# Convergence Analysis of a Two-Grid Method for Nonsymmetric Positive Definite Problems

Xuefeng Xu

Manuel Batalha

Technische Universität Berlin

Multilevel Methods

# Outline

# Introduction

$$Ax = b$$

where $A \in \mathbb{R}^{n \times n}$ is nonsymmetric but positive definite, namely

$$v^\mathsf{T} A v > 0 \qquad \forall v \in \mathbb{R}^n \setminus \{0\}$$

e.g. discretization of convection-diffusion equations.

# A-splittings

$$A = A_{sym} + A_{skew}$$

where

$$A_{sym} = \frac{1}{2}(A + A^{\mathsf{T}}) \qquad A_{skew} = \frac{1}{2}(A - A^{\mathsf{T}})$$

n.b.

$A_{sym}$ is positive definite iif $A$ is positive definite.

# A-splittings

$$A = M - N$$

$$x_{k+1} = M^{-1} N x_k + M^{-1} b$$

$$= (I - M^{-1} A) x_k + M^{-1} b$$

$$= x_k + M^{-1} (b - A x_k)$$

recall

convergence iif $\rho(I - M^{-1} A) < 1$

# Assumptions

- Let $M \in \mathbb{R}^{n \times n}$ be an SPD smoother such that $\|I - M^{-1}A\|_M \leq 1$
- Let $R \in \mathbb{R}^{n_c \times n}$ be a restriction matrix of rank $n_c$
- Let $P \in \mathbb{R}^{n \times n_c}$ be a prolongation matrix of rank $n_c$
- The coarse-grid matrix $A_c := R\,A\,P \in \mathbb{R}^{n_c \times n_c}$ is nonsingular

<u>aim</u>: find the "ideal" $R$ and $P$

# Remark

$$\|I - M^{-1}A\|_M^2 = \|M^{-\frac{1}{2}}(I - M^{-1}A)M^{\frac{1}{2}}\|_2^2$$

$$= \lambda_{max}((I - M^{-\frac{1}{2}}AM^{\frac{1}{2}})(I - M^{-\frac{1}{2}}A^{\intercal}M^{\frac{1}{2}})) \leq 1$$

that is

$$\lambda_{min}(M^{-\frac{1}{2}}\widetilde{A}M^{\frac{1}{2}}) \geq 0$$

where

$$\widetilde{A} = A + A^{\intercal} - AM^{-1}A^{\intercal}$$

n.b. $\|I - M^{-1}A\|_M \leq 1$ iif $\widetilde{A}$ is positive semidefinite.

# Algorithm: exact two-grid method

1. smoothing : $x^{(1)} \leftarrow x^{(0)} + M^{-1}(b - Ax^{(0)})$
2. restriction : $r_c \leftarrow R(b - Ax^{(1)})$
3. coarse-grid correction : $e_c \leftarrow A_c^{-1}\, r_c$
4. prolongation : $x_{TG} \leftarrow x^{(1)} + P\, e_c$

# Error reduction processes

$$x - x^{(1)} = (I - M^{-1}A)(x - x^{(0)})$$

$$x - x_{TG} = (I - \Pi_A)(x - x^{(1)})$$

where

$$\Pi_A := PA_c^{-1}RA$$

then

$$\|x - x^{(1)}\|_M \leq \|I - M^{-1}A\|_M \|x - x^{(0)}\|_M$$

$$\|x - x_{TG}\|_M \leq \|I - \Pi_A\|_M \|x - x^{(1)}\|_M$$

# Projection $\Pi_A := P\,A_c^{-1}R\,A$

- $\Pi_A$ is a projection along *null*$(RA)$ onto the *range*$(P)$
- For any sub-multiplicative norm: $\|I - \Pi_A\| \geq 1$
- $\|I - \Pi_A\|_M = 1 \iff \Pi_A = M^{-1}\Pi_A^{\intercal}M$
- remark: in SPD problems, the usual choice $P = R^{\intercal}$ yields

$$\|I - \Pi_A\|_A = \|I - R^{\intercal}(RAR^{\intercal})^{-1}RA\|_A = 1$$

# Prolongation $P_*$

It suffices to choose prolongation $P$ such that $\Pi_A = M^{-1}A^\intercal R^\intercal (RAP)^{-\intercal} P^\intercal M$

n.b.

$$range(P) = range(M^{-1}A^\intercal R^\intercal)$$

$$null(RA) = null(P^\intercal M)$$

Satisfied by $P = M^{-1}A^\intercal R^\intercal W$, with $W$ nonsingular.

