JÜLICH FORSCHUNGSZENTRUM

Using non-Galerkin coarse grid operators in multigrid methods

General considerations and case studies for circulant matrices

12 September 2008 | Matthias Bolten

Outline

Algebraic theory for non-Galerkin coarse grid operators

- Motivation and classical results
- Application and breakdown of the theory
- Convergence result for the non-Galerkin case

Replacement operators for certain circulant matrices

- Multigrid for circulant matrices
- Replacement of the Galerkin operator
- Numerical examples

Conclusion and outlook

Outline

- Algebraic theory for non-Galerkin coarse grid operators
  - Motivation and classical results
  - Application and breakdown of the theory
  - Convergence result for the non-Galerkin case
- Replacement operators for certain circulant matrices
  - Multigrid for circulant matrices
  - Replacement of the Galerkin operator
  - Numerical examples
- Conclusion and outlook

Motivation

- Multigrid methods are efficient solvers for a broad range of matrices, including matrices with a lot of structure.
- Algebraic theory uses the variational property of the *Galerkin coarse grid operator* given by

$$A_{k-1} = R_k A_k P_k.$$

- Fulfilling the variational property, this choice is optimal, but the *operator complexity* can be very high.
- The use of the Galerkin operator is the basis of the theory of AMG.
- In geometric multigrid rediscretizations of the PDE are used on the coarser levels; these are in general not equivalent to the Galerkin operator.

Motivation

- Multigrid methods are efficient solvers for a broad range of matrices, including matrices with a lot of structure.
- Algebraic theory uses the variational property of the *Galerkin coarse grid operator* given by

$$A_{k-1} = R_k A_k P_k.$$

- Fulfilling the variational property, this choice is optimal, but the *operator complexity* can be very high.
- The use of the Galerkin operator is the basis of the theory of AMG.
- In geometric multigrid rediscretizations of the PDE are used on the coarser levels; these are in general not equivalent to the Galerkin operator.

Motivation

- Multigrid methods are efficient solvers for a broad range of matrices, including matrices with a lot of structure.
- Algebraic theory uses the variational property of the *Galerkin coarse grid operator* given by

$$A_{k-1} = R_k A_k P_k.$$

- Fulfilling the variational property, this choice is optimal, but the *operator complexity* can be very high.
- The use of the Galerkin operator is the basis of the theory of AMG.
- In geometric multigrid rediscretizations of the PDE are used on the coarser levels; these are in general not equivalent to the Galerkin operator.

Motivation

- Multigrid methods are efficient solvers for a broad range of matrices, including matrices with a lot of structure.
- Algebraic theory uses the variational property of the *Galerkin coarse grid operator* given by

$$A_{k-1} = R_k A_k P_k.$$

- Fulfilling the variational property, this choice is optimal, but the *operator complexity* can be very high.
- The use of the Galerkin operator is the basis of the theory of AMG.
- In geometric multigrid rediscretizations of the PDE are used on the coarser levels; these are in general not equivalent to the Galerkin operator.

Motivation

- Multigrid methods are efficient solvers for a broad range of matrices, including matrices with a lot of structure.
- Algebraic theory uses the variational property of the *Galerkin coarse grid operator* given by

$$A_{k-1} = R_k A_k P_k.$$

- Fulfilling the variational property, this choice is optimal, but the *operator complexity* can be very high.
- The use of the Galerkin operator is the basis of the theory of AMG.
- In geometric multigrid rediscretizations of the PDE are used on the coarser levels; these are in general not equivalent to the Galerkin operator.

Simple example

Consider the 5-point discretization of the Laplacian given by

$$\frac{1}{h^2}\begin{bmatrix} & 1 & \\ 1 & -4 & 1 \\ & 1 & \end{bmatrix}.$$

The Galerkin coarse grid operator is given by

$$\frac{1}{4h^2}\begin{bmatrix} \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \\ \frac{1}{8} & -\frac{3}{4} & \frac{1}{8} \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \end{bmatrix}.$$

So we have a 9-point stencil instead of a 5-point stencil and the complexity per unknown is almost twice as large.

Simple example

Consider the 5-point discretization of the Laplacian given by

$$\frac{1}{h^2}\begin{bmatrix} & 1 & \\ 1 & -4 & 1 \\ & 1 & \end{bmatrix}.$$

The Galerkin coarse grid operator is given by

$$\frac{1}{4h^2}\begin{bmatrix} \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \\ \frac{1}{8} & -\frac{3}{4} & \frac{1}{8} \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \end{bmatrix}.$$

So we have a 9-point stencil instead of a 5-point stencil and the complexity per unknown is almost twice as large.

Simple example

Consider the 5-point discretization of the Laplacian given by

$$\frac{1}{h^2}\begin{bmatrix} & 1 & \\ 1 & -4 & 1 \\ & 1 & \end{bmatrix}.$$

The Galerkin coarse grid operator is given by

$$\frac{1}{4h^2}\begin{bmatrix} \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \\ \frac{1}{8} & -\frac{3}{4} & \frac{1}{8} \\ \frac{1}{16} & \frac{1}{8} & \frac{1}{16} \end{bmatrix}.$$

So we have a 9-point stencil instead of a 5-point stencil and the complexity per unknown is almost twice as large.

