Algebraic multigrid for stochastic matrices

Ivana Pultarová

Department of Mathematics, Faculty of Civil Engineering
Czech Technical University in Prague

PANM 16, Horní Maxov, 2012

Outline

1. Algebraic multigrid (AMG)
2. Stochastic matrices, Markov chains, stationary probability distribution vector
3. Iterative aggregation - disaggregation (IAD) methods
4. Convergence and divergence of IAD
5. Conclusions and open questions

Algebraic multigrid (AMG)

We have to solve $Ax = b$, where $A \in \mathcal{R}^{N\times N}$ is

- symmetric and positive definite
- sparse
- $A\mathbb{1} \approx 0$, where $\mathbb{1}$ is all ones vector and 0 is a zero vector

Then the eigenpairs of A are

- $\lambda_k \approx 0$ and $v_k \approx (\sin(2\pi kj/N))_j$ for $k = 0, 1, \ldots$
- $\lambda_k \gg 0$ and $v_k \approx (\sin(2\pi kj/N))_j$ for $k = \ldots, N/2-2, N/2-1$.

Iterative methods, e.g. Jacobi method or Richardson method:

$$x^{n+1} = (I - A)x^n + b$$

and thus for error vectors $e^n = x - x^n$

$$e^{n+1} = (I - A)e^n.$$

The eigenvalues of the iteration matrix $M = I - A$ are then in $(-1, 1)$ or in $(-\frac{1}{2}, 1)$ if e.g. $c \cdot Ax = c \cdot b$ is considered instead of $Ax = b$.

Suppose

$$e^0 = \sum c_k v_k.$$

The iteration process

$$e^{n+1} = (I - A)e^n$$

with the iteration matrix $M = I - A$ ”smoothes” the error:

- high frequency components v_k of the error are annilihated faster than that with small frequencies,
- or, after some iterations with M the error vector e^k contains only the eigenvectors of A with low eigenmodes,
- then $||r^n|| = ||b - Ax^n|| = ||A(x - x^n)|| = ||Ae^n|| \ll ||e^n||$.

How does the error vector e^n look like?

Note that

$$v^T A v = \frac{1}{2} \sum_{i,j} -A_{ij}(v_i - v_j)^2 + \sum_{i,j} A_{ij} v_i^2.$$

The last term is almost zero if row sums of A are almost null.

Denote

$$e^n = x - x^n, \qquad r^n = b - Ax^n = Ae^n,$$

and

$$E_n = e^{n\,T} r^n = e^{n\,T} A e^n, \qquad R_n = \sum_i (r_i^n)^2 / A_{ii}.$$

Thus

$$||r^n|| \ll ||e^n||$$

is equivalent to

$$R_n \ll E_n.$$

We have

$$E_n^2 = (e^{n\,T} r^n)^2 \leq \sum (e_i^n)^2 A_{ii} \cdot \sum (r_k^n)^2 / A_{kk} = R_n \sum (e_i^n)^2 A_{ii} \ll E_n \sum (e_i^n)^2 A_{ii}$$

thus $E_n \ll \sum (e_i^n)^2 A_{ii}$, which yields

$$E_n = e^{n\,T} A e^n = \frac{1}{2} \sum_{i,j} -A_{ij}(e_i^n - e_j^n)^2 \ll \sum (e_i^n)^2 A_{ii}.$$

There is almost no difference between "strongly connected" components of the error vector in case $R_n \ll E_n$.

Summary. For $e^n = x - x^n$

- e^n contains low frequency vectors,
- $e_i \approx e_k$ whenever $A_{ik} \ll 0$.

Need for eliminating the error components with low frequencies leads to construction and solution of a smaller ”coarser” problems.

Two main approaches:

- **geometric MG** - exploiting the properties of the equation, of physics and of the geometry of the underlying problem,
- **algebraic MG** - like a black-box; easier to apply; but no hint from the original problem,

  special approach **aggregation based algebraic MG**

aggregation based algebraic MG - how to construct the coarse problem

Let rows of $R \in \mathcal{R}^{N_c \times N}$ be (approximations of) low frequency vectors of A, $N_c < N$.
Let the coarse matrix and the coarse right hand side be

$$A_c = RAR^T, \qquad r_c = Rr^n,$$

then

$$A_c u_c = r_c$$

is a restriction of the problem to a "coarser mesh" and

$$x^n_{new} = x^n + R^T u_c$$

is a better approximation to x.
We need A_c sparse, thus R must contain many zeros,

$$R = \begin{pmatrix} \times & \times & 0 & \dots & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & \dots & \dots & 0 & \times & \times \end{pmatrix},$$

i.e. R must have piecewise constant rows.
For example, ones in positions of strong connections.