Taking $W = I$ yields

$$P_* = M^{-1}A^\intercal R^\intercal$$

# Error propagation matrix $E_{TG}$

$$x - x_{TG} = E_{TG}(x - x^{(0)})$$

where

$$E_{TG} = (I - \Pi_A)(I - M^{-1}A)$$

Then

$$\|x - x_{TG}\|_M \leq \|E_{TG}\|_M \|x - x^{(0)}\|_M$$

with

$$\|E_{TG}\|_M = \|(I - M^{\frac{1}{2}}\,\Pi_A\,M^{-\frac{1}{2}})(I - M^{-\frac{1}{2}}\,A\,M^{-\frac{1}{2}})\|_2$$

# A technical lemma ...

## Lemma 3.1

Let $\widetilde{A}$ and $\Pi_A$ be defined as above. Then

$$null(\widetilde{A}) \cap null(RA) = \{0\} \qquad (3.1)$$

if and only if $rank(\widetilde{A}^{\frac{1}{2}}(I - \Pi_A)) = n - n_c$.

remark:

... which provides a sufficient and necessary condition for

$$\|E_{TG}\|_M < 1$$

# $M$-convergence factor

## Theorem 3.3

Under the assumptions of the exact two-grid algorithm, the condition (3.1), and using prolongation $P_*$ , the $M$-convergence factor can be characterized as

$$\|E_{TG}\|_M = \sqrt{1 - \sigma_{TG}}$$

with

$$\sigma_{TG} = \lambda^+_{min}(M^{-1}\widetilde{A}\,(I - \Pi_A))$$

where $\widetilde{A}$ and $\Pi_A$ are defined as above.

# Proof of theorem 3.3

If $P = P_*$, then

$$A_c = RAP_* = P_*^\mathsf{T} M P_*$$

and

$$\Pi_A = P_* A_c^{-1} RA = P_* (P_*^\mathsf{T} M P_*)^{-1} P_*^\mathsf{T} M$$

Let

$$\Pi = M^{\frac{1}{2}} \Pi_A M^{-\frac{1}{2}}$$

Then

$$\Pi = M^{\frac{1}{2}} P_* (P_*^\mathsf{T} M P_*)^{-1} P_*^\mathsf{T} M^{\frac{1}{2}}$$

# Proof of theorem 3.3

Since $\Pi^\intercal = \Pi = \Pi^2$ and $rank(\Pi) = n_c$, there exists an $n \times n$ orthogonal matrix $Q$ such that

$$Q^\intercal \Pi Q = \begin{bmatrix} I_{n_c} & 0 \\ 0 & 0 \end{bmatrix}$$

Recall

$$\|E_{TG}\|_M^2 = \|(I-\Pi)(I - M^{-\frac{1}{2}} A M^{-\frac{1}{2}})\|_2^2$$

$$= \lambda_{max}((I - M^{-\frac{1}{2}} A^\intercal M^{-\frac{1}{2}})(I-\Pi)(I - M^{-\frac{1}{2}} A M^{-\frac{1}{2}}))$$

$$= \lambda_{max}((I - M^{-\frac{1}{2}} A M^{-\frac{1}{2}})(I - M^{-\frac{1}{2}} A^\intercal M^{-\frac{1}{2}})(I-\Pi))$$

$$= \lambda_{max}((I - M^{-\frac{1}{2}} \widetilde{A} M^{-\frac{1}{2}})(I-\Pi))$$

$$= 1 - \lambda_{min}(\Pi - M^{-\frac{1}{2}} \widetilde{A} M^{-\frac{1}{2}}(I-\Pi))$$

# Proof of theorem 3.3

Let

$$Q^{\mathsf{T}} M^{-\frac{1}{2}} \widetilde{A} M^{-\frac{1}{2}} Q = \begin{bmatrix} X_1 & X_2 \\ X_2^{\mathsf{T}} & X_3 \end{bmatrix}$$

where $X_1 \in \mathbb{R}^{n_c \times n_c}$, $X_2 \in \mathbb{R}^{n_c \times (n-n_c)}$, and $X_3 \in \mathbb{R}^{(n-n_c)\times(n-n_c)}$.

The positive semidefiniteness of $Q^{\mathsf{T}} M^{-\frac{1}{2}} \widetilde{A} M^{-\frac{1}{2}} Q$ implies that of $X_3$. From the relation

$$\begin{aligned} I - Q^{\mathsf{T}} M^{-\frac{1}{2}} \widetilde{A} M^{-\frac{1}{2}} Q &= Q^{\mathsf{T}}(I - M^{-\frac{1}{2}} \widetilde{A} M^{-\frac{1}{2}})Q \\ &= Q^{\mathsf{T}}(M^{-\frac{1}{2}} A M^{-\frac{1}{2}})(M^{-\frac{1}{2}} A^{\mathsf{T}} M^{-\frac{1}{2}})Q \end{aligned}$$

we deduce that $I_{(n-n_c)\times(n-n_c)} - X_3$ is SPSD. It follows that $\lambda(X_3) \in [0, 1]$.