Definitions

We define the following:

- $n_k \in \mathbb{N}, k = 0, 1, \ldots$ are the system sizes, where n_0 is the size of the coarsest system,
- $A_k \in \mathbb{C}^{n_k \times n_k}, k = 0, 1, \ldots$ are the system matrices, which we expect to be hermitian positive definite,
- $R_k \in \mathbb{C}^{n_{k-1} \times n_k}, k = 1, 2, \ldots$ are the *restriction operators* from level k to level $k-1$,
- $P_k \in \mathbb{C}^{n_k \times n_{k-1}}, k = 1, 2, \ldots$ are the *prolongation operators* from level $k-1$ to level k, we choose $P_k = R_k^H$,
- $T_k = I - P_k A_{k-1}^{-1} R_k A_k, k = 1, 2, \ldots$ is the iteration matrix of the *coarse grid correction*, and
- $S_k, k = 1, 2, \ldots$ is the iteration matrix of an iterative method used as a *smoother*.

Definitions

We define the following:

- $n_k \in \mathbb{N}, k = 0, 1, \ldots$ are the system sizes, where n_0 is the size of the coarsest system,
- $A_k \in \mathbb{C}^{n_k \times n_k}, k = 0, 1, \ldots$ are the system matrices, which we expect to be hermitian positive definite,
- $R_k \in \mathbb{C}^{n_{k-1} \times n_k}, k = 1, 2, \ldots$ are the *restriction operators* from level k to level $k-1$,
- $P_k \in \mathbb{C}^{n_k \times n_{k-1}}, k = 1, 2, \ldots$ are the *prolongation operators* from level $k-1$ to level k, we choose $P_k = R_k^H$,
- $T_k = I - P_k A_{k-1}^{-1} R_k A_k, k = 1, 2, \ldots$ is the iteration matrix of the *coarse grid correction*, and
- $S_k, k = 1, 2, \ldots$ is the iteration matrix of an iterative method used as a *smoother*.

Definitions

We define the following:

- $n_k \in \mathbb{N}, k = 0, 1, \ldots$ are the system sizes, where n_0 is the size of the coarsest system,
- $A_k \in \mathbb{C}^{n_k \times n_k}, k = 0, 1, \ldots$ are the system matrices, which we expect to be hermitian positive definite,
- $R_k \in \mathbb{C}^{n_{k-1} \times n_k}, k = 1, 2, \ldots$ are the *restriction operators* from level k to level $k-1$,
- $P_k \in \mathbb{C}^{n_k \times n_{k-1}}, k = 1, 2, \ldots$ are the *prolongation operators* from level $k-1$ to level k, we choose $P_k = R_k^H$,
- $T_k = I - P_k A_{k-1}^{-1} R_k A_k, k = 1, 2, \ldots$ is the iteration matrix of the *coarse grid correction*, and
- $S_k, k = 1, 2, \ldots$ is the iteration matrix of an iterative method used as a *smoother*.

Definitions

We define the following:

- $n_k \in \mathbb{N}, k = 0, 1, \dots$ are the system sizes, where n_0 is the size of the coarsest system,
- $A_k \in \mathbb{C}^{n_k \times n_k}, k = 0, 1, \dots$ are the system matrices, which we expect to be hermitian positive definite,
- $R_k \in \mathbb{C}^{n_{k-1} \times n_k}, k = 1, 2, \dots$ are the *restriction operators* from level k to level $k-1$,
- $P_k \in \mathbb{C}^{n_k \times n_{k-1}}, k = 1, 2, \dots$ are the *prolongation operators* from level $k-1$ to level k, we choose $P_k = R_k^H$,
- $T_k = I - P_k A_{k-1}^{-1} R_k A_k, k = 1, 2, \dots$ is the iteration matrix of the *coarse grid correction*, and
- $S_k, k = 1, 2, \dots$ is the iteration matrix of an iterative method used as a *smoother*.

Definitions

We define the following:

- $n_k \in \mathbb{N}, k = 0, 1, \dots$ are the system sizes, where n_0 is the size of the coarsest system,
- $A_k \in \mathbb{C}^{n_k \times n_k}, k = 0, 1, \dots$ are the system matrices, which we expect to be hermitian positive definite,
- $R_k \in \mathbb{C}^{n_{k-1} \times n_k}, k = 1, 2, \dots$ are the *restriction operators* from level k to level $k - 1$,
- $P_k \in \mathbb{C}^{n_k \times n_{k-1}}, k = 1, 2, \dots$ are the *prolongation operators* from level $k - 1$ to level k, we choose $P_k = R_k^H$,
- $T_k = I - P_k A_{k-1}^{-1} R_k A_k, k = 1, 2, \dots$ is the iteration matrix of the *coarse grid correction*, and
- $S_k, k = 1, 2, \dots$ is the iteration matrix of an iterative method used as a *smoother*.

