given by $\{\lambda_1, \ldots, \lambda_N\}$ and GMRES(A, b) yields residuals with the prescribed nonincreasing sequence $(x_0 = 0)$

$$\|r_0\| \geq \|r_1\| \geq \cdots \geq \|r_{N-1}\| > \|r_N\| = 0\,.$$

2° Let C be the spectral companion matrix, $h = (h_1, \ldots, h_N)^T$, $h_i^2 = \|r_{i-1}\|^2 - \|r_i\|^2$, $i = 1, \ldots, N$. Let R be a nonsingular upper triangular matrix such that $Rs = h$ with s being the first column of C^{-1}, and let W be a unitary matrix. Then

$$A = W R C R^{-1} W^* \quad \text{and} \quad b = W h\,.$$

Greenbaum, Pták, Arioli and S (1994 - 98); Liesen (1999); Eiermann and Ernst (2001); Meurant (2012); Meurant and Tebbens (2012, 2014); .....

Theorem

Let the spectrum of A be given by $\{\lambda_1, \ldots, \lambda_N\}$ and $\mathrm{GMRES}(A, b)$ yields residuals with the prescribed non-increasing sequence $(x_0 = 0)$

$$\|r_0\| \geq \|r_1\| \geq \cdots \geq \|r_{N-1}\| > \|r_N\| = 0\,.$$

1° There is always a GMRES-equivalent unitary matrix B.

2° Let zero is out of the field of values of the matrix A.
Then there is always a GMRES-equivalent
Hermitian positive definite matrix B.

Greenbaum and S (1994)

5. Mathematical structure preserved at the presence of numerical errors

5 Numerical behavior of algebraic solvers

In finite precision arithmetic, the characteristics used for linear (nonadaptive) iterative solvers remain (with a slight modification) descriptive:

- J^n for the linear stationary iterative solvers
- $|\chi_n(0)|^{-1}$ for the Chebyshev semiiterative method.

However, the description of numerical behavior of the method of conjugate gradients in finite precision arithmetics is significantly different.

5 Mathematical structure preserved in FP Lanczos/CG computations

- Practical computation generates a sequence of (nested) Jacobi matrices T_n, $n = 1, 2, \ldots$

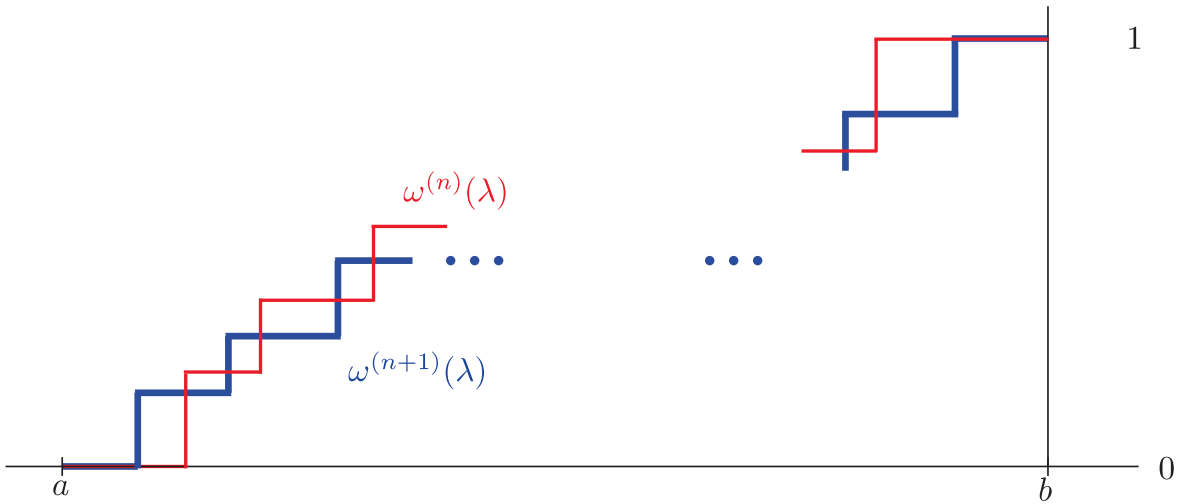
- Whenever the bottom element of a normalized eigenvector of T_n vanishes, the associated eigenvalue of T_n closely approximates an eigenvalue of A and an analogous approximation must exist for T_{n+1}, T_{n+2} etc; see Paige (1971 -1980). This result is highly nontrivial.

