Decomposition into Subspaces and Operator Preconditioning

Zdeněk Strakoš

Charles University, Prague

ICMS Edinburgh, April 11, 2018

Many thanks

to the coauthors

Tomáš Gergelits,
Jakub Hrnčíř,
Josef Málek,
Jan Papež,
Ivana Pultarová,

and to

Barbara Wohlmuth,
Uli Rüde,

for the discussions and references that has motivated this work.

Preconditioning deals with the problem, not with the method

Preconditioning of a linear algebraic system

$$\mathbf{A}\,\mathbf{x} \;=\; \mathbf{b}$$

means its transformation to another system with more favourable properties for its numerical solution. Standard textbook introduction considers $\mathbf{A}$ SPD and takes an SPD matrix $\mathbf{B} \approx \mathbf{A}$ with decomposition $\mathbf{B} = \mathbf{L}\mathbf{L}^*$, giving

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

In order to technically apply an iterative method (CG) to the transformed system, its algorithm is reformulated in terms of the original variables which is better resembled by

$$\mathbf{B}^{-1}\mathbf{A}\,\mathbf{x} \;=\; \mathbf{B}^{-1}\mathbf{b}\,.$$

Reference choice $\mathbf{B} = \mathbf{B}^{1/2}\mathbf{B}^{1/2}$

Given SPD matrix $\mathbf{B}$, this schema will work with any decomposition $\mathbf{B} = \mathbf{L}\mathbf{L}^*$.
For later convenience, consider the special (reference) choice

$$\mathbf{B} = \mathbf{B}^{1/2}\mathbf{B}^{1/2}\,.$$

Then for any other decomposition $\mathbf{B} = \mathbf{L}\mathbf{L}^*$ we have

$$\mathbf{L}^{-1}\mathbf{B}\mathbf{L}^{*\,-1} = (\mathbf{L}^{-1}\mathbf{B}^{1/2})(\mathbf{B}^{1/2}\mathbf{L}^{*\,-1}) = \mathbf{I}\,,$$

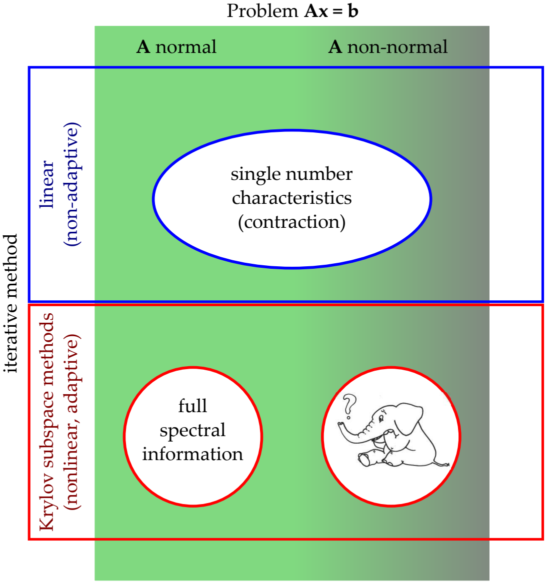
and taking the unitary matrix

$$\mathbf{Q} := \mathbf{L}^{-1}\mathbf{B}^{1/2}\,, \quad \mathbf{Q}^{-1} = \mathbf{Q}^* = \mathbf{B}^{-1/2}\mathbf{L} = \mathbf{B}^{1/2}\mathbf{L}^{*\,-1}\,,$$

we have the unitary transformation from $\mathbf{L}$ to $\mathbf{B}^{1/2}$ and vice versa

$$\mathbf{L} = \mathbf{B}^{1/2}\mathbf{Q}^*\,, \quad \mathbf{B}^{1/2} = \mathbf{L}\mathbf{Q}\,.$$

Which goal should preconditioning target in transforming the problem?



Condition and spectral numbers?

- It can indeed be useful to investigate condition and spectral numbers providing that this is not considered, in general, the end of the story. See Faber, Manteuffel and Parter (1990).

Condition and spectral numbers?