JÜLICH FORSCHUNGSZENTRUM

Definitions

We define the following:

- $n_k \in \mathbb{N}, k = 0, 1, \ldots$ are the system sizes, where n_0 is the size of the coarsest system,
- $A_k \in \mathbb{C}^{n_k \times n_k}, k = 0, 1, \ldots$ are the system matrices, which we expect to be hermitian positive definite,
- $R_k \in \mathbb{C}^{n_{k-1} \times n_k}, k = 1, 2, \ldots$ are the *restriction operators* from level k to level $k-1$,
- $P_k \in \mathbb{C}^{n_k \times n_{k-1}}, k = 1, 2, \ldots$ are the *prolongation operators* from level $k-1$ to level k, we choose $P_k = R_k^H$,
- $T_k = I - P_k A_{k-1}^{-1} R_k A_k, k = 1, 2, \ldots$ is the iteration matrix of the *coarse grid correction*, and
- $S_k, k = 1, 2, \ldots$ is the iteration matrix of an iterative method used as a *smoother*.

Smoothing property and approximation property

Smoothing property

An iterative method $\phi_S^{(k)}$ with iteration matrix S_k fulfills the *smoothing property* if there exists an $\alpha > 0$ such that for all $\mathbf{e}_k \in \mathbb{C}^{n_k}$ it holds

$$\|S_k \mathbf{e}_k\|_{A_k}^2 \leq \|\mathbf{e}_k\|_{A_k}^2 - \alpha \|\mathbf{e}_k\|_*^2. \tag{1}$$

Approximation property

Let T_k be the iteration matrix of the coarse grid correction $\phi_{CGC}^{(k)}$. If there exists a β for all $\mathbf{e}_k \in \mathbb{C}^{n_k}$ such that

$$\|T_k \mathbf{e}_k\|_{A_k}^2 \leq \beta \|\mathbf{e}_k\|_*^2, \tag{2}$$

then $\phi_{CGC}^{(k)}$ fulfills the *approximation property*.

Smoothing property and approximation property

Smoothing property

An iterative method $\phi_{\mathrm{S}}^{(k)}$ with iteration matrix S_k fulfills the *smoothing property* if there exists an $\alpha > 0$ such that for all $\mathbf{e}_k \in \mathbb{C}^{n_k}$ it holds

$$\|S_k \mathbf{e}_k\|_{A_k}^2 \le \|\mathbf{e}_k\|_{A_k}^2 - \alpha \|\mathbf{e}_k\|_*^2. \qquad (1)$$

Approximation property

Let T_k be the iteration matrix of the coarse grid correction $\phi_{\mathrm{CGC}}^{(k)}$. If there exists a β for all $\mathbf{e}_k \in \mathbb{C}^{n_k}$ such that

$$\|T_k \mathbf{e}_k\|_{A_k}^2 \le \beta \|\mathbf{e}_k\|_*^2, \qquad (2)$$

then $\phi_{\mathrm{CGC}}^{(k)}$ fulfills the *approximation property*.

V-cycle convergence using Galerkin

Theorem

If in a modified coarse grid correction $\bar{\phi}^{(k)}_{CGC}$ *the defect equation is solved by a linear iterative method*

$$\bar{\phi}^{(k-1)}(\mathbf{x}_{k-1}, \mathbf{b}_{k-1}) = \bar{M}_{k-1}\mathbf{x}_{k-1} + \bar{N}_{k-1}\mathbf{x}_{k-1}$$

with convergence rate $\bar{\eta} := \|I - \bar{N}_{k-1}A_{k-1}\|_{A_{k-1}} < 1$*, then the (post-smoothing) two grid method using* $\bar{\phi}^{(k)}_{CGC}$ *converges with convergence factor of at most* $\max\{\bar{\eta}, \sqrt{1-\delta}\}$*, i.e.*

$$\|S_k^{\nu_2}\bar{T}_k\mathbf{e}_k\|_{A_k} \leq \max\{\bar{\eta}, \sqrt{1-\delta}\}\|\mathbf{e}_k\|_{A_k},$$

where $\delta = \alpha/\beta$ *with* α *and* β *from the smoothing and approximation property.*

Replacement in the two-grid case

- Theorem above makes no assumption on the method used to solve the coarse grid equation.
- As a result, we can solve it using any approximation.
- For our purpose we assume that $\hat{A}_{k-1}, k = 1, 2, \dots$ are hermitian and positive definite approximations to the Galerkin coarse grid operators $A_{k-1} = R_k A_k R_k^H$.
- If we have

  $$\eta := \|I - \hat{A}_{k-1}^{-1} A_{k-1}\|_{A_{k-1}} < 1,$$

  then the theorem is applicable and the two-grid cycle converges.
- Result does *not* directly carry over to the case, where the modified correction equation is solved approximately.

Replacement in the two-grid case

- Theorem above makes no assumption on the method used to solve the coarse grid equation.
- As a result, we can solve it using any approximation.
- For our purpose we assume that $\hat{A}_{k-1}, k = 1, 2, \ldots$ are hermitian and positive definite approximations to the Galerkin coarse grid operators $A_{k-1} = R_k A_k R_k^H$.
- If we have

$$\eta := \| I - \hat{A}_{k-1}^{-1} A_{k-1} \|_{A_{k-1}} < 1,$$

then the theorem is applicable and the two-grid cycle converges.
- Result does *not* directly carry over to the case, where the modified correction equation is solved approximately.