The resulting iterative process is

$$\begin{aligned} x^{k+1,1} &= (I - A)x^k + b \\ x^{k+1} &= \left(I - R^T A_c^{-1} RA\right) x^{k+1,1} + R^T A_c^{-1} Rb. \end{aligned}$$

and the resulting iteration matrix is

$$M_{\mathrm{MG}} = (I - R^T A_c^{-1} RA)(I - A).$$

Spectra of $I - A$ and of M_{MG} are

$\sigma(I - A)$ $\quad -1 \ll \lambda_N \leq \lambda_{N-1} \leq \ldots\ldots\cdots \leq \lambda_2 \leq \lambda_1 < 1$

$\sigma(M_{\mathrm{MG}})$ $\quad -1 \ll \lambda_N \leq \lambda_{N-1} \leq \cdots \leq \lambda_{(\approx N_c+1)} \ll 1$

MG - rules for aggregation

geometric MG - according to the location of elements, properties of FEs and to the operator of the problem, [P. Vaněk, ... many papers].

algebraic MG - according to the "strength of the connection", size of the corresponding offdiagonal elements, [e.g. A. Brandt, Algebraic multigrid theory: The symmetric case, 1983]. Advantageous in case of singularities, narrow shapes, etc.

Examples prepared by E. Dvořáková according to [Y. Notay, 2011].

a) Laplace operator with anizotropy,
b) linear elasticity,

in both cases Dirichlet boundary conditions are used.

We compare spectral radii of $I - A$ and of M_{MG} for various choices of groups.

a) Laplace operator with anizotropy

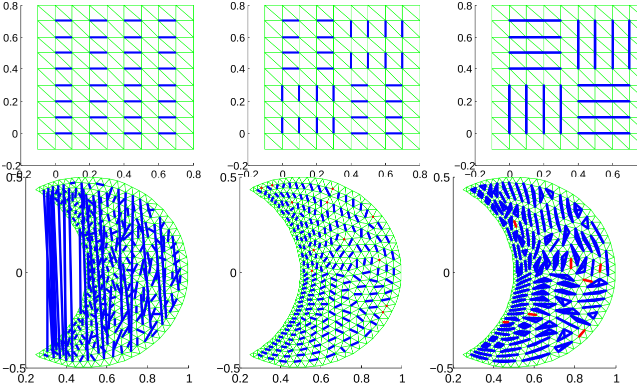


Table: Spectral radii of iteration matrices.

mesh	0.1	0.05	0.025	0.0125
$I - A$	0.5525	0.8775	0.9721	0.9466
MG according to numbering of nodes	0.3202	0.7428	0.8647	0.9466
MG with slow eigenvectors	0.1926	0.2341	0.3586	0.4127
MG with Notay's pairs	0.2362	0.4639	0.6915	0.7741
MG with Notay's pairs of pairs	0.2450	0.6086	0.7919	0.8653

b) linear elasticity

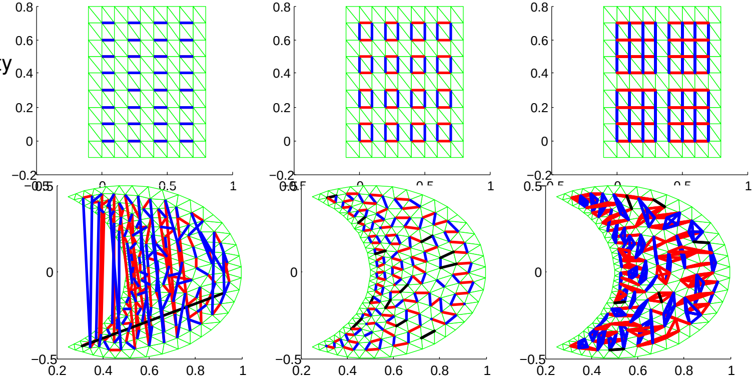


Table: Spectral radii of iteration matrices.

mesh	0.1	0.05	0.025
$I - A$	0.4918	0.8653	0.9705
MG according to numbering of nodes	0.3040	0.6663	0.7819
MG with slow eigenvectors	0.1427	0.2431	0.3204
MG with pairs of x-y displacements	0.4558	0.8338	0.9604
MG with Notay's pairs	0.2107	0.7113	0.7870
MG with Notay's pairs of pairs	0.2399	0.7880	0.8813

Symmetric vs. nonsymmetric MG

. . . It means, MG for symmetric and nonsymmetric marices.