- It can indeed be useful to investigate condition and spectral numbers providing that this is not considered, in general, the end of the story. See Faber, Manteuffel and Parter (1990).

- Rutishauser (1959) as well as Lanczos (1952) considered CG principally different in their nature from the method based on Chebyshev polynomials.

Condition and spectral numbers?

- It can indeed be useful to investigate condition and spectral numbers providing that this is not considered, in general, the end of the story. See Faber, Manteuffel and Parter (1990).

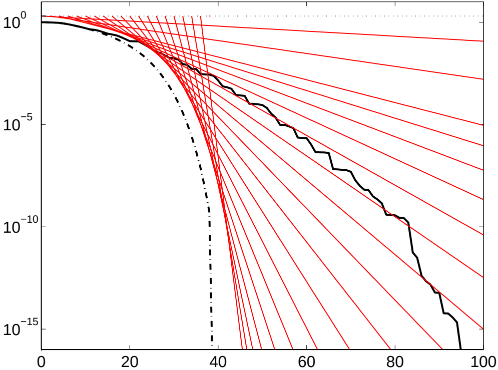
- Rutishauser (1959) as well as Lanczos (1952) considered CG principally different in their nature from the method based on Chebyshev polynomials.

- 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 A lies inside the interval $[\lambda_1, \lambda_N]$, we can do no better than Theorem 1.2.2."

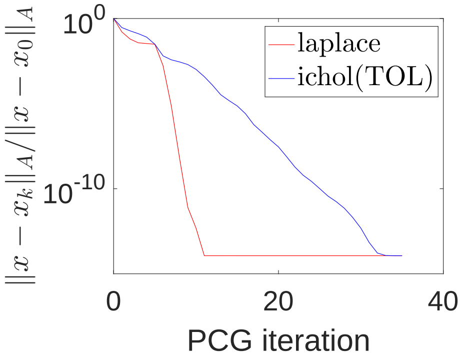
  That means that the Chebyshev polynomials-based bound holds for any distribution of eigenvalues between λ_1 and λ_N and for any distribution of the components of the initial residuals in the individual invariant subspaces.

Adaptive Chebyshev bound principally fails to resolve the matter



The finite precision computation (the thick black line) is not captured quantitatively nor described qualitatively!

Better conditioning does not necessarily mean faster convergence!



Nonhomogeneous diffusion function, uniform mesh.
ICHOLPCG (drop-off tolerance 1e-02); Laplace operator PCG.
Condition numbers of $\mathbf{A}_{t,h}$: 1.6e01, 1.61e02.

Outline

1. Operator preconditioning
2. Discretization
3. Decomposition into subspaces and preconditioning
4. Conclusions

1 Extensive literature related to operator preconditioning

Gunn, D'yakonov, Faber, Manteuffel, Parter, Klawonn, Arnold, Falk, Winther, Axelsson, Karátson, Hiptmair, Vassilevski, Neytcheva, Notay, Elmann, Silvester, Wathen, Zulehner, Simoncini, Oswald, Griebel, Rüde, Steinbach, Wohlmuth, Bramble, Pasciak, Xu, Kraus, Nepomnyaschikh, Dahmen, Kunoth, Yserentant, Mardal, Nordbotten, Rees, Smears, Pearson, ...

Details, proofs and (certainly far from complete) references can be found in

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

1 Basic setting on the Hilbert space V

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

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

Equation in the functional space $V^\#$

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

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

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

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

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

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

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

Instead of the equation in the functional space $V^{\#}$

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

we solve the equation in the solution space V

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

i.e.

