Algebraic multigrid and subdivision

Maria Charina

University of Vienna, Austria

joint work with

Marco Donatelli, Lucia Romani, Valentina Turati

Algebraic multigrid

Algebraic multigrid (Ruge, Stüben (1980-...)) is an iterative method for solving linear systems

$$Ax = b, \quad A \in \mathbb{C}^{n \times n}, \quad b \in \mathbb{C}^n, \quad n \in \mathbb{N},$$

with A symmetric, positive definite, sparse, and $\lambda_{min}(A) \approx 0$.

Algebraic multigrid (Ruge, Stüben (1980-...)) is an iterative method for solving linear systems

$$Ax = b, \quad A \in \mathbb{C}^{n\times n}, \quad b \in \mathbb{C}^n, \quad n \in \mathbb{N},$$

with A symmetric, positive definite, sparse, and $\lambda_{min}(A) \approx 0$.

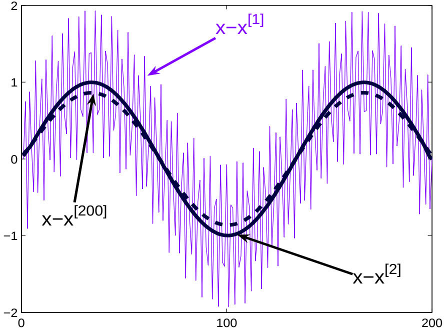
There are other capable iterative solvers, e.g., ω-Jacobi

$$x^{[\ell+1]} = (I - \frac{\omega}{2}A)x^{[\ell]} + \frac{\omega}{2}b, \quad \ell \in \mathbb{N}_0, \quad 0 < \omega \leq 1.$$

Drawback of ω-Jacobi: slow convergence. If $\lambda_{min}(A) \approx 0$, then

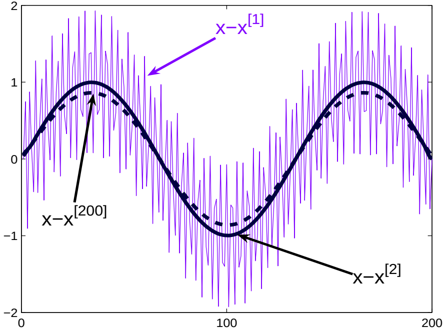
$$\lambda_{max}(I - \frac{\omega}{2}A) \approx 1.$$

Error of $\frac{1}{2}$—Jacobi



$$x - x^{[\ell]} = \sum_{j=1}^{n} c_j \lambda_j^{\ell} v_j, \quad (I - \frac{\omega}{2} A) v_j = \lambda_j v_j, \quad j = 1, \dots, n.$$

Error of $\frac{1}{2}$−Jacobi



$$x - x^{[\ell]} = \sum_{j=1}^{n} c_j \lambda_j^{\ell} v_j, \quad (I - \frac{\omega}{2} A) v_j = \lambda_j v_j, \quad j = 1, \dots, n.$$

- Multigrid philosophy: keep the solver simple and accelerate its convergence by multilevel error correction

$$x \approx x^{[2]} + e_1 + e_2 + \dots + e_k, \quad k \leq \log(n).$$

Oversimplified idea of multilevel error correction

$$n = \frac{n}{2} + \frac{n}{4} + \frac{n}{8} + \dots$$

Oversimplified idea of multilevel error correction

$$n = \frac{n}{2} + \frac{n}{4} + \frac{n}{8} + \dots$$

allows us to compute

$$x \approx x^{[2]} + e_1 + e_2 + \dots + e_k, \quad k \leq \log(n),$$

in $O(n)$ computational steps.

