Sparsity, local and global information in numerical solution of PDEs

Zdeněk Strakoš, Josef Málek and Jan Papež

Nečas Center for Mathematical Modeling

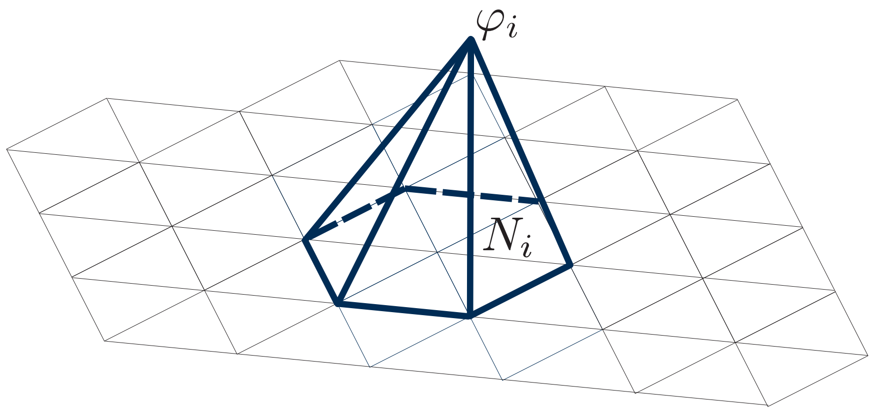
Charles University in Prague and Czech Academy of Sciences

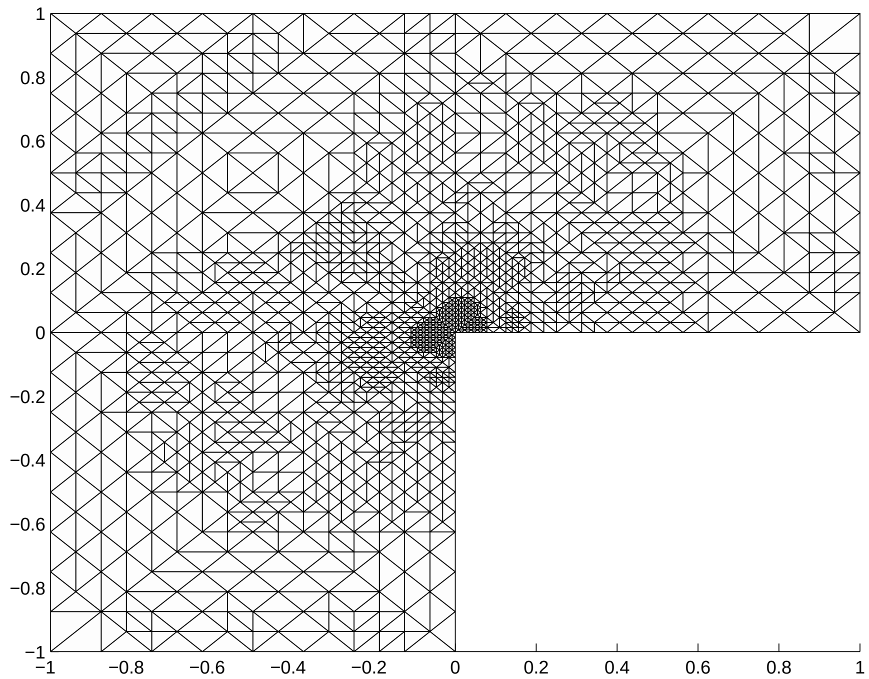
Jörg Liesen

TU Berlin

Banff, BIRS, October 2014

FEM with piecewise polynomial basis functions





Quote: *“piecewise polynomials lead naturally to a sparse stiffness matrix, allowing [the resulting linear algebraic system] to be solved efficiently”*

Local discretisation and global computation

Discrete (piecewise polynomial) FEM approximation $\quad x_h = \Phi_h \, \mathbf{x}_h \, .$

- If the algebraic solution $\mathbf{x}_h$ is known exactly, then the accuracy of approximating the solution x by x_h over the given domain is fully determined by the local approximation property of the piecewise polynomial FEM space.