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

Theorem (Norm equivalence and condition number)

Assuming that the linear, bounded, coercive and self-adjoint operators $\mathcal{A}$ and $\mathcal{B}$ are $V^{\#}$-*norm equivalent* on V, i.e. there exist $0 < \alpha \leq \beta < \infty$ such that

$$\alpha \leq \frac{\|\mathcal{A}w\|_{V^{\#}}}{\|\mathcal{B}w\|_{V^{\#}}} \leq \beta, \quad \text{for all } \; w \in V,\, w \neq 0\,.$$

Then

$$\kappa(\mathcal{B}^{-1}\mathcal{A}) \;:=\; \|\mathcal{B}^{-1}\mathcal{A}\|_{\mathcal{L}(V,V)}\|\mathcal{A}^{-1}\mathcal{B}\|_{\mathcal{L}(V,V)} \;\leq\; \frac{\beta}{\alpha}\,.$$

Theorem (Spectral equivalence and spectral number)

Assuming that the linear, bounded, coercive and self-adjoint operators $\mathcal{A}$ and $\mathcal{B}$ are *spectrally equivalent* on V, i.e. there exist $0 < \gamma \leq \delta < \infty$ such that

$$\gamma \leq \frac{\langle \mathcal{A}w, w\rangle}{\langle \mathcal{B}w, w\rangle} \leq \delta, \quad \text{for all } w \in V,\, w \neq 0\,.$$

Then

$$\hat{\kappa}(\mathcal{A}, \mathcal{B}) := \frac{\sup_{z\in V,\, \|z\|_V=1} \Big((\tau\mathcal{B})^{-1/2}\tau\mathcal{A}\,(\tau\mathcal{B})^{-1/2}z, z\Big)_V}{\inf_{v\in V,\, \|v\|_V=1} \left((\tau\mathcal{B})^{-1/2}\tau\mathcal{A}\,(\tau\mathcal{B})^{-1/2}v, v\right)_V} \leq \frac{\delta}{\gamma}\,.$$

Outline

1. Operator preconditioning
2. Discretization
3. Decomposition into subspaces and preconditioning
4. Conclusions

Consider N-dimensional subspace $V_h \subset V$ and look for $u_h \in V_h$, $u_h \approx u \in V$ such that

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

Restrictions $\mathcal{A}_h : V_h \to V_h^{\#}$, $b_h : V_h \to \mathbb{R}$ give the problem in $V_h^{\#}$

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

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 in V_h

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

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

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

where

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

or

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

Using (an arbitrary) decomposition $\mathbf{M}_h = \mathbf{L}_h \mathbf{L}_h^*$, the resulting preconditioned algebraic system can be 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.,

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

Consider

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

i.e. orthogonalization of the basis with respect to the inner product $(\cdot,\cdot)_{\mathcal{B}}$. Then

$$\tilde{\Phi}_{t,h} = \Phi_h \mathbf{M_h}^{-1/2}, \quad \tilde{\Phi}^{\#}_{t,h} = \Phi^{\#}_h \mathbf{M_h}^{1/2}$$

gives immediately the preconditioned system $\tilde{\mathbf{A}}_{t,h}\, \tilde{\mathbf{x}}^t_h \;=\; \tilde{\mathbf{b}}^t_h$ corresponding to $\mathbf{L_h} := \mathbf{M_h}^{1/2}$. Any other choice

$$\Phi_{t,h} = \Phi_h \mathbf{L}^{*\,-1}_{\mathbf{h}}, \quad \Phi^{\#}_{t,h} = \Phi^{\#}_h \mathbf{L_h}$$

is given via orthogonal transformation

$$\Phi_{t,h} = \tilde{\Phi}_{t,h} \mathbf{Q}^*\,, \quad \mathbf{Q}^* = \mathbf{M_h}^{1/2} \mathbf{L}^{*\,-1}_{\mathbf{h}}\,, \quad \mathbf{Q}^*\mathbf{Q} = \mathbf{I}\,.$$

2 Points that are worth noticing

- Transformation of the discretization basis (preconditioning) is different from a change of the algebraic basis (similarity transformation).
- Any algebraic preconditioning can be put into the operator preconditioning framework by transformation of the discretization basis and the associated change of the inner product in the infinite dimensional Hilbert space V.