Algebraic multigrid (Ruge, Stüben (1980-...)):

- complexity of one iteration is $O(n)$, due to multilevel error correction

size e.g. $n = 2^k$: $A_0 = A, \; \tilde{x}^{[0]} = x^{[2]}$ (two steps of ω-Jacobi)

for $j = 1, \ldots, k$

size $2^{-j} n$: $\begin{cases} A_j = P_j^T A_{j-1} P_j \\ \text{Solve} \quad A_j \, \tilde{e}_j = P_j^T \cdots P_1^T (b - A \, \tilde{x}^{[j-1]}) \end{cases}$

size n: $\boxed{e_j = P_1 \cdots P_j \tilde{e}_j}$ and $\tilde{x}^{[j]} = \tilde{x}^{[j-1]} + e_j$

end

Algebraic multigrid (Ruge, Stüben (1980-...)):

- complexity of one iteration is $O(n)$, due to multilevel error correction

size e.g. $n = 2^k$: $A_0 = A, \ \tilde{x}^{[0]} = x^{[2]}$ (two steps of ω-Jacobi)

for $j = 1, \ldots, k$

$$\text{size } 2^{-j} n: \begin{cases} A_j = P_j^T A_{j-1} P_j \\ \text{Solve} \quad A_j \, \tilde{e}_j = P_j^T \cdots P_1^T (b - A \, \tilde{x}^{[j-1]}) \end{cases}$$

$$\text{size } n: \boxed{e_j = P_1 \cdots P_j \tilde{e}_j} \quad \text{and} \quad \tilde{x}^{[j]} = \tilde{x}^{[j-1]} + e_j$$

end

- operator-dependent: P_j depend on the properties of A.

Algebraic multigrid (Ruge, Stüben (1980-...)):

- complexity of one iteration is $O(n)$, due to multilevel error correction

size e.g. $n = 2^k$: $\quad A_0 = A,\ \tilde{x}^{[0]} = x^{[2]} \quad$ (two steps of ω-Jacobi)

for $j = 1, \ldots, k$

size $2^{-j}n$: $\begin{cases} A_j = P_j^T A_{j-1} P_j \\ \text{Solve} \quad A_j\, \tilde{e}_j = P_j^T \cdots P_1^T (b - A\, \tilde{x}^{[j-1]}) \end{cases}$

size n: $\boxed{e_j = P_1 \cdots P_j \tilde{e}_j} \quad$ and $\quad \tilde{x}^{[j]} = \tilde{x}^{[j-1]} + e_j$

end

- operator-dependent: P_j depend on the properties of A.

Goal: define full rank, sparse P_j, $j = 1, \ldots, k$, such that

$$\lim_{L \to \infty} \|x - \tilde{x}^{[k,L]}\|_A = 0, \quad \tilde{x}^{[k,L]} = x^{[2,L]} + e_{1,L} + \ldots + e_{k,L},$$

and number of iterations L for reaching TOL is independent of n.

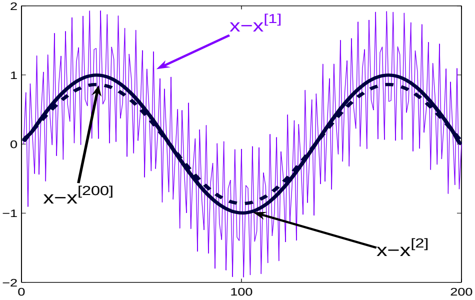
Multigrid and subdivision (Ch., Donatelli, Romani, Turati (2016))

Multigrid and subdivision in a nutshell:

- multilevel error correction with wavelet flavor (Brandt (1986): for $\omega = \frac{1}{2}$ the error is smooth)

$$x \approx \underbrace{x^{[2]}}_{oscillatory} + \underbrace{e_1 + \ldots + e_k}_{smooth\ \ error}$$

Error of $\frac{1}{2}$−Jacobi



Properties of A dictated by applications (e.g. elliptic PDEs):

- $A = T(f)$ is (block) Toeplitz with symbol
  - 2-dim Laplacian (finite differences)

    $$f(x,y) = \frac{2 - \cos x - \cos y + c(2 - \cos(x+y) - \cos(x-y))}{2+2c}, \quad c \geq 0,$$

    or

    $$f(x,y) = (1 - \cos x) + b(1 - \cos(y)), \quad b > 0.$$

  - 1-dim Laplacian (isogeometric approach with μ-th order B-splines)

    $$f_\mu(x) = (2 - 2\cos x)h_\mu(x), \quad h_\mu > 0 \text{ trig. polynomial.}$$

- A is symmetric and sparse (f real trigonometric polynomial)
- A positive definite, $\lambda_{min}(A) \approx 0$ ($f(0) = 0$, $f > 0$ otherwise)