Replacement in the two-grid case

- Theorem above makes no assumption on the method used to solve the coarse grid equation.
- As a result, we can solve it using any approximation.
- For our purpose we assume that $\hat{A}_{k-1}, k = 1, 2, \ldots$ are hermitian and positive definite approximations to the Galerkin coarse grid operators $A_{k-1} = R_k A_k R_k^H$.
- If we have
$$\eta := \|I - \hat{A}_{k-1}^{-1} A_{k-1}\|_{A_{k-1}} < 1,$$
then the theorem is applicable and the two-grid cycle converges.
- Result does *not* directly carry over to the case, where the modified correction equation is solved approximately.

Replacement in the two-grid case

- Theorem above makes no assumption on the method used to solve the coarse grid equation.
- As a result, we can solve it using any approximation.
- For our purpose we assume that $\hat{A}_{k-1}, k = 1, 2, \ldots$ are hermitian and positive definite approximations to the Galerkin coarse grid operators $A_{k-1} = R_k A_k R_k^H$.
- If we have

$$\eta := \|I - \hat{A}_{k-1}^{-1} A_{k-1}\|_{A_{k-1}} < 1,$$

  then the theorem is applicable and the two-grid cycle converges.
- Result does *not* directly carry over to the case, where the modified correction equation is solved approximately.

Replacement in the two-grid case

- Theorem above makes no assumption on the method used to solve the coarse grid equation.
- As a result, we can solve it using any approximation.
- For our purpose we assume that $\hat{A}_{k-1}, k = 1, 2, \ldots$ are hermitian and positive definite approximations to the Galerkin coarse grid operators $A_{k-1} = R_k A_k R_k^H$.
- If we have

$$\eta := \| I - \hat{A}_{k-1}^{-1} A_{k-1} \|_{A_{k-1}} < 1,$$

  then the theorem is applicable and the two-grid cycle converges.
- Result does *not* directly carry over to the case, where the modified correction equation is solved approximately.

Replacement in the multigrid case

We now analyze the modified coarse grid correction

$$\hat{T}_k := I - R_k^H \hat{A}_{k-1}^{-1} R_k A_k .$$

Modified coarse grid equation is solved by iterative method

$$\tilde{\phi}_{k-1}(\mathbf{x}_{k-1}, \mathbf{b}_{k-1}) = \tilde{M}_{k-1}\mathbf{x}_{k-1} + \tilde{N}_{k-1}\mathbf{b}_{k-1} .$$

With zero initial approximation we obtain

$$\tilde{T}_k = I - R_k^H \tilde{N}_k R_k A_k .$$

Furthermore we assume that $\tilde{\eta}$ is the convergence rate of $\tilde{\phi}_{k-1}$ and we define $\tilde{\mathbf{d}}_{k-1} := \tilde{N}_{k-1} R_k A_k \mathbf{e}_k$.

Replacement in the multigrid case

We now analyze the modified coarse grid correction

$$\hat{T}_k := I - R_k^H \hat{A}_{k-1}^{-1} R_k A_k.$$

Modified coarse grid equation is solved by iterative method

$$\tilde{\phi}_{k-1}(\mathbf{x}_{k-1}, \mathbf{b}_{k-1}) = \tilde{M}_{k-1}\mathbf{x}_{k-1} + \tilde{N}_{k-1}\mathbf{b}_{k-1}.$$

With zero initial approximation we obtain

$$\tilde{T}_k = I - R_k^H \tilde{N}_k R_k A_k.$$

Furthermore we assume that $\tilde{\eta}$ is the convergence rate of $\tilde{\phi}_{k-1}$ and we define $\tilde{\mathbf{d}}_{k-1} := \tilde{N}_{k-1} R_k A_k \mathbf{e}_k$.

Replacement in the multigrid case

We now analyze the modified coarse grid correction

$$\hat{T}_k := I - R_k^H \hat{A}_{k-1}^{-1} R_k A_k.$$

Modified coarse grid equation is solved by iterative method

$$\tilde{\phi}_{k-1}(\mathbf{x}_{k-1}, \mathbf{b}_{k-1}) = \tilde{M}_{k-1}\mathbf{x}_{k-1} + \tilde{N}_{k-1}\mathbf{b}_{k-1}.$$

With zero initial approximation we obtain

$$\tilde{T}_k = I - R_k^H \tilde{N}_k R_k A_k.$$

Furthermore we assume that $\tilde{\eta}$ is the convergence rate of $\tilde{\phi}_{k-1}$ and we define $\tilde{\mathbf{d}}_{k-1} := \tilde{N}_{k-1} R_k A_k \mathbf{e}_k$.

Replacement in the multigrid case

We now analyze the modified coarse grid correction

$$\hat{T}_k := I - R_k^H \hat{A}_{k-1}^{-1} R_k A_k.$$

Modified coarse grid equation is solved by iterative method

$$\tilde{\phi}_{k-1}(\mathbf{x}_{k-1}, \mathbf{b}_{k-1}) = \tilde{M}_{k-1}\mathbf{x}_{k-1} + \tilde{N}_{k-1}\mathbf{b}_{k-1}.$$

With zero initial approximation we obtain

$$\tilde{T}_k = I - R_k^H \tilde{N}_k R_k A_k.$$

Furthermore we assume that $\tilde{\eta}$ is the convergence rate of $\tilde{\phi}_{k-1}$ and we define $\tilde{\mathbf{d}}_{k-1} := \tilde{N}_{k-1} R_k A_k \mathbf{e}_k$.