- However, apart from trivial cases, $\mathbf{x}_h$ that supply the global information is not known exactly, either because of truncation (stopping the iterative process) or because of roundoff or because of both. Then

$$\underbrace{x - x_h^{(n)}}_{\text{total error}} = \underbrace{x - x_h}_{\text{discretisation error}} + \underbrace{x_h - x_h^{(n)}}_{\text{algebraic error}} \, .$$

Two talks

1. Does the locality of the basis functions go hand in hand with efficient solution of the resulting algebraic system? Preconditioning and discretization are linked together.

2. The computation is inexact. Locality of the basis functions allows to express the approximation (discretization) error locally (at least in theory). Algebraic computation takes care for the transfer of the global information across the domain. Expressing algebraic error locally then means being able to perform a componentwise forward error analysis. Do we know how to do it?

Outline

1. Introduction
2. CG in infinite dimensional Hilbert spaces
3. Discretization by restriction to a finite dimensional subspace
4. Algebraic preconditioning and the functional spaces
5. Outlook

Motto (from Thomas):

“First work on the infinite-dimensional setting, then discretize, instead of discretize first.”

1 Basic notation

Let V be a real (infinite dimensional) Hilbert space with the inner product

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

$V^\#$ be the dual space of bounded (continuous) linear functionals on V with the duality pairing

$$\langle \cdot, \cdot \rangle : V^\# \times V \to \mathbb{R} \,.$$

For each $f \in V^\#$ there exists a unique $\tau f \in V$ such that

$$\langle f, v \rangle = (\tau f, v)_V \quad \text{for all } v \in V \,.$$

In this way the inner product $(\cdot, \cdot)_V$ determines the Riesz map

$$\tau : V^\# \to V \,.$$

1 Numerical solution of PDEs

Consider a PDE problem described in the form of the functional equation

$$\mathcal{A}x = b, \quad \mathcal{A} : V \to V^{\#}, \quad x \in V, \quad b \in V^{\#},$$

where the linear, bounded, and coercive operator $\mathcal{A}$ is self-adjoint with respect to the duality pairing $\langle \cdot, \cdot \rangle$.

Standard approach to solving boundary-value problems using the preconditioned conjugate gradient method (PCG) preconditions the algebraic problem,

$$\mathcal{A}, \langle b, \cdot \rangle \to \mathbf{A}, \mathbf{b} \to \text{preconditioning} \to \text{PCG with } \mathbf{A}\mathbf{x} = \mathbf{b},$$

i.e., discretization and preconditioning are often considered separately.

1 Galerkin discretization and preconditioning

Finite dimensional solution subspace $V_h \subset V$. The restriction to V_h gives the approximation $x_h \in V_h$ to $x \in V$,

$$\langle \mathcal{A} x_h, v \rangle = \langle b, v \rangle \qquad \text{for all} \quad v \in V_h \, .$$

With the basis $\Phi_h = (\phi_1^{(h)}, \ldots, \phi_N^{(h)})$ of V_h, the discretization gives the algebraic system

$$\mathbf{A}_h \mathbf{x}_h = \mathbf{b}_h \quad \text{with} \quad \mathbf{x}_h = \mathbf{A}_h^{-1} \mathbf{b}_h$$

and the algebraic preconditioning (PCG) is derived using the transformed algebraic problem using some matrix preconditioner $\widehat{\mathbf{M}} = \widehat{\mathbf{L}} \widehat{\mathbf{L}}^*$,

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

1 Sparsity, iterations and preconditioning

$$\mathbf{A}_h^{-1} \quad \text{in} \quad \mathbf{x}_h = \mathbf{A}_h^{-1}\mathbf{b}_h$$

is not sparse and the individual components of $\mathbf{x}_h$ substantially depend, in general, on all or most components of $\mathbf{b}_h$.

Using the local information transfers in

$$\mathbf{A}_h\mathbf{b}_h, \; \mathbf{A}_h(\mathbf{A}_h\mathbf{b}_h), \; \mathbf{A}_h(\mathbf{A}_h(\mathbf{A}_h\mathbf{b}_h)), \; \ldots$$

it takes for a sparse $\mathbf{A}_h$ many iterative steps to approximate $\mathbf{x}_h$.

Preconditioning is needed in order to cure the computational inefficiency to which the FEM discretization has contributed.