Multigrid and subdivision in a nutshell:

- multilevel error correction with wavelet flavor (Brandt (1986): for $\omega = \frac{1}{2}$ the error is smooth)

$$x \approx \underbrace{x^{[2]}}_{oscillatory} + \underbrace{e_1 + \ldots + e_k}_{smooth\ \ error}.$$

- for Toeplitz $A = T(f)$ with $f(0) = 0$, subdivision task is coarse-to-fine propagation of smooth errors

$$P_j : \mathbb{R}^{2^{-j}n} \to \mathbb{R}^{2^{-j+1}n}, \quad e_j = P_1 \cdots P_j \, \tilde{e}_j, \quad j = 1, \ldots, k.$$

Multigrid and subdivision in a nutshell:

- multilevel error correction with wavelet flavor (Brandt (1986): for $\omega = \frac{1}{2}$ the error is smooth)

$$x \approx \underbrace{x^{[2]}}_{oscillatory} + \underbrace{e_1 + \ldots + e_k}_{smooth\ \ error}.$$

- for Toeplitz $A = T(f)$ with $f(0) = 0$, subdivision task is coarse-to-fine propagation of smooth errors

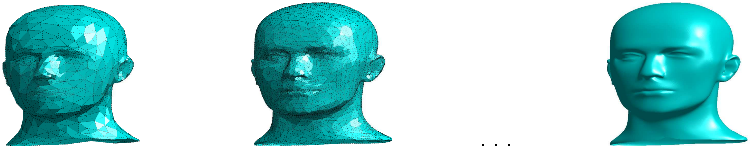
$$P_j : \mathbb{R}^{2^{-j}n} \to \mathbb{R}^{2^{-j+1}n}, \quad e_j = P_1 \cdots P_j \tilde{e}_j, \quad j = 1, \ldots, k.$$

- Convergence and optimality of multigrid are influenced by properties of subdivision.

(Optimality: number of iterations for reaching TOL is independent of n.)

Subdivision (de Rahm (1956)):

- iterative method for local and smoothing mesh refinement.
- Application: computer animation.



$c \in \mathbb{R}^{N_0 \times 3}$ $\qquad$ $\mathcal{P}_j c \in \mathbb{R}^{N_1 \times 3}$ $\qquad \cdots \qquad$ $\mathcal{P}_1 \cdots \mathcal{P}_j c \in \mathbb{R}^{N_j \times 3}$

- Multigrid error propagation via subdivision

$$P_j : \mathbb{R}^{2^{-j} n} \to \mathbb{R}^{2^{-j+1} n}, \quad e_j = P_1 \cdots P_j \, \tilde{e}_j, \quad j = 1, \ldots, k,$$

where sparse and full rank P_j is a rectangular sub-matrix of $\mathcal{P}_j$.

A Toeplitz matrix $A = T(f)$ is defined by symbol f, e.g.

$$T(f) = \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}, \quad f(x) = -e^{-ix} + 2 - e^{ix}.$$

Subdivision step consists of up-sampling and convolution. E.g.

$$P_j = \underbrace{T_j(p)}_{Toeplitz} \begin{pmatrix} 1 & 0 & & & \\ 0 & 0 & & & \\ 0 & 1 & & & \\ & & \ddots & & \\ & & & 1 & 0 \\ & & & 0 & 0 \\ & & & 0 & 1 \\ & & & 0 & 0 \end{pmatrix} = \frac{1}{2} \begin{pmatrix} 2 & & & \\ 1 & 1 & & \\ & \ddots & & \\ & & 2 & \\ & & 1 & 1 \\ & & & 2 \\ & & & 1 \end{pmatrix}$$

with symbol $p(x) = \frac{1}{2}\left(e^{-ix} + 2 + e^{ix}\right)$. Goal: match f and p.