Requirements for V-cycle convergence

Let

$$\hat{A}_{k-1} \geq R_k A_k R_k^H,$$

$$\| \hat{T}_k \tilde{T}_k \cdot \|_* \geq \mu_k \| \hat{T}_k \cdot \|_*$$

and

$$\ker(\hat{T}_k^H A_k \hat{T}_k) \subset \ker((\tilde{T}_k - \hat{T}_k)^H A_k \hat{T}_k + \hat{T}_k^H A_k (\tilde{T}_k - \hat{T}_k)),$$

$$\lambda_k := \max_{\mathbf{e}_k \in \mathbb{C}^{n_k} / \ker(\cdots)} \frac{\langle (\tilde{T}_k - \hat{T}_k)^H A_k \hat{T}_k + \hat{T}_k^H A_k (\tilde{T}_k - \hat{T}_k) \mathbf{e}_k, \mathbf{e}_k \rangle}{\langle \hat{T}_k^H A_k \hat{T}_k \mathbf{e}_k, \mathbf{e}_k \rangle},$$

with $\alpha_k / \hat{\beta}_k > \lambda_k$.

JÜLICH FORSCHUNGSZENTRUM

Convergence of method using replacement operator

Theorem

Let S_k fulfill the smoothing property (1) *and let $\hat{T}_k$ fulfill the approximation property* (2), *i.e.*

$$\|\hat{T}_k \mathbf{e}_k\|_{A_k}^2 \leq \hat{\beta} \|\mathbf{e}_k\|_*^2.$$

Under the assumptions of the previous slide and with the given definitions we have

$$\|S^{\nu_2} \tilde{T} \mathbf{e}_k\|_{A_k} \leq \max \left\{ \tilde{\eta}, \sqrt{(1+\lambda_k) - \hat{\alpha_k}/\hat{\beta}_k} \right\} \|\mathbf{e}_k\|_{A_k}.$$

V-cycle multigrid convergence proof

$$
\begin{aligned}
\|\tilde{T}_k \mathbf{e}_k\|^2_{A_k} &= \|\hat{T}_k \mathbf{e}_k + R_k^H(\hat{\mathbf{d}}_{k-1} - \tilde{\mathbf{d}}_{k-1})\|^2_{A_k} \\
&= \|\hat{T}_k \mathbf{e}^k\|^2_{A_k} + \|R_k^H(\hat{\mathbf{d}}_{k-1} - \tilde{\mathbf{d}}_{k-1})\|^2_{A_k} \\
&\leq \|\hat{T}_k \mathbf{e}_k\|^2_{A_k} + \tilde{\eta}^2 \|R_k^H \hat{\mathbf{d}}_{k-1}\|^2_{A_k}
\end{aligned}
$$

and

$$
\begin{aligned}
\|S_k \tilde{T}_k \mathbf{e}_k\|^2_{A_k} &\leq \|\tilde{T}_k \mathbf{e}_k\|^2_{A_k} - \frac{\alpha}{\hat{\beta}} \|\hat{T}_k \mathbf{e}_k\|^2_{A_k} \\
&= (1 - \tilde{\eta}^2 - \frac{\alpha}{\hat{\beta}}) \|\hat{T}_k \mathbf{e}_k\|^2_{A_k} + \tilde{\eta}^2 \|\mathbf{e}_k\|^2_{A_k} \\
&\leq \max\{1 - \frac{\alpha}{\hat{\beta}}, \tilde{\eta}^2\} \|\mathbf{e}_k\|^2_{A_k}.
\end{aligned}
$$

V-cycle multigrid convergence proof (replacement)

$$
\begin{aligned}
\|\tilde{T}_k \mathbf{e}_k\|_{A_k}^2 &= \|\hat{T}_k \mathbf{e}_k + R_k^H(\hat{\mathbf{d}}_{k-1} - \tilde{\mathbf{d}}_{k-1})\|_{A_k}^2 \\
&\leq (1+\lambda_k)\|\hat{T}_k \mathbf{e}^k\|_{A_k}^2 + \|R_k^H(\hat{\mathbf{d}}_{k-1} - \tilde{\mathbf{d}}_{k-1})\|_{A_k}^2 \\
&\leq (1+\lambda_k)\|\hat{T}_k \mathbf{e}_k\|_{A_k}^2 + \tilde{\eta}^2\|R_k^H \hat{\mathbf{d}}_{k-1}\|_{A_k}^2
\end{aligned}
$$

and

$$
\begin{aligned}
\|S_k \tilde{T}_k \mathbf{e}_k\|_{A_k}^2 &\leq \|\tilde{T}_k \mathbf{e}_k\|_{A_k}^2 - \frac{\alpha}{\hat{\beta}}\|\hat{T}_k \mathbf{e}_k\|_{A_k}^2 \\
&\leq (1+\lambda_k - \tilde{\eta}^2 - \frac{\alpha}{\hat{\beta}})\|\hat{T}_k \mathbf{e}_k\|_{A_k}^2 + \tilde{\eta}^2\|\mathbf{e}_k\|_{A_k}^2 \\
&\leq \max\{1+\lambda_k - \frac{\alpha}{\hat{\beta}}, \tilde{\eta}^2\}\|\mathbf{e}_k\|_{A_k}^2.
\end{aligned}
$$