The fundamental question: What distribution function is behind this? Greenbaum (1989) gave a beautiful answer. For a given iteration step n the associated distribution function is

$$\omega_{1-n}(\lambda)$$

having clusters of the points of increase whenever the given eigenvalue of A is within the steps 1 to n multiply approximated.

5 Interlocking property for the modified distribution functions



$$\int \lambda^j \, d\omega(\lambda) \;\to\; \int \lambda^j \, d\omega_{1-n}(\lambda) \;\approx\; \int \lambda^j \, d\widehat{\omega}(\lambda)\,.$$

- Mathematical structure preserved for the methods with short recurrences? Complex Jacobi matrices, Gauss quadrature in the complex plane?
- Mathematical structure preserved for Arnoldi, FOM and GMRES? Hessenberg matrices?

6. Discretization and preconditioning

For more details and references:

- J. Málek and Z.S., *Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs.* SIAM Spotlight Series, SIAM (2015)
- I. Pultarová, Z.S., *Decomposition into subspaces and operator preconditioning* (2017?)

6 Basic setting

Consider real (complete) Hilbert space V with the inner product

$$(\cdot,\cdot)_V : V \times V \to \mathbb{R}\,, \quad \|\cdot\|_V\,,$$

dual space $V^\#$ of bounded linear functionals on V with the duality pairing and the associated Riesz map

$$\langle\cdot,\cdot\rangle : V^\# \times V \to \mathbb{R}\,, \quad \tau : V^\# \to V \quad \text{such that} \quad (\tau f, v)_V := \langle f, v\rangle \quad \text{for all } v \in V.$$

Consider a problem

$$\mathcal{A}u \;=\; b$$

with a linear, bounded, coercive, and self-adjoint operator

$$\mathcal{A} : V \to V^\#\,, \quad a(u,v) := \langle \mathcal{A}u, v\rangle\,,$$

$$\begin{aligned} C_{\mathcal{A}} &:= \sup_{v\in V,\,\|v\|_V=1} \|\mathcal{A}v\|_{V^\#} < \infty\,, \\ c_{\mathcal{A}} &:= \inf_{v\in V,\,\|v\|_V=1} \langle \mathcal{A}v, v\rangle > 0\,. \end{aligned}$$

$$M_{\mathcal{A}} := \sup_{u\in V,\, \|u\|_V=1} \langle \mathcal{A}u, u\rangle\,, \qquad m_{\mathcal{A}} := \inf_{u\in V,\, \|u\|_V=1} \langle \mathcal{A}u, u\rangle\,.$$

Theorem.

The spectrum of $\tau\mathcal{A}$ is enclosed in $[m_{\mathcal{A}}, M_{\mathcal{A}}]$. Moreover, $m_{\mathcal{A}}$ and $M_{\mathcal{A}}$ belong to the spectrum, and

$$C_{\mathcal{A}} = \|\mathcal{A}\|_{\mathcal{L}(V,V^\#)} = \sup_{u\in V,\, \|u\|_V=1} \langle \mathcal{A}u, u\rangle = M_{\mathcal{A}}\,,$$

$$c_{\mathcal{A}} = m_{\mathcal{A}} = \frac{1}{\sup_{f\in V^\#,\, \|f\|_{V^\#}=1} \|\mathcal{A}^{-1}f\|_V} = \left\{\|\mathcal{A}^{-1}\|_{\mathcal{L}(V^\#,V)}\right\}^{-1}.$$

6 Operator preconditioning

Linear, bounded, coercive, and self-adjoint $\mathcal{B}$,

$$C_{\mathcal{B}} := \sup_{v \in V,\, \|v\|_V = 1} \|\mathcal{B}v\|_{V^\#} < \infty\,, \qquad c_{\mathcal{B}} := \inf_{v \in V,\, \|v\|_V = 1} \langle \mathcal{B}v, v \rangle\,.$$