Not all f and p match. For Toeplitz $A = T(f)$ with

$$f = \left(-e^{-ix} + 2 - e^{ix}\right)^2, \quad x \in [0, 2\pi),$$

we get

Subdivision scheme defining P_j	$n = 2^{10}$ *iter*	$n = 2^{11}$ *iter*	$n = 2^{12}$ *iter*
Linear Bspline	617	744	801
Cubic Bspline	40	43	45
Interp. 6-point	13	13	14

Multigrid is convergent, but is not optimal for linear B-spline subdivision scheme. Why?

(For simplicity, a univariate version.)

Theorem: (Ch, Donatelli, Romani, Turati (2016)):

Assume that the symbol of the system matrix $A = T(f)$ satisfies

$$D^m f(0) = 0, \;\; m = 0, \ldots, M, \quad \textit{and} \quad f(x) \neq 0, \;\; x \in (0, 2\pi).$$

If

$$(i) \;\; D^m p(\pi) = 0, \quad m = 0, \ldots, M, \;\; \textit{(polynomial generation)}$$

and

$$(ii) \;\; |p(x)| > 0 \textit{ for } x \in [-\tfrac{\pi}{2}, \tfrac{\pi}{2}], \qquad \textit{(stability)}$$

then multigrid is convergent and optimal.

Not all f and p match. For Toeplitz $A = T(f)$ with

$$f = \left(-e^{-ix} + 2 - e^{ix}\right)^2, \quad x \in [0, 2\pi),$$

that has a double zero at 0 (i.e. $M = 1$), we get

Subdivision scheme defining P_j	$n = 2^{10}$ *iter*	$n = 2^{11}$ *iter*	$n = 2^{12}$ *iter*
Linear Bspline	617	744	801
Cubic Bspline	40	43	45
Interp. 6-point	13	13	14

Multigrid is convergent, but is not optimal for linear B-spline subdivision scheme. Reason: the corresponding symbol

$$p(x) = \frac{1}{2}(e^{-ix} + 2 + e^{ix}), \quad x \in [0, 2\pi),$$

has a simple zero at π.

(For simplicity, univariate version (several zeroes of f).)

Theorem: (Ch, Donatelli, Romani, Turati (2016)):

Assume that the symbol f of the system matrix $A = T(f)$ satisfies

$$\begin{cases} D^m f(y) = 0, \quad y \in \{0, \pi\}, \quad m = 0, \ldots, M, \quad M \in \mathbb{N}_0, \\ f(x) \neq 0, \ x \in (0, \pi) \cup (\pi, 2\pi). \end{cases}$$

If

(*i*) $D^m p(\frac{2\pi}{3}) = 0, \quad D^m p(\frac{4\pi}{3}) = 0, \quad m = 0, \ldots, M,$

and

(*ii*) $|p(x)| > 0$ *for* $x \in [-\frac{\pi}{3}, \frac{\pi}{3}]$,

then multigrid is convergent and optimal.

Isogeometric approach with μ-th order B-splines (Donatelli, Garoni, Manni, Serra-Capizzano, Spellers (2015)) deals with $A = T(f_\mu)$,

$$f_\mu = \left(-e^{-ix} + 2 - e^{ix}\right) h_\mu(x), \qquad \lim_{\mu\to\infty} h_\mu(\pi) = 0.$$

We get

Subdivision scheme defining P_j	$\mu = 3$ *iter*	$\mu = 10$ *iter*	$\mu = 16$ *iter*
Binary, interp. 6-point	8	13	126
Ternary, interp. 4-point	30	17	49

Ternary schemes remove the singularity at π.