Theorem. The MG method converges for symmetric matrix A for any kind of aggregation and for any number of smoothing steps within one multigrid cycle.

No analogous statement has been proved for the nonsymmetric case - stochastic matrices.

However, the MG algorithms for nonsymmetric problems exist and can be studied. ”Mostly they converge.” The heuristical explanation of their fast convergence is based on similarity with the symmetric case.
[H. De Sterck, T. A. Manteuffel, S. F. McCormick, Q. Nguyen and J. Ruge . . . many papers 2008 - 2012]

But, there are basic differences between symmetric and nonsymmetric MG.

Application of MG to the solution of Markov chains - special name: **iterative aggregation - disaggregation (IAD) methods**.

Problem description.

We assume an **irreducible** $N \times N$ **column stochastic** matrix B, i.e.

$$B \geq 0 \qquad \text{and} \qquad e^T B = e^T.$$

Find **stationary probability distribution vector** (Perron eigenvector) of a column stochastic matrix, i.e. find x such that

$$Bx = x, \qquad e^T x = 1$$

or

$$(I - B)x = 0, \qquad e^T x = 1.$$

Perron-Frobenius theorem.

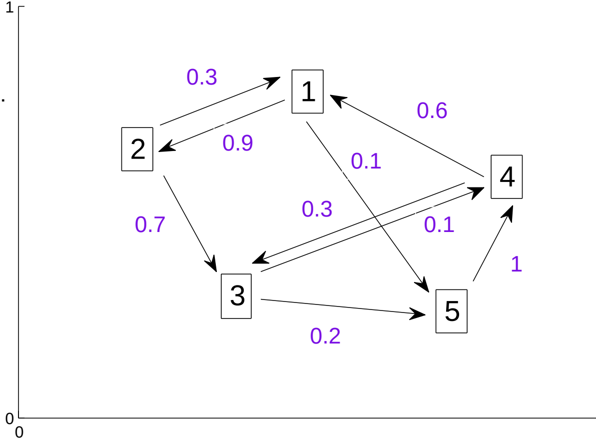
Solution

- Direct solvers.
- Numerical solution. Power method, Jacobi m., Gauss-Seidel m., their block modifications. Iteration matrix $T = M^{-1}W$, where $I - B = M - W$, is a regular splitting ($M^{-1} \geq 0$, $W \geq 0$). For example: $M = I$, $M =$ block-diagonal of $I - B$, $M =$ block-upper-triangle of $I - B$, ...
- Multilevel methods, based on aggregation of states: iterative aggregation - disaggregation (IAD) methods. Motivation from PDEs. BUT no symmetry here.

Example. Find stationary probability distribution of the random process with five states characterized by probabilities B_{ij} of transition from j to i.

$$B = \begin{pmatrix} 0 & 0.3 & 0 & 0.6 & 0 \\ 0.9 & 0 & 0 & 0 & 0 \\ 0 & 0.7 & 0.7 & 0.3 & 0 \\ 0 & 0 & 0.1 & 0.1 & 1 \\ 0.1 & 0 & 0.2 & 0 & 0 \end{pmatrix}.$$

Vector x is $x \approx \begin{pmatrix} 0.1390 \\ 0.1251 \\ 0.4609 \\ 0.1691 \\ 0.1061 \end{pmatrix}.$



1 Algebraic multigrid (AMG)

2 Stochastic matrices, Markov chains, stationary probability distribution vector

3 Iterative aggregation - disaggregation (IAD) methods

4 Convergence and divergence of IAD

5 Conclusions and open questions

Notation

Set of events $\{1, 2, 3, \ldots, N\}$ divided into n groups $G_1, G_2, \ldots, G_{N_c}$,

$$\cup_{j=1}^{N_c} G_j, \qquad G_j \cap G_k = \emptyset, \quad \text{when} \quad j \neq k.$$