Outline

Algebraic theory for non-Galerkin coarse grid operators

- Motivation and classical results
- Application and breakdown of the theory
- Convergence result for the non-Galerkin case

Replacement operators for certain circulant matrices

- Multigrid for circulant matrices
- Replacement of the Galerkin operator
- Numerical examples

Conclusion and outlook

Circulant matrices

- Special class of structured matrices
- Described by their *generating symbols*, univariate 2π-periodic functions f
- With Fourier matrix of dimension $n \times n$, i.e.

$$(F_n)_{j,k=0}^{n-1}, (F_n)_{j,k} = \frac{1}{\sqrt{n}} e^{-2\pi i \frac{jk}{n}},$$

  circulant matrix $A \in \mathbb{C}^{n \times n}$ is given by

$$A = \mathcal{A}(f) = F_n \mathrm{diag}\left((f(2\pi j/n))_{j=0}^{n-1} \right) F_n^H.$$

- Extension to multilevel case using tensorial arguments
- Multigrid for circulant matrices well analyzed by Serra Cappizano and Tablino-Possio, extended by Aricó and Donatelli. Based on use of Galerkin operator.

Circulant matrices

- Special class of structured matrices
- Described by their *generating symbols*, univariate 2π-periodic functions f
- With Fourier matrix of dimension $n \times n$, i.e.

$$(F_n)_{j,k=0}^{n-1}, (F_n)_{j,k} = \frac{1}{\sqrt{n}} e^{-2\pi i \frac{jk}{n}},$$

  circulant matrix $A \in \mathbb{C}^{n \times n}$ is given by

$$A = \mathcal{A}(f) = F_n \text{diag}\left((f(2\pi j/n))_{j=0}^{n-1} \right) F_n^H.$$

- Extension to multilevel case using tensorial arguments
- Multigrid for circulant matrices well analyzed by Serra Cappizano and Tablino-Possio, extended by Aricó and Donatelli. Based on use of Galerkin operator.

JÜLICH FORSCHUNGSZENTRUM

Circulant matrices

- Special class of structured matrices
- Described by their *generating symbols*, univariate 2π-periodic functions f
- With Fourier matrix of dimension $n \times n$, i.e.

$$(F_n)_{j,k=0}^{n-1}, (F_n)_{j,k} = \frac{1}{\sqrt{n}} e^{-2\pi i \frac{jk}{n}},$$

  circulant matrix $A \in \mathbb{C}^{n \times n}$ is given by

$$A = \mathcal{A}(f) = F_n \mathrm{diag}\left((f(2\pi j/n))_{j=0}^{n-1} \right) F_n^H.$$

- Extension to multilevel case using tensorial arguments
- Multigrid for circulant matrices well analyzed by Serra Cappizano and Tablino-Possio, extended by Aricó and Donatelli. Based on use of Galerkin operator.

Circulant matrices

- Special class of structured matrices
- Described by their *generating symbols*, univariate 2π-periodic functions f
- With Fourier matrix of dimension $n \times n$, i.e.

  $$(F_n)_{j,k=0}^{n-1}, (F_n)_{j,k} = \frac{1}{\sqrt{n}} e^{-2\pi i \frac{jk}{n}},$$

  circulant matrix $A \in \mathbb{C}^{n \times n}$ is given by

  $$A = \mathcal{A}(f) = F_n \text{diag}\left((f(2\pi j/n))_{j=0}^{n-1} \right) F_n^H.$$
- Extension to multilevel case using tensorial arguments
- Multigrid for circulant matrices well analyzed by Serra Cappizano and Tablino-Possio, extended by Aricó and Donatelli. Based on use of Galerkin operator.

Circulant matrices

- Special class of structured matrices
- Described by their *generating symbols*, univariate 2π-periodic functions f
- With Fourier matrix of dimension $n \times n$, i.e.

$$(F_n)_{j,k=0}^{n-1}, (F_n)_{j,k} = \frac{1}{\sqrt{n}} e^{-2\pi i \frac{jk}{n}},$$

  circulant matrix $A \in \mathbb{C}^{n \times n}$ is given by

$$A = \mathcal{A}(f) = F_n \mathrm{diag}\left((f(2\pi j/n))_{j=0}^{n-1} \right) F_n^H.$$

- Extension to multilevel case using tensorial arguments
- Multigrid for circulant matrices well analyzed by Serra Cappizano and Tablino-Possio, extended by Aricó and Donatelli. Based on use of Galerkin operator.

Replacement of the Galerkin operator

Let f_{k-1} be the nonnegative generating symbol of the Galerkin operator and let $\hat{f}_{k-1}$ be the generating symbol of a replacement.

Assume that the following holds:

- For all $\mathbf{x}$ where $f(\mathbf{x}) \neq 0$, we have $\hat{f}_{k-1} > f_{k-1}$,
- f_{k-1} and $\hat{f}_{k-1}$ only have isolated common zeros of order 2, and
- for any of these zeros $\mathbf{x}_0$ we have $\nabla^2 f_{k-1}(\mathbf{x}_0)$ and $\nabla^2 \hat{f}_{k-1}(\mathbf{x}_0)$ are symmetric and positive definite.