Define

$$(\cdot,\cdot)_{\mathcal{B}} : V \times V \to \mathbb{R}, \qquad (w,v)_{\mathcal{B}} := \langle \mathcal{B}w, v \rangle \qquad \text{for all } w, v \in V\,,$$

$$\tau_{\mathcal{B}} : V^\# \to V, \qquad (\tau_{\mathcal{B}} f, v)_{\mathcal{B}} := \langle f, v \rangle \qquad \text{for all } f \in V^\#,\ v \in V\,.$$

Instead of

$$\mathcal{A}u = b$$

we solve

$$\tau_{\mathcal{B}}\, \mathcal{A}\, u = \tau_{\mathcal{B}}\, b$$

or

$$\mathcal{B}^{-1}\, \mathcal{A}\, u = \mathcal{B}^{-1} b.$$

We are interested in the condition number (but recall *Málek, S, 2015, Chapter 11*)

$$\kappa(\mathcal{B}^{-1}\mathcal{A}) := \|\mathcal{B}^{-1}\mathcal{A}\|_V \|\mathcal{A}^{-1}\mathcal{B}\|_V \leq \kappa(\mathcal{A})\kappa(\mathcal{B})$$

and in the spectral number

$$\hat{\kappa}(\mathcal{B}^{-1}\mathcal{A}) := \frac{\sup_{u\in V,\, \|u\|_V=1}(\mathcal{B}^{-1}\mathcal{A}u, u)_V}{\inf_{v\in V,\, \|v\|_V=1}(\mathcal{B}^{-1}\mathcal{A}v, v)_V}\,.$$

Assuming the norm equivalence of $\mathcal{A}^{-1}$ and $\mathcal{B}^{-1}$,

$$\alpha \leq \frac{\|\mathcal{A}^{-1}f\|_V}{\|\mathcal{B}^{-1}f\|_V} \leq \beta \quad \text{for all } f \in V^{\#},\ f \neq 0\,,$$

we get

$$\kappa(\mathcal{B}^{-1}\mathcal{A}) \leq \frac{\beta}{\alpha}\,.$$

N-dimensional subspace $V_h \subset V$; abstract Galerkin discretization gives $u_h \in V_h$, $u_h \approx u \in V$ satisfying Galerkin orthogonality

$$\langle \mathcal{A} u_h - b, v \rangle \;=\; 0 \qquad \text{for all } v \in V_h \,.$$

Considering the restrictions $\mathcal{A}_h : V_h \to V_h^{\#}$, $b_h : V_h \to \mathbb{R}$,

$$\mathcal{A}_h u_h \;=\; b_h, \qquad u_h \in V_h, \quad b_h \in V_h^{\#} \,.$$

With the inner product $(\cdot,\cdot)_{\mathcal{B}}$ and the associated restricted Riesz map

$$\tau_{\mathcal{B},h} : V_h^{\#} \to V_h$$

we get the abstract form of the preconditioned discretized problem

$$\tau_{\mathcal{B},h}\; \mathcal{A}_h \; u_h \;=\; \tau_{\mathcal{B},h}\; b_h \,.$$

Using the discretization basis $\Phi_h = (\phi_1^{(h)}, \ldots, \phi_N^{(h)})$ of V_h
and the canonical dual basis $\Phi_h^{\#} = (\phi_1^{(h)\#}, \ldots, \phi_N^{(h)\#})$ of $V_h^{\#}$, $(\Phi_h^{\#})^* \Phi_h = \mathbf{I}_N$,
we get the linear algebraic system

$$\mathbf{M}_h^{-1}\, \mathbf{A}_h\, \mathbf{x}_h \;=\; \mathbf{M}_h^{-1}\, \mathbf{b}_h,$$

where

$$\begin{aligned}
\mathbf{A}_h,\, \mathbf{M}_h \;&\in\; \mathbb{R}^{N\times N}, \quad \mathbf{x}_h, \mathbf{b}_h \in \mathbb{R}^N, \\
(\mathbf{x}_h)_i \;&=\; \langle \phi_i^{(h)\#}, u_h \rangle, \quad (\mathbf{b}_h)_i \;=\; \langle b, \phi_i^{(h)} \rangle, \\
(\mathbf{A}_h)_{ij} \;&=\; \Big(a(\phi_j^{(h)}, \phi_i^{(h)}) \Big)_{i,j=1,\ldots,N} \;=\; \Big(\langle \mathcal{A}\phi_j^{(h)}, \phi_i^{(h)} \rangle \Big)_{i,j=1,\ldots,N}, \\
(\mathbf{M}_h)_{ij} \;&=\; \Big(\langle \mathcal{B}\phi_j^{(h)}, \phi_i^{(h)} \rangle \Big)_{i,j=1,\ldots,N},
\end{aligned}$$