Communication matrices R (fine to coarse level) and $S(x)$ (coarse to fine) are e.g. for $G_1 = \{1, 2\}$, $G_2 = \{3, 4, 5\}$

$$R = \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix}, \qquad S(y) = \begin{pmatrix} 1/3 & 0 \\ 2/3 & 0 \\ 0 & 2/6 \\ 0 & 3/6 \\ 0 & 1/6 \end{pmatrix}, \qquad \text{if} \quad y = \frac{1}{12} \begin{pmatrix} 2 \\ 4 \\ 2 \\ 3 \\ 1 \end{pmatrix}.$$

Aggregated matrix $RBS(y) =$

$$= \begin{pmatrix} 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 & 0 & 0.1 & 0 \\ 0 & 0.3 & 0.5 & 0 & 1 \\ 0 & 0 & 0.5 & 0 & 0 \\ 1 & 0.4 & 0 & 0 & 0 \\ 0 & 0.3 & 0 & 0.9 & 0 \end{pmatrix} \begin{pmatrix} 1/3 & 0 \\ 2/3 & 0 \\ 0 & 2/6 \\ 0 & 3/6 \\ 0 & 1/6 \end{pmatrix} = \begin{pmatrix} 1/5 & 23/60 \\ 4/5 & 37/60 \end{pmatrix}.$$

Projection $P(y) = S(y)R$.

Algorithm of IAD method for $Bx = x$.

1. Choose initial approximation x^0. Set $k := 0$.
2. Solve (small) $N_c \times N_c$ problem
$$RBS(x^k)z = z.$$
(*z* is carried out exactly.)
3. Prolong z from $\mathcal{R}^{N_c}$ to the original size $\mathcal{R}^N$,
$$y = S(x^k)z,$$
and apply ν steps of (large) basic iteration $T \in \mathcal{R}^{N \times N}$,
$$x^{k+1} = T^\nu y.$$
4. If $||x^{k+1} - x^k||$ small then STOP, else $k := k + 1$ and GOTO Step 2.

[W. J. Stewart, Introduction to the Numerical Solutions of Markov Chains, 1994,
P. Buchholz, T. Dayar, G. Horton, S. T. Leutenegger, U. R. Krieger, A. N. Langville, C. D. Meyer]

Error propagation formula

$$x^{k+1} - x = J(x^k)(x^k - x),$$

where
$$J(x^k) = T^\nu \left(I - P(x^k)(B - xe^T)\right)^{-1} \left(I - P(x^k)\right), \qquad P(x^k) = S(x^k)R.$$

[I. Marek, P. Mayer, 1998]. (Exploited for local convergence proofs.)

Comparison of the error propagation matrices of AMG and of IAD

Note

$$I - B \approx A, \qquad B \approx I - A.$$

AMG:

$$M_{\mathrm{MG}} = (I - A)\left(I - R^T (RAR^T)^{-1} RA\right)$$

IAD:

$$\begin{aligned} J(x^k) &= B\left(I - P(x^k)(B - xe^T)\right)^{-1} (I - P(x^k)) \\ &= \dots \\ &= B\left(I - S(x^k)\left(R(I - Z)S(x^k)\right)^{-1} R(I - B)\right) \end{aligned}$$

where

$$Z = B - xe^T.$$

Small example
Stochastic matrix

$$B = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}, \qquad x = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$$

- **Power method** yields a convergent sequence x^k for all starting x^0. Note $\rho(B - xe^T) = 1/2$.
- **The IAD method** with $T = B$, $\nu = 1$ and aggregation

$$B = \left(\begin{array}{c|cc} 1/2 & 0 & 1/2 \\ \hline 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{array}\right)$$

  yields divergent (oscillating) sequence x^k for almost all starting x^0. Note $\rho(J(x)) = 1$.
- **The IAD method** with $T = B$, $\nu = 1$ and aggregation

$$B = \left(\begin{array}{cc|c} 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \\ \hline 0 & 1 & 0 \end{array}\right)$$

  yields exact solution after the second step $x^2 = x$ for any starting x^0. Note $\rho(J(x)) = 0$.