Theorem (Norm equivalence and condition number)

Let the linear, bounded, coercive and self-adjoint operators $\mathcal{A}$ and $\mathcal{B}$ from V to $V^\#$ be $V^\#$-norm equivalent with the lower and upper bounds α and β, respectively, i.e.

$$\alpha \leq \frac{\|\mathcal{A}w\|_{V^\#}}{\|\mathcal{B}w\|_{V^\#}} \leq \beta \quad \text{for all } w \in V\,,\ w \neq 0, \quad 0 < \alpha \leq \beta < \infty\,.$$

Let $\mathbf{S}_h$ be the Gram matrix of the discretization basis $\Phi_h = (\phi_1, \ldots, \phi_N)$ of $V_h \subset V$,

$$(\mathbf{S}_h)_{ij} = (\phi_i, \phi_j)_V\ .$$

Then the condition number of the matrix $\mathbf{M}_h^{-1}\mathbf{A}_h$ is bounded as

$$\kappa(\mathbf{M}_h^{-1}\mathbf{A}_h) := \|\mathbf{M}_h^{-1}\mathbf{A}_h\|\,\|\mathbf{A}_h^{-1}\mathbf{M}_h\| \leq \frac{\beta}{\alpha}\,\kappa(\mathbf{S}_h)\,.$$

Theorem (Spectral equivalence and spectral number)

Let the linear, bounded, coercive and self-adjoint operators $\mathcal{A}$ and $\mathcal{B}$ be spectrally equivalent with the lower and upper bounds γ and δ respectively, i.e.

$$\gamma \leq \frac{\langle \mathcal{A}w, w\rangle}{\langle \mathcal{B}w, w\rangle} \leq \delta \quad \text{for all } w \in V\,, \quad 0 < \gamma \leq \delta < \infty\,.$$

Then the spectral number $\hat{\kappa}(\mathbf{A}_h, \mathbf{M}_h)$, which is equal to the condition number of the matrix $\mathbf{A}_{t,h} = \mathbf{L}_h^{-1}\mathbf{A}_h(\mathbf{L}_h^*)^{-1}$ for any $\mathbf{L}_h$ such that $\mathbf{M}_h = \mathbf{L}_h\mathbf{L}_h^*$, is bounded as

$$\hat{\kappa}(\mathbf{A}_h, \mathbf{M}_h) := \frac{\sup_{\mathbf{z}\in\mathbb{R}^N,\, \|\mathbf{z}\|=1} \left(\mathbf{M}_h^{-1/2}\mathbf{A}_h\mathbf{M}_h^{-1/2}\mathbf{z}, \mathbf{z}\right)}{\inf_{\mathbf{v}\in\mathbb{R}^N,\, \|\mathbf{v}\|=1} \left(\mathbf{M}_h^{-1/2}\mathbf{A}_h\mathbf{M}_h^{-1/2}\mathbf{v}, \mathbf{v}\right)} = \kappa(\mathbf{A}_{t,h}) \leq \frac{\delta}{\gamma}\,.$$

Outline

Decomposition with non-unique representation of elements in V

$$V = \sum_{j \in J} V_j, \quad \text{i.e.,} \quad v = \sum_{j \in J} v_j, \;\; v_j \in V_j, \quad \text{for all } \; v \in V, \; J \; \text{is finite;}$$

Sufficient condition for $V^\# \subset V_j^\#$:

$$c_{V_j} \, \|v\|_V^2 \;\le\; \|v\|_j^2 \quad \text{for all } \; v \in V_j, \;\; 0 < c_{V_j}, \;\; j \in J;$$

Other side inequality:

$$\|v\|_{\mathrm{S}}^2 \;:=\; \inf_{v = \sum_{j \in J} v_j} \left\{ \sum_{j \in J} \|v_j\|_j^2 \right\} \;\le\; C_{\mathrm{S}} \, \|v\|_V^2 \,, \quad \text{for all } \; v \in V \,.$$