or

$$\mathbf{A}_h \;=\; (\mathcal{A}\Phi_h)^* \Phi_h, \qquad \mathbf{M}_h \;=\; (\mathcal{B}\Phi_h)^* \Phi_h\,.$$

Indeed, for the restricted Riesz map $\tau_{\mathcal{B},h}$ for $\mathbf{v}$ and $\mathbf{f}$, with $f = \Phi_h^{\#}\mathbf{f}$, $v = \Phi_h\mathbf{v}$

$$(\tau_{\mathcal{B},h} f, v)_{\mathcal{B}} = (\tau_{\mathcal{B},h}\Phi_h^{\#}\mathbf{f}, \Phi_h\mathbf{v})_{\mathcal{B}} = (\Phi_h\mathbf{M}_\tau\mathbf{f}, \Phi_h\mathbf{v})_{\mathcal{B}} = \langle \mathcal{B}\Phi_h\mathbf{M}_\tau\mathbf{f}, \Phi_h\mathbf{v}\rangle = \mathbf{v}^*\mathbf{M}_h\mathbf{M}_\tau\mathbf{f},$$
$$(\tau_{\mathcal{B},h} f, v)_{\mathcal{B}} = \langle f, v\rangle = \mathbf{v}^*\mathbf{f}$$

and therefore

$$\mathbf{M}_\tau \;=\; \mathbf{M}_h^{-1}\,.$$

Using the Cholesky decomposition $\mathbf{M}_h \;=\; \mathbf{L}_h\mathbf{L}_h^*$, the resulting preconditioned algebraic system is transformed into

$$\mathbf{L}_h{}^{-1}\mathbf{A}_h(\mathbf{L}_h^*)^{-1}(\mathbf{L}_h^*\mathbf{x}_h) \;=\; \mathbf{L}_h^{-1}\mathbf{b}_h\;,$$

i.e.,

$$\mathbf{A}_{t,h}\,\mathbf{x}_h^t \;=\; \mathbf{b}_h^t\;.$$

$$\hat{\kappa}(\mathbf{M}_h^{-1}\mathbf{A}_h) \;:=\; \frac{|\lambda_{\max}(\mathbf{M}_h^{-1}\mathbf{A}_h)|}{|\lambda_{\min}(\mathbf{M}_h^{-1}\mathbf{A}_h)|} \;=\; \frac{\max_{u\in V_h,\,\|u\|_{\mathcal{B}}=1}\langle \mathcal{A}u,u\rangle}{\min_{v\in V_h,\,\|v\|_{\mathcal{B}}=1}\langle \mathcal{A}v,v\rangle} \;\le\; \kappa(\mathcal{B})\kappa(\mathcal{A})\,.$$

If the operators $\mathcal{A}$ and $\mathcal{B}$ are spectrally equivalent, i.e.,

$$\alpha \;\le\; \frac{\langle \mathcal{A}w,w\rangle}{\langle \mathcal{B}w,w\rangle} \;\le\; \beta \qquad\qquad \text{for all } \; w\in V,\; w\neq 0\,,$$