# Proof of theorem 3.3

Direct computation yields

$$\Pi + M^{-\frac{1}{2}} \widetilde{A} M^{-\frac{1}{2}} (I - \Pi) = Q \begin{bmatrix} I_{n_c} & X_2 \\ 0 & X_3 \end{bmatrix} Q^T$$

Then

$$\|E_{TG}\|_M^2 = 1 - \lambda_{min}(\Pi + M^{-\frac{1}{2}} \widetilde{A} M^{-\frac{1}{2}} (I - \Pi))$$

$$= 1 - min\{1, \lambda_{min}(X_3)\}$$

$$= 1 - \lambda_{min}(X_3)$$

# Proof of theorem 3.3

Since

$$(I-\Pi)M^{-\frac{1}{2}}\widetilde{A}M^{-\frac{1}{2}}(I-\Pi) = Q\begin{bmatrix} 0 & 0 \\ 0 & X_3 \end{bmatrix}Q^T$$

we obtain

$$rank(X_3) = rank((I-\Pi)M^{-\frac{1}{2}}\widetilde{A}M^{-\frac{1}{2}}(I-\Pi))$$

$$= rank(\widetilde{A}^{\frac{1}{2}}M^{-\frac{1}{2}}(I-\Pi))$$

$$= rank(\widetilde{A}^{\frac{1}{2}}(I-\Pi)) = n - n_c$$

where we have used Lemma 3.1. This means $X_3$ is SPD, i.e. $\lambda_{min}(X_3) > 0$.

# Proof of theorem 3.3

Due to

$$M^{-\frac{1}{2}}\widetilde{A}M^{-\frac{1}{2}}(I-\Pi) = Q\begin{bmatrix} 0 & X_2 \\ 0 & X_3 \end{bmatrix}Q^{\mathsf{T}}$$

it follows that

$$\lambda_{min}(X_3) = \lambda_{min}^{+}(M^{-\frac{1}{2}}\widetilde{A}M^{-\frac{1}{2}}(I-\Pi))$$

Thus,

$$\begin{aligned}\|E_{TG}\|_M^2 &= 1-\lambda_{min}^{+}(M^{-\frac{1}{2}}\widetilde{A}M^{-\frac{1}{2}}(I-\Pi)) \\ &= 1-\lambda_{min}^{+}(M^{-1}\widetilde{A}M^{-\frac{1}{2}}(I-\Pi)M^{\frac{1}{2}}) \\ &= 1-\lambda_{min}^{+}(M^{-1}\widetilde{A}(I-\Pi_A))\end{aligned}$$

$\square$

# Optimal restriction

## Theorem 3.5

Let $\widetilde{A}$ be defined as above, and let $\{(\mu_i, v_i)\}_{i=1}^{n}$ be the eigenpairs of the generalized eigenvalue problem

$$\widetilde{A} = \mu M v$$

, where

$$\mu_1 \leq \mu_2 \leq \ldots \leq \mu_n\,, \qquad v_i^T M v_j = \begin{cases} 1 & \textit{if} \;\; i = j\,, \\ 0 & \textit{if} \;\; i \neq j\,. \end{cases}$$

Under the assumptions of Theorem 3.3, it holds that

$$\|E_{TG}\|_M^2 \geq \sqrt{1 - \mu_{n_c+1}}$$

with equality if $null(RA) = span\{v_{n_c+1}, \ldots, v_n\}$.

# Summary

- Choice of the smoother induced *M*-norm to measure the two-grid convergence

$$\|x - x_{TG}\|_M \leq \|E_{TG}\|_M \|x - x^{(0)}\|_M , \qquad E_{TG} = (I - \Pi_A)(I - M^{-1}A)$$

- Prolongation operator $P_* = M^{-1}A^{\intercal}R^{\intercal}$ such that $\|I - \Pi_A\|_M = 1$
- Lemma 3.1 provides a sufficient and necessary condition for $\|E_{TG}\|_M < 1$
- Theorem 3.3 establishes an identity for characterizing the convergence factor

$$\|E_{TG}\|_M = \sqrt{1 - \sigma_{TG}} , \qquad \sigma_{TG} = \lambda_{min}^{+}(M^{-1}\widetilde{A}(I - \Pi_A))$$

- Analysis on the influence of $range(R^{\intercal})$ on $\|E_{TG}\|_M$, based on the generalized eigenvalue problem

$$\widetilde{A} = \mu M v$$

# Questions?

thank you for your attention!