Outline

2. CG in infinite dimensional Hilbert spaces
3. Discretization by restriction to a finite dimensional subspace
4. Algebraic preconditioning and the functional spaces
5. Outlook

2 Krylov subspaces in Hilbert spaces

Using the Riesz map, $\tau\mathcal{A} : V \rightarrow V$. One can form for $g \in V$ the Krylov sequence

$$g,\ \tau\mathcal{A}g\,,\ (\tau\mathcal{A})^2 g\,,\ \ldots \qquad \text{in } V$$

and define Krylov subspace methods in the Hilbert space operator setting (here CG) such that with $r_0 = b - \mathcal{A}x_0 \in V^{\#}$ the approximations x_n to the solution x , $n = 1, 2, \ldots$ belong to the Krylov manifolds in V

$$x_n \in x_0 + K_n(\tau\mathcal{A}, \tau r_0) \equiv$$

$$x_0 + \text{span}\{\tau r_0, \tau\mathcal{A}(\tau r_0), (\tau\mathcal{A})^2(\tau r_0), \ldots, (\tau\mathcal{A})^{n-1}(\tau r_0)\}\,.$$

Approximating the solution $x = (\tau\mathcal{A})^{-1}\tau b$ using Krylov subspaces is not the same as approximating the operator inverse $(\tau\mathcal{A})^{-1}$ by the operators $I, \tau\mathcal{A}, (\tau\mathcal{A})^2, \ldots$

2 Minimization of the energy functional

Defining the energy functional

$$J(v) := \frac{1}{2} \langle \mathcal{A}v, v \rangle - \langle b, v \rangle \,, \qquad v \in V,$$

the solution is equivalently given by the condition

$x \in V$ minimizes the functional J over V .

The Galerkin solution (of the discretized problem) then solves

$x_h \in V_h$ minimizes the functional J over V_h .

Observation: Minimization of the energy functional over the sequence of Krylov subspaces defines the iterates of the conjugate gradient method.

2 Self-adjoint $\mathcal{A}$ wrt the duality pairing

The approximate solution x_n minimizing the energy functional J over $x_0 + K_n$ is equivalently expressed as

$$\|x - x_n\|_a = \min_{z \in x_0 + K_n} \|x - z\|_a \,,$$

where $\|z\|_a^2 = a(z, z) = \langle \mathcal{A} z, z \rangle$, or by the Galerkin orthogonality condition

$$\langle b - \mathcal{A} x_n, w \rangle = \langle r_n, w \rangle = 0 \qquad \text{for all } w \in K_n \equiv K_n(\tau \mathcal{A}, \tau r_0) \,.$$

Since K_n is finite dimensional, this provides in a straightforward way the discretization of the problem matching the maximal number, i.e. $2n$, of moments. The first n steps of the infinite dimensional CG in Hilbert spaces can always be expressed using the n by n linear algebraic system with the Jacobi matrix $\mathbf{T}_n$. Structured sparse representation.

2 Operator preconditioned CG in Hilbert spaces

$$r_0 = b - \mathcal{A}x_0 \in V^{\#}, \quad p_0 = \tau r_0 \in V$$

For $n = 1, 2, \ldots, n_{\max}$

$$\alpha_{n-1} = \frac{\langle r_{n-1}, \tau r_{n-1} \rangle}{\langle \mathcal{A}p_{n-1}, p_{n-1} \rangle} = \frac{(\tau r_{n-1}, \tau r_{n-1})_V}{(\tau \mathcal{A}p_{n-1}, p_{n-1})_V}$$

$$x_n = x_{n-1} + \alpha_{n-1} p_{n-1} \, ,$$ stop when the stopping criterion is satisfied

$$r_n = r_{n-1} - \alpha_{n-1} \mathcal{A} p_{n-1}$$

$$\beta_n = \frac{\langle r_n, \tau r_n \rangle}{\langle r_{n-1}, \tau r_{n-1} \rangle} = \frac{(\tau r_n, \tau r_n)_V}{(\tau r_{n-1}, \tau r_{n-1})_V}$$

$$p_n = \tau r_n + \beta_n p_{n-1}$$

End

Hayes (1954); ... ; Glowinski (2003); Axelsson and Karatson (2009);
Mardal and Winther (2011); Günnel, Herzog and Sachs (2013)