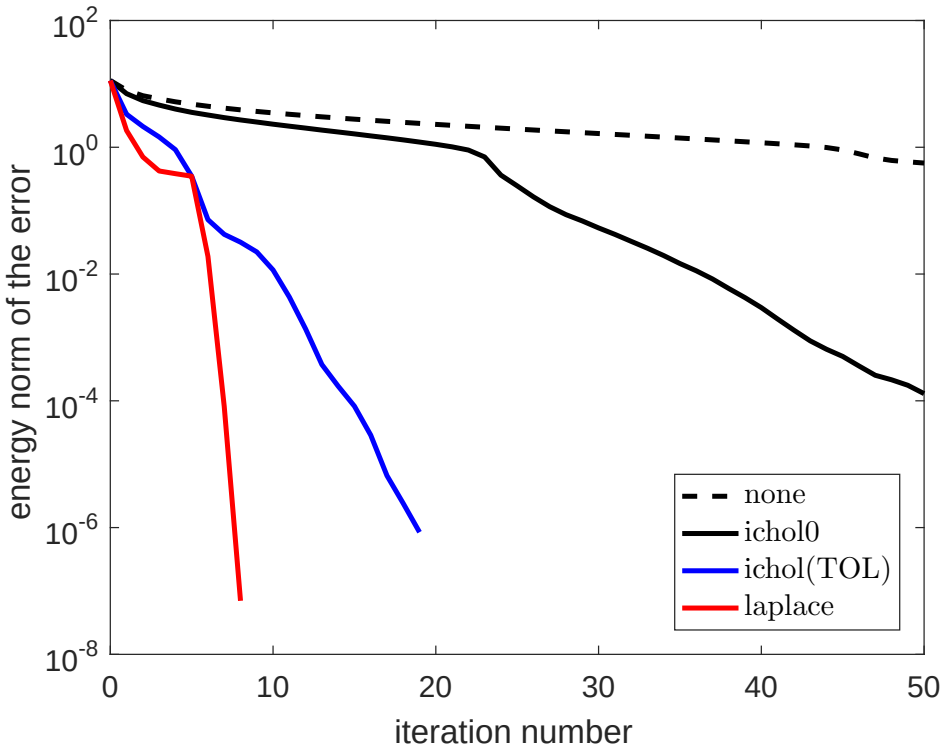
we get

$$\hat{\kappa}(\mathbf{M}_h^{-1}\mathbf{A}_h) \;=\; \kappa(\mathbf{A}_{t,h}) \;\le\; \frac{\beta}{\alpha}\,.$$

$$\kappa(\mathbf{M}_h^{-1}\mathbf{A}_h) \;=\; \|\mathbf{M}_h^{-1}\mathbf{A}_h\|_V\|\mathbf{A}_h^{-1}\mathbf{M}_h\|_V \;\neq\; \frac{|\lambda_{\max}(\mathbf{M}_h^{-1}\mathbf{A}_h)|}{|\lambda_{\min}(\mathbf{M}_h^{-1}\mathbf{A}_h)|} \;=\; \hat{\kappa}(\mathbf{M}_h^{-1}\mathbf{A}_h)\,.$$

Equality holds iff $\mathbf{M}_h$ and $\mathbf{A}_h$ commute (then $\mathbf{M}_h^{-1}\mathbf{A}_h = \mathbf{M}_h^{-1/2}\mathbf{A}_h\mathbf{M}_h^{-1/2}$).

6 Does better conditioning mean faster convergence?



Nonhomogeneous diffusion tensor, uniform mesh. Unpreconditioned CG; ichol PCG (no fill-in); ichol PCG (drop-off tolerance 1e-02); Laplace operator PCG. Condition numbers of $\mathbf{A}_{t,h}$: 6.75e04, 4.31e02, 1.6e01, 1.61e02.

6 Preconditioning as transformation of the discretization basis

Transformation of the discretization basis

$$\Phi_h \to \Phi_{t,h} \quad \text{such that} \quad (\mathcal{B}\Phi_{t,h})^* \Phi_{t,h} = \mathbf{I}$$

i.e. orthogonalization of the basis with respect to the inner product $(\cdot,\cdot)_{\mathcal{B}}$, transforms (preconditions) the system $\mathbf{A}_h\, \mathbf{x}_h \;=\; \mathbf{b}_h$.

For example, the transformed basis

$$\Phi_{t,h} = \Phi_h (\mathbf{L}_h^*)^{-1}, \quad \Phi_{t,h}^{\#} = \Phi_h^{\#} \mathbf{L}_h$$

gives

$$\mathbf{A}_{t,h}\, \mathbf{x}_h^t \;=\; \mathbf{b}_h^t \;.$$

6 Analytic results on preconditioning non-self-adjoint problem

???

Analytic results concerning convergence behaviour for non-Hermitian algebraic problems will be highly appreciated. They are prerequisite for developing an analytic theory of preconditioning.