Consider local preconditioners

$$\mathcal{B}_j : V_j \to V_j^{\#}, \qquad \langle \mathcal{B}_j w, z \rangle \; = \; \langle \mathcal{B}_j z, w \rangle \qquad \text{for all} \quad w, z \in V_j \,,$$

with $C_{\mathcal{B}_j}$, $c_{\mathcal{B}_j}$ defined as above. Then $\mathcal{B}_j^{-1} : V_j^{\#} \to V_j$, $V^{\#} \subset V_j^{\#}$, and

$$\mathcal{M}^{-1} \; := \; \sum_{j \in J} \mathcal{B}_j^{-1}, \qquad \mathcal{M}^{-1} : V^{\#} \to V$$

gives the global preconditioner. The preconditioned (equivalent?) problem

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

Boundedness and coercivity of $\mathcal{M}^{-1}$

$$\|\mathcal{M}^{-1}\|_{\mathcal{L}(V^\#,V)} = \sup_{f\in V^\#,\, \|f\|_{V^\#}=1} \|\mathcal{M}^{-1} f\|_V \le C_{\mathcal{M}^{-1}} := \sum_{j\in J} \frac{1}{c_{\mathcal{B}_j} c_{V_j}} < \infty\,,$$

$$\inf_{f\in V^\#,\, \|f\|_{V^\#}=1} \langle f, \mathcal{M}^{-1} f\rangle \ge c_{\mathcal{M}^{-1}} := \frac{1}{C_{\mathrm{S}}\, \max_{j\in J} C_{\mathcal{B}_j}} > 0\,,$$

gives equivalence of $\mathcal{A}u = b$ and $\mathcal{M}^{-1}\mathcal{A}u = \mathcal{M}^{-1}b$.

Moreover, we can get norm equivalence and spectral equivalence of $\mathcal{A}$ and $\mathcal{M}$.

Theorem

For any $v \in V \approx u$

$$a\left(\mathcal{M}^{-1}\mathcal{A}\,(v-u),\, v-u\right) \;=\; \sum_{j\in J} \|\bar{r}_j\|^2_{\mathcal{B}_j}\,,$$

$$\frac{\min_{j\in J} c_{\mathcal{B}_j}}{C^2_{\mathcal{A}}}\left(\sum_{k\in J}\frac{1}{c_{V_k}\,c_{\mathcal{B}_k}}\right)^{-1}\sum_{j\in J}\|\bar{r}_j\|^2_j \quad\leq$$

$$\|v-u\|^2_V \quad\leq\quad \frac{C_{\mathrm{S}}\,(\max_{j\in J} C_{\mathcal{B}_j})^2}{c^2_{\mathcal{A}}}\sum_{j\in J}\|\bar{r}_j\|^2_j\,,$$

where $\bar{r}_j := \mathcal{B}_j^{-1}\mathcal{A}\,v - \mathcal{B}_j^{-1}b$ are the locally preconditioned residuals of v.

Theorem

If we consider the stable splitting

$$c_{\mathrm{S}} \|v\|_V^2 \;\leq\; \|v\|_{\mathrm{S}}^2 \;\leq\; C_{\mathrm{S}} \|v\|_V^2 \qquad \text{for all } v \in V,$$

then

$$\frac{c_{\mathcal{A}}}{C_{\mathrm{S}} \max_{j \in J} C_{\mathcal{B}_j}} \;\leq\; \frac{\langle \mathcal{A} v, v\rangle}{\langle \mathcal{M} v, v\rangle} \;\leq\; \frac{C_{\mathcal{A}}}{c_{\mathrm{S}} \min_{j \in J} c_{\mathcal{B}_j}} \quad \text{for all } v \in V,\; v \neq 0,$$

$$\frac{c_{\mathrm{S}} \min_{j \in J} c_{\mathcal{B}_j}}{C_{\mathcal{A}}} \;\leq\; \frac{\|\mathcal{A}^{-1} f\|_V}{\|\mathcal{M}^{-1} f\|_V} \;\leq\; \frac{C_{\mathrm{S}} \max_{j \in J} C_{\mathcal{B}_j}}{c_{\mathcal{A}}} \qquad \text{for all } f \in V^{\#},\; f \neq 0.$$

Outline

1. Operator preconditioning
2. Discretization
3. Decomposition into subspaces and preconditioning
4. Conclusions

- Given framework may help in comparison of existing approaches (work in progress).