Sketch of the proof

How to match f and p for multigrid convergence and optimality

Fiorentino, Serra (1991), Böttcher et. al. (2006)

In our case, implementation of algebraic multigrid

- is done with positive definite (block) Toeplitz $A = T(f)$

Analysis of algebraic multigrid

- is done for (block) circulant semi-positive definite $A = C(f)$
- Circulant matrices form a matrix algebra ($A_j = C(f_j)$)
- Circulant matrices are diagonalizable via Fourier transform

$$A_j = C(f_j) = F_j \operatorname{diag}\left(f_j(x_r) \ : \ x_r = \frac{2\pi r}{2^{-j} n} \right) F_j^*$$

$$P_j = C_j(p) K_j = F_j \operatorname{diag}\left(p(x_r) \ : \ x_r = \frac{2\pi r}{2^{-j} n} \right) F_j^* K_j$$

- Circulant matrices approximate Toeplitz matrices well

Algebraic two grid method yields $x \approx x^{[2]} + e_1$, where

size n : $x^{[2]}$ = two steps of ω-Jacobi

size $\frac{n}{2}$: $\begin{cases} A_1 = P^T A P \\ \tilde{e}_1 = A_1^{-1} P^T (b - A x^{[2]}) \end{cases}$

size n : $e_1 = P \tilde{e}_1$

via subdivision

Algebraic two grid method yields $x \approx x^{[2]} + e_1$, where

size n : $x^{[2]}$ = two steps of ω-Jacobi

size $\frac{n}{2}$: $\begin{cases} A_1 = P^T A P \\ \tilde{e}_1 = A_1^{-1} P^T (b - A x^{[2]}) \end{cases}$

size n : $e_1 = P \tilde{e}_1$ via subdivision

- Exact solution $x = x^{[2]} + e$, iff $x - x^{[2]} \in \text{Range}(P)$. Due to

$$x - x^{[2]} - e = \underbrace{(I - P A_1^{-1} P^T A)}_{\substack{orthogonal\ \ projection \\ w.r.t\ \ (A\cdot,\cdot)}} (x - x^{[2]}).$$

Algebraic two grid method yields $x \approx x^{[2]} + e_1$, where

size n : $x^{[2]} =$ two steps of ω-Jacobi

size $\frac{n}{2}$: $\begin{cases} A_1 = P^T A P \\ \tilde{e}_1 = A_1^{-1} P^T (b - A\,x^{[2]}) \end{cases}$

size n : $e_1 = P\tilde{e}_1$

via subdivision

- Exact solution $x = x^{[2]} + e$, iff $x - x^{[2]} \in \text{Range}(P)$. Due to

$$x - x^{[2]} - e = \underbrace{(I - PA_1^{-1}P^T A)}_{\substack{orthogonal\ \ projection \\ w.r.t\ \ (A\cdot,\cdot)}}(x - x^{[2]}).$$

- Best approximation (Ruge, Stuben):

$$\|e - PA_1^{-1}P^T A\,e\|_A = \min_{\tilde{e}\in\mathbb{C}^{n/2}} \|e - P\tilde{e}\|_A \quad \text{for all } e \in \mathbb{C}^n.$$

Algebraic two grid method yields $x \approx x^{[2]} + e_1$, where

size n : $x^{[2]}$ = two steps of ω-Jacobi

size n : $e_1 = P\tilde{e}_1$ via subdivision

size $\frac{n}{2}$: $\begin{cases} A_1 = P^T A P \\ \tilde{e}_1 = A_1^{-1} P^T (b - A x^{[2]}) \end{cases}$

- Exact solution $x = x^{[2]} + e$, iff $x - x^{[2]} \in \text{Range}(P)$. Due to

$$x - x^{[2]} - e = (\underbrace{I - PA_1^{-1}P^T A}_{\substack{orthogonal\ \ projection \\ w.r.t\ \ (A\cdot,\cdot)}})(x - x^{[2]}).$$

- Best approximation (Ruge, Stuben):

$$\|e - PA_1^{-1}P^T A\, e\|_A = \min_{\tilde{e} \in \mathbb{C}^{n/2}} \|e - P\tilde{e}\|_A \quad \text{for all } e \in \mathbb{C}^n.$$

Match f and p: $\exists C > 0$ independent of n such that

$$\|(I - PA_1^{-1}P^T A)(I - \frac{\omega}{2} A)\|_A = C < 1.$$

Similarly, for multigrid.