Then we can show the assumptions of the convergence theorem introduced before.

Replacement of the Galerkin operator

Let f_{k-1} be the nonnegative generating symbol of the Galerkin operator and let $\hat{f}_{k-1}$ be the generating symbol of a replacement.

Assume that the following holds:

- For all $\mathbf{x}$ where $f(\mathbf{x}) \neq 0$, we have $\hat{f}_{k-1} > f_{k-1}$,
- f_{k-1} and $\hat{f}_{k-1}$ only have isolated common zeros of order 2, and
- for any of these zeros $\mathbf{x}_0$ we have $\nabla^2 f_{k-1}(\mathbf{x}_0)$ and $\nabla^2 \hat{f}_{k-1}(\mathbf{x}_0)$ are symmetric and positive definite.

Then we can show the assumptions of the convergence theorem introduced before.

Replacement of the Galerkin operator

Let f_{k-1} be the nonnegative generating symbol of the Galerkin operator and let $\hat{f}_{k-1}$ be the generating symbol of a replacement.
Assume that the following holds:

- For all $\mathbf{x}$ where $f(\mathbf{x}) \neq 0$, we have $\hat{f}_{k-1} > f_{k-1}$,
- f_{k-1} and $\hat{f}_{k-1}$ only have isolated common zeros of order 2, and
- for any of these zeros $\mathbf{x}_0$ we have $\nabla^2 f_{k-1}(\mathbf{x}_0)$ and $\nabla^2 \hat{f}_{k-1}(\mathbf{x}_0)$ are symmetric and positive definite.

Then we can show the assumptions of the convergence theorem introduced before.

Replacement of the Galerkin operator

Let f_{k-1} be the nonnegative generating symbol of the Galerkin operator and let $\hat{f}_{k-1}$ be the generating symbol of a replacement.

Assume that the following holds:

- For all $\mathbf{x}$ where $f(\mathbf{x}) \neq 0$, we have $\hat{f}_{k-1} > f_{k-1}$,
- f_{k-1} and $\hat{f}_{k-1}$ only have isolated common zeros of order 2, and
- for any of these zeros $\mathbf{x}_0$ we have $\nabla^2 f_{k-1}(\mathbf{x}_0)$ and $\nabla^2 \hat{f}_{k-1}(\mathbf{x}_0)$ are symmetric and positive definite.

Then we can show the assumptions of the convergence theorem introduced before.

Replacement of the Galerkin operator

Let f_{k-1} be the nonnegative generating symbol of the Galerkin operator and let $\hat{f}_{k-1}$ be the generating symbol of a replacement.

Assume that the following holds:

- For all $\mathbf{x}$ where $f(\mathbf{x}) \neq 0$, we have $\hat{f}_{k-1} > f_{k-1}$,
- f_{k-1} and $\hat{f}_{k-1}$ only have isolated common zeros of order 2, and
- for any of these zeros $\mathbf{x}_0$ we have $\nabla^2 f_{k-1}(\mathbf{x}_0)$ and $\nabla^2 \hat{f}_{k-1}(\mathbf{x}_0)$ are symmetric and positive definite.

Then we can show the assumptions of the convergence theorem introduced before.

Replacement operator for 2-level 9-point stencil

Theory allows definition of replacement schemes.

Consider the 9-point stencil of the 2-level circulant matrix described by generating symbol with single isolated zero at the origin:

$$\begin{bmatrix} c & b & c \\ a & -2(a+b)-4c & a \\ c & b & c \end{bmatrix}.$$

It may be replaced by the stencil

$$\begin{bmatrix} & b+2c & \\ a+2c & -2(a+b)-8c & a+2c \\ & b+2c & \end{bmatrix}.$$

Replacement operator for 2-level 9-point stencil

Theory allows definition of replacement schemes.

Consider the 9-point stencil of the 2-level circulant matrix described by generating symbol with single isolated zero at the origin:

$$\begin{bmatrix} c & b & c \\ a & -2(a+b)-4c & a \\ c & b & c \end{bmatrix}.$$

It may be replaced by the stencil

$$\begin{bmatrix} & b+2c & \\ a+2c & -2(a+b)-8c & a+2c \\ & b+2c & \end{bmatrix}.$$

Replacement operator for 2-level 9-point stencil

Theory allows definition of replacement schemes.

Consider the 9-point stencil of the 2-level circulant matrix described by generating symbol with single isolated zero at the origin:

$$\begin{bmatrix} c & b & c \\ a & -2(a+b)-4c & a \\ c & b & c \end{bmatrix}.$$

It may be replaced by the stencil

$$\begin{bmatrix} & b+2c & \\ a+2c & -2(a+b)-8c & a+2c \\ & b+2c & \end{bmatrix}.$$