- Results guaranteeing fast convergence in practice are based on the subspace splitting and construction of preconditioning that use information on (the inner structure of) the operator $\mathcal{A}$.

- Relationship between the operators $\mathcal{A}$ and $\mathcal{B}$? What can be said about the whole spectrum of the matrix $B^{-1}A$? (Work in progress).

- Adaptation to the problem is the key to efficient solvers.
  Adaptation in many ways!

- $\mathcal{O}(n)$ reliable approximate solvers? A posteriori error analysis leading to efficient and reliable balancing the errors of various origin (including the inaccuracy of algebraic computations).

Theorem

Let $\mathcal{A} : V \to V^{\#}$ be a linear, bounded, coercive and self-adjoint operator. Then its boundedness constant $C_{\mathcal{A}}$ and the coercivity constant $c_{\mathcal{A}}$ can be expressed as

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

$$c_{\mathcal{A}} = \inf_{v \in V,\ \|v\|_V = 1} \langle \mathcal{A}v, v \rangle = \frac{1}{\sup_{f \in V^{\#},\ \|f\|_{V^{\#}} = 1} \|\mathcal{A}^{-1} f\|_V} = \frac{1}{\|\mathcal{A}^{-1}\|_{\mathcal{L}(V^{\#},V)}}. \tag{2}$$

Proof of (1)

Statement (1) follows from

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

where we used the fact that for any self-adjoint operator S in a Hilbert space V

$$\begin{aligned} \|S\|_{\mathcal{L}(V,V)} &= \sup_{z\in V,\, \|z\|_V=1} \|Sz\|_V = \sup_{z\in V,\, \|z\|_V=1} (Sz, Sz)_V^{1/2} \\ &= \sup_{z\in V,\, \|z\|_V=1} |(Sz, z)_V|. \end{aligned}$$

Proof of (2)

$$\frac{1}{\sup_{f\in V^\#,\,\|f\|_{V^\#}=1}\|\mathcal{A}^{-1}f\|_V} = \inf_{v\in V,\,\|v\|_V=1}\|\mathcal{A}v\|_{V^\#} = \inf_{v\in V,\,\|v\|_V=1}\|\tau\mathcal{A}v\|_V$$

We have to prove

$$m_{\mathcal{A}} := \inf_{v\in V,\,\|v\|_V=1}(\tau\mathcal{A}v,v)_V = \inf_{v\in V,\,\|v\|_V=1}\|\tau\mathcal{A}v\|_V.$$

Here "$\leq$" is trivial. We will show that "$<$" leads to a contradiction. Since $m_{\mathcal{A}}$ belongs to the spectrum of $\tau\mathcal{A}$, there exists a sequence $v_1, v_2, \dots \in V$, $\|v_k\|_V = 1$, such that

$$\lim_{k\to\infty}\|\tau\mathcal{A}v_k - m_{\mathcal{A}}v_k\|_V^2 = 0. \tag{3}$$

Assuming

$$m_{\mathcal{A}} < \inf_{v\in V,\,\|v\|_V=1}\|\tau\mathcal{A}v\|_V - \triangle, \quad \triangle > 0,$$

we get

$$\begin{aligned}\|\tau\mathcal{A}v_k - m_{\mathcal{A}}v_k\|_V^2 &= \|\tau\mathcal{A}v_k\|_V^2 + m_{\mathcal{A}}^2 - 2m_{\mathcal{A}}(\tau\mathcal{A}v_k, v_k)_V \\ &\geq \|\tau\mathcal{A}v_k\|_V^2 + m_{\mathcal{A}}^2 - 2m_{\mathcal{A}}\|\tau\mathcal{A}v_k\|_V = (\|\tau\mathcal{A}v_k\|_V - m_{\mathcal{A}})^2 > \triangle^2.\end{aligned}$$

Thank you for your kind patience!