2 CG $\equiv$ Gauss-Christoffel quadrature

$$\tau\mathcal{A}, w_1 = \tau r_0/\|\tau r_0\|_V \quad \longleftrightarrow \quad \omega(\lambda), \quad \int f(\lambda)\, d\omega(\lambda)$$

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

$$\mathbf{T}_n, \mathbf{e}_1 \quad \longleftrightarrow \quad \omega^{(n)}(\lambda), \quad \sum_{i=1}^{n} \omega_i^{(n)} f\left(\theta_i^{(n)}\right)$$

Using $f(\lambda) = \lambda^{-1}$ gives

$$\int \lambda^{-1}\, d\omega(\lambda) \;=\; \sum_{i=1}^{n} \omega_i^{(n)} \left(\theta_i^{(n)}\right)^{-1} + \frac{\|x - x_n\|_a^2}{\|\tau r_0\|_V^2}$$

Condition number bounds should always be checked against this CG - Gauss-Christoffel quadrature equivalence.

Outline

3. Discretization by restriction to a finite dimensional subspace

4. Algebraic preconditioning and the functional spaces

5. Outlook

3 Finite dimensional CG and matrix formulation

Let $\Phi_h = (\phi_1^{(h)}, \dots, \phi_N^{(h)})$ be the basis of the subspace $V_h \subset V$,
let $\Phi_h^\# = (\phi_1^{(h)\#}, \dots, \phi_N^{(h)\#})$ be the canonical basis of its dual $V_h^\#$,
(recall $V_h^\# = \mathcal{A} V_h$). Using the coordinates in Φ_h and in $\Phi_h^\#$,

$$\langle f, v \rangle \rightarrow \mathbf{v}^* \mathbf{f} \,,$$

$$(u, v)_V \rightarrow \mathbf{v}^* \mathbf{M} \mathbf{u}, \qquad (\mathbf{M}_{ij}) = ((\phi_j, \phi_i)_V)_{i,j=1,\dots,N} \,,$$

$$\mathcal{A} u \rightarrow \mathbf{A} \mathbf{u} \,, \qquad \mathcal{A} u = \mathcal{A} \Phi_h \mathbf{u} = \Phi_h^\# \mathbf{A} \mathbf{u} \,; \quad (\mathbf{A}_{ij}) = (a(\phi_j, \phi_i))_{i,j=1,\dots,N} \,,$$

$$\tau f \rightarrow \mathbf{M}^{-1} \mathbf{f} \,, \qquad \tau f = \tau \Phi_h^\# \mathbf{f} = \Phi_h \mathbf{M}^{-1} \mathbf{f} \,;$$

we get with $b = \Phi_h^\# \mathbf{b}$, $x_n = \Phi_h \, \mathbf{x}_n$, $p_n = \Phi_h \, \mathbf{p}_n$, $r_n = \Phi_h^\# \, \mathbf{r}_n$

3 Preconditioned algebraic CG

$\mathbf{r}_0 = \mathbf{b} - \mathbf{A}\mathbf{x}_0,$ solve $\mathbf{M}\mathbf{z}_0 = \mathbf{r}_0,\ \mathbf{p}_0 = \mathbf{z}_0$

For $n = 1, \ldots, n_{\max}$

$$\alpha_{n-1} = \frac{\mathbf{z}_{n-1}^* \mathbf{r}_{n-1}}{\mathbf{p}_{n-1}^* \mathbf{A}\mathbf{p}_{n-1}}$$

$$\mathbf{x}_n = \mathbf{x}_{n-1} + \alpha_{n-1}\mathbf{p}_{n-1}\,,$$ stop when the stopping criterion is satisfied

$$\mathbf{r}_n = \mathbf{r}_{n-1} - \alpha_{n-1}\mathbf{A}\mathbf{p}_{n-1}$$

$$M\mathbf{z}_n = \mathbf{r}_n\,,$$ solve for $\mathbf{z}_n$

$$\beta_n = \frac{\mathbf{z}_n^* \mathbf{r}_n}{\mathbf{z}_{n-1}^* \mathbf{r}_{n-1}}$$

$$\mathbf{p}_n = \mathbf{z}_n + \beta_n \mathbf{p}_{n-1}$$

End

3 Observations

- Unpreconditioned CG, i.e. $\mathbf{M} = \mathbf{I}$ corresponds to the basis Φ orthonormal wrt $(\cdot, \cdot)_V$.