7. Myths about Krylov subspace methods

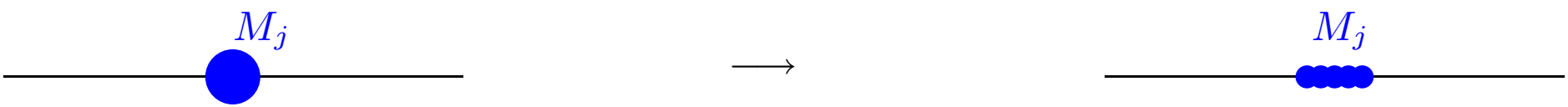
Myth: A belief given uncritical acceptance by the members of a group especially in support of existing or traditional practices and institutions.

Webster's Third New International Dictionary, Enc. Britannica Inc., Chicago (1986)

7 Comments concerning widespread myths

1. Minimal polynomials can rarely be linked to practical computations.
2. Chebyshev semi-iterative method and CG are very different.
3. Clustering of eigenvalues can be misleading even in the SPD case.
4. Infinite dimensional operator-based bounds may not serve in finite dimensional analysis.
5. Finite precision computations can not be seen as minor modifications of the exact arithmetic results.
6. Linearization of nonlinear phenomenon can eliminate the main point that should be analyzed – adaptation to the problem.
7. Short term recurrences can not guarantee well conditioned basis due to rounding errors. This is true even for symmetric positive definite problems, and it remains true also for nonsymmetric problems.
8. Sparsity can have positive as well as negative effects to computations.

7 Clustering of eigenvalues in the SPD case



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).

If it does not, then it means that CG can not adapt to the problem, and it converges almost linearly. **In such cases - is it worth using?**

- It is not true that CG (or other Krylov subspace methods used for solving systems of linear algebraic equations with symmetric matrices) applied to a matrix with t distinct well separated tight clusters of eigenvalues produces in general a large error reduction after t steps; see Sections 5.6.5 and 5.9.1 of Liesen, S (2013). The associated myth has been proved false more than 25 years ago; see Greenbaum (1989); S (1991); Greenbaum, S (1992). Still it is persistently repeated in literature as an obvious fact.

- With no information on the structure of invariant subspaces it can not be claimed that distribution of eigenvalues provides insight into the asymptotic behavior of Krylov subspace methods (such as GMRES) applied to systems with generally nonsymmetric matrices; see Sections 5.7.4, 5.7.6 and 5.11 of Liesen, S (2013). As above, the relevant results Greenbaum, S (1994); Greenbaum, Pták, S (1996) and Arioli, Pták, S (1998) are more than 20 years old.

7 How the mathematical myths are created?

- Rutishauser (1959) as well as Lanczos (1952) considered CG principally different in their nature from the method based on the Chebyshev polynomials.
- Daniel (1967) did not identify the CG convergence with the Chebyshev polynomials-based bound. He carefully writes (modifyling slightly his notation)

  "assuming only that the spectrum of the matrix $\mathbf{A}$ lies inside the interval $[\lambda_1, \lambda_N]$, we can do no better than Theorem 1.2.2."
- That means that the Chebyshev polynomials-based bound holds for any distribution of eigenvalues between λ_1 and λ_1 and for any distribution of the components of the initial residuals in the individual invariant subspaces.
- Why we do not read the original works? They are many times most valuable sources of insight, that can be gradually forgotten and can be overshadowed by commonly accepted myth ...

7 Analogy with a priori and a posteriori numerical PDE analysis

- Think of a priori and a posteriori numerical PDE analysis!
- The Chebyshev bound is a typical a priori bound; it uses no a posteriori information.
- A priori bounds are useful for the purpose they have been derived to. They can not take over the role of the a posteriori bounds.

8 Concluding remarks and outlook

- We wish to understand the behaviour of Krylov subspace methods within the first few steps.
- We are dealing with highly nonlinear finite dimensional phenomena.
- Analytic theory is in the self-adjoint case based on the spectral decomposition of operators.
- In the non-self-adjoint case theoretical results are fragmented and a synthesis is largely missing.
- Beautiful results on the effects of numerical errors,
  but still a lot of work ahead ...
- Historia Magistra Vitae

Thank you very much for your kind patience!