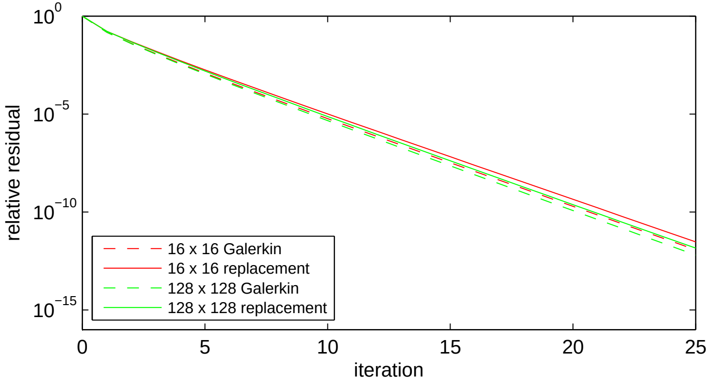
Numerical example: 5-point Laplacian

Operator on finest level:
$$\begin{bmatrix} & -1 & \\ -1 & 4 & -1 \\ & -1 & \end{bmatrix}$$

1. Galerkin operator:
$$\begin{bmatrix} -\frac{1}{64} & -\frac{1}{32} & -\frac{1}{64} \\ -\frac{1}{32} & \frac{3}{16} & -\frac{1}{32} \\ -\frac{1}{64} & -\frac{1}{32} & -\frac{1}{64} \end{bmatrix}$$

1. replacement operator:
$$\begin{bmatrix} & -\frac{1}{16} & \\ -\frac{1}{16} & \frac{1}{4} & -\frac{1}{16} \\ & -\frac{1}{16} & \end{bmatrix}$$

Numerical example: 5-point Laplacian



JÜLICH FORSCHUNGSZENTRUM

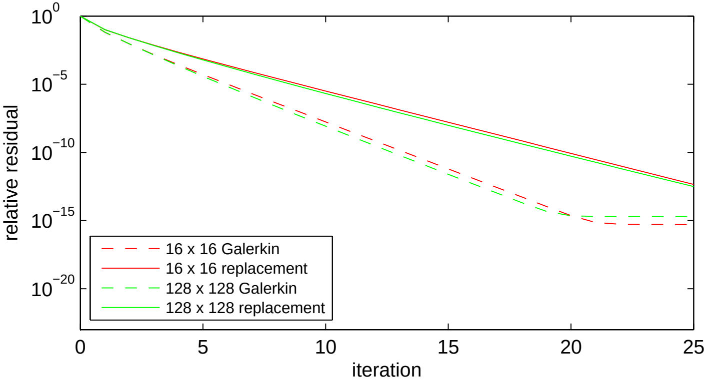
Numerical example: 9-point Laplacian

Operator on finest level: $\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$

1. Galerkin operator: $\begin{bmatrix} -\frac{1}{16} & -\frac{1}{16} & -\frac{1}{16} \\ -\frac{1}{16} & \frac{1}{2} & -\frac{1}{16} \\ -\frac{1}{16} & -\frac{1}{16} & -\frac{1}{16} \end{bmatrix}$

1. replacement operator: $\begin{bmatrix} & -\frac{3}{16} & \\ -\frac{3}{16} & \frac{3}{4} & -\frac{3}{16} \\ & -\frac{3}{16} & \end{bmatrix}$

Numerical example: 9-point Laplacian



Outline

Algebraic theory for non-Galerkin coarse grid operators

- Motivation and classical results
- Application and breakdown of the theory
- Convergence result for the non-Galerkin case

Replacement operators for certain circulant matrices

- Multigrid for circulant matrices
- Replacement of the Galerkin operator
- Numerical examples

Conclusion and outlook

Conclusion and outlook

- Replacement of the Galerkin coarse grid operator in multigrid methods is possible without loosing prerequisites for linear convergence
- The sufficient conditions can be fulfilled for certain circulant matrices
- Replacing the Galerkin operator can reduce compute time
- If necessary, implementation can be adopted to include zeros at other locations
- In the future we will investigate the application of the theory to other classes of matrices

Conclusion and outlook

- Replacement of the Galerkin coarse grid operator in multigrid methods is possible without loosing prerequisites for linear convergence
- The sufficient conditions can be fulfilled for certain circulant matrices
- Replacing the Galerkin operator can reduce compute time
- If necessary, implementation can be adopted to include zeros at other locations
- In the future we will investigate the application of the theory to other classes of matrices

Conclusion and outlook

- Replacement of the Galerkin coarse grid operator in multigrid methods is possible without loosing prerequisites for linear convergence
- The sufficient conditions can be fulfilled for certain circulant matrices
- Replacing the Galerkin operator can reduce compute time
- If necessary, implementation can be adopted to include zeros at other locations
- In the future we will investigate the application of the theory to other classes of matrices

Conclusion and outlook

- Replacement of the Galerkin coarse grid operator in multigrid methods is possible without loosing prerequisites for linear convergence
- The sufficient conditions can be fulfilled for certain circulant matrices
- Replacing the Galerkin operator can reduce compute time
- If necessary, implementation can be adopted to include zeros at other locations
- In the future we will investigate the application of the theory to other classes of matrices

JÜLICH
FORSCHUNGSZENTRUM

Conclusion and outlook

- Replacement of the Galerkin coarse grid operator in multigrid methods is possible without loosing prerequisites for linear convergence
- The sufficient conditions can be fulfilled for certain circulant matrices
- Replacing the Galerkin operator can reduce compute time
- If necessary, implementation can be adopted to include zeros at other locations
- In the future we will investigate the application of the theory to other classes of matrices

Thanks for your attention!