- Operator preconditioning on the discrete space can be interpreted as orthogonalization of the discretization basis. Consequence: the resulting matrix is not sparse!

- Interpretation of the algebraic preconditioning with the preconditioner

$$\widehat{\mathbf{M}} = \widehat{\mathbf{L}}\widehat{\mathbf{L}}^*$$

different from the discretized operator preconditioner

$$\mathbf{M} = \mathbf{L}\mathbf{L}^* \ ?$$

Outline

4. Algebraic preconditioning and the functional spaces

5. Outlook

4 Interpretation of the algebraic preconditioning

Consider the matrix formulation of the finite dimensional CG using the transformed discretization bases

$$\widehat{\Phi} = \Phi\,(\widehat{\mathbf{L}}^*)^{-1}, \quad \widehat{\Phi}^{\#} = \Phi^{\#}\,\widehat{\mathbf{L}}\,.$$

together with the change of the inner product in V_h
(recall $(u,v)_V = \mathbf{v}^*\mathbf{M}\mathbf{u}$)

$$(u,v)_{\text{new},V_h} = (\widehat{\Phi}\widehat{\mathbf{u}}, \widehat{\Phi}\widehat{\mathbf{v}})_{\text{new},V_h} := \widehat{\mathbf{v}}^*\widehat{\mathbf{u}} = \mathbf{v}^*\widehat{\mathbf{L}}\widehat{\mathbf{L}}^*\mathbf{u} = \mathbf{v}^*\widehat{\mathbf{M}}\mathbf{u}\,.$$

Then the discretized Hilbert space formulation of CG gives the algebraically preconditioned matrix formulation of CG (in particular, the unpreconditioned CG applied to the algebraically preconditioned discretized system) with the preconditioner $\widehat{\mathbf{M}}$.

4 Summary

Sparsity of matrices of the algebraic systems is always presented as an advantage of the FEM discretizations.

Sparsity means locality of information. In order to solve the problem, we need the appropriate global transfer of information. This can be accomplished., e.g., by the appropriate globally supported basis functions (cf. hierarchical bases preconditioning, DD with coarse space components, multilevel methods, ...). Recall Rüde (2009).

Preconditioning can be interpreted in part as addressing the difficulty related to sparsity (locality of the supports of the basis functions).

Outline

5. Outlook

5 Discretization via Krylov subspaces

- Coarse grid components, inverted dense blocks etc. means handling global information. It might be worth to give a thought to the automatic and general focus on locality of FEM bases.

- Sparse representations in bases with globally supported components.

- What if an approximation to the the n-th Krylov subspace K_n is taken as the finite dimensional subspace $V_h \subset V$ in

$$\{\mathcal{A}, b, \tau\} \;\rightarrow\; \{\tau \mathcal{A}_n \,:\, K_n \rightarrow K_n\} \;\rightarrow\; \text{PCG with } \{\mathbf{A}_h, \mathbf{M}_h\}\ ?$$

- Can the efficiency of algebraic preconditioners be evaluated using algebraic stopping criteria?

References

- J. Málek and Z.S., Preconditioning and the Conjugate Gradient Method in the Context of Solving PDEs. SIAM Spolight Series, SIAM (2014), in print
- J. Liesen and Z.S., Krylov Subspace Methods, Principles and Analysis. Oxford University Press (2013)
- J. Papez, J. Liesen and Z.S., On distribution of the discretization and algebraic error in numerical solution of partial differential equations, LAA 449, 89-114 (2014)
- M. Arioli, J. Liesen, A. Miedlar, and Z.S., Interplay between discretization and algebraic computation in adaptive numerical solution of elliptic PDE problems, GAMM Mitteilungen 36, 102-129 (2013)
- T. Gergelits and Z.S., Composite convergence bounds based on Chebyshev polynomials and finite precision conjugate gradient computations, Numerical Algorithms 65, 759-782 (2013)

Thank you very much for your kind patience!