Theorem (Ruge, Stuben 1987)

Let A be positive definite. If, for all $j = 1, \ldots, k$,

(i) *"iteration matrices $I - \frac{1}{4}A_j$ are contractive w.r.t. $\|\cdot\|_{A_j}$ and their contraction constants are independent of n",*

(ii) *coarse grid correction operators are uniformly bounded, i.e. $\exists \gamma_j > 0$ independent of n such that*

$$\|(I - P_j A_{j+1}^{-1} P_j^T A_j)\, x\|_{A_j} \leq \gamma_j \|x\|_{A_j^2} \quad \forall x,$$

then multigrid is convergent and optimal.

(ii) After the Fourier transform and algebraic manipulations, it is left to show that $\exists \gamma_j > 0$, $j = 0, \ldots, k$, independent of n and

$$\sup_{x \in [0,2\pi)} \frac{|p(x+\pi)|^2 f_j(x+\pi)}{\Big(|p(x)|^2 f_j(x) + |p(x+\pi)|^2 f_j(x+\pi)\Big) f_j(x)} \leq \gamma_j$$

and

$$\sup_{x \in [0,2\pi)} \frac{p(x)\overline{p(x+\pi)}}{|p(x)|^2 f_j(x) + |p(x+\pi)|^2 f_j(x+\pi)} \leq \gamma_j.$$

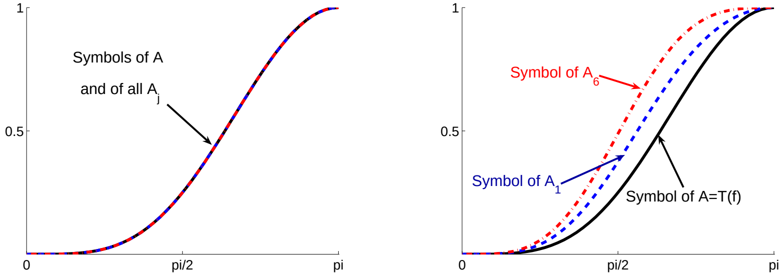
The choice of p:

- For $f_j(0) = 0$ choose p such that $p(0) \neq 0$ and $p(\pi) = 0$.
- To ensure $f_j(0) = 0$ choose p to satisfy

$$|p(x)|^2 + |p(x+\pi)|^2 > 0, \quad x \in [0, 2\pi).$$

□

What do coarse-to-fine operators P_j really do?

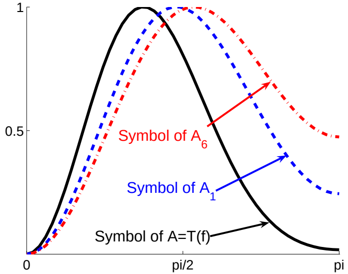


bad P_j good P_j

Good subdivision schemes (that define P_j) improve the conditioning of

$$A_j = P_j^T A_{j-1} P_j, \quad j = 1, \dots, k.$$

What do coarse-to-fine operators P_j really do?



Bad (binary) P_j Good (ternary) P_j

Good subdivision schemes remove the singularity at π

$$A_j = P_j^T A_{j-1} P_j, \quad j = 1, \ldots, k.$$

Summary:

- Algebraic properties of multivariate subdivision symbols influence convergence and optimality of multigrid for system matrices $A = T(f)$, $f(0) = 0$.
- The choice of associated dilation Matrix influences the conditioning of multigrid.
- Still to do:
  - dual subdivision and face-centered discretizations of PDE's
  - anisotropic dilations and semi-coarsening
  - ...

Summary:

- Algebraic properties of multivariate subdivision symbols influence convergence and optimality of multigrid for system matrices $A = T(f)$, $f(0) = 0$.
- The choice of associated dilation Matrix influences the conditioning of multigrid.
- Still to do:
  - dual subdivision and face-centered discretizations of PDE's
  - anisotropic dilations and semi-coarsening
  - ...

Thank you for your attention!