Small example
Stochastic matrix

$$B = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}, \qquad x = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$$

- **Power method** yields a convergent sequence x^k for all starting x^0. Note $\rho(B - xe^T) = 1/2$.
- **The IAD method** with $T = B$, $\nu = 1$ and aggregation

$$B = \left(\begin{array}{c|cc} 1/2 & 0 & 1/2 \\ \hline 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{array}\right)$$

yields divergent (oscillating) sequence x^k for almost all starting x^0. Note $\rho(J(x)) = 1$.

- **The IAD method** with $T = B$, $\nu = 1$ and aggregation

$$B = \left(\begin{array}{cc|c} 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \\ \hline 0 & 1 & 0 \end{array}\right)$$

yields exact solution after the second step $x^2 = x$ for any starting x^0. Note $\rho(J(x)) = 0$.

Small example

Stochastic matrix

$$B = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}, \qquad x = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$$

- **Power method** yields a convergent sequence x^k for all starting x^0. Note $\rho(B - xe^T) = 1/2$.
- **The IAD method** with $T = B$, $\nu = 1$ and aggregation

$$B = \left(\begin{array}{c|cc} 1/2 & 0 & 1/2 \\ \hline 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{array}\right)$$

  yields divergent (oscillating) sequence x^k for almost all starting x^0. Note $\rho(J(x)) = 1$.

- **The IAD method** with $T = B$, $\nu = 1$ and aggregation

$$B = \left(\begin{array}{cc|c} 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \\ \hline 0 & 1 & 0 \end{array}\right)$$

  yields exact solution after the second step $x^2 = x$ for any starting x^0. Note $\rho(J(x)) = 0$.

Small example

Stochastic matrix

$$B = \begin{pmatrix} 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{pmatrix}, \qquad x = \begin{pmatrix} 1/3 \\ 1/3 \\ 1/3 \end{pmatrix}$$

- **Power method** yields a convergent sequence x^k for all starting x^0. Note $\rho(B - xe^T) = 1/2$.
- **The IAD method** with $T = B$, $\nu = 1$ and aggregation

$$B = \left(\begin{array}{c|cc} 1/2 & 0 & 1/2 \\ \hline 1/2 & 0 & 1/2 \\ 0 & 1 & 0 \end{array}\right)$$

  yields divergent (oscillating) sequence x^k for almost all starting x^0. Note $\rho(J(x)) = 1$.
- **The IAD method** with $T = B$, $\nu = 1$ and aggregation

$$B = \left(\begin{array}{cc|c} 1/2 & 0 & 1/2 \\ 1/2 & 0 & 1/2 \\ \hline 0 & 1 & 0 \end{array}\right)$$

  yields exact solution after the second step $x^2 = x$ for any starting x^0. Note $\rho(J(x)) = 0$.

Multi-level IAD procedure (input: B, x; output: y)

1. Construct T and apply μ steps of pre-smoothing: $x := T^{\mu}x$.
2. If $size(B) < \tau$ solve $RBS(x^k)z = z$

   else call **Multi-level IAD procedure** (input: $RBS(x)$, Rx; output: z).
3. Prolong z to $x := S(x)z$.
4. Apply ν steps of post-smoothing $y := T^{\nu}x$.

[H. De Sterck, T. A. Manteuffel, S. F. McCormick, Q. Nguyen and J. Ruge, 2009, 2010, E. Treister, I. Yavneh, 2011, . . .]

Choice of aggregates is very important.

Mostly according to ”strong connections” between states.

NCD Markov chains

Nearly completely decomposable (NCD) Markov chains:
Diagonal blocks of much larger magnitude than the off-diagonal blocks; their largest eigenvalues close to one; other eigenvalues separated from one.
Sufficient for global convergence of IAD [W. J. Stewart, 1994].
Main drawback: hard to recognize whether a matrix is NCD or not.

Non-zero pattern of B

B has a positive row or column or diagonal, and $T = \alpha B + (1 - \alpha) I$, $\nu = 1$ are sufficient for local convergence. [I. Marek, I. Pultarová, 2006].

Special choice of groups ("1,1,...,1,N_1") and $T = B$, $\nu = 1$: necessary and sufficient cond. for global convergence. Estimate of the asymptotic rate of convergence. [I. Ipsen, S. Kirkland, 2006].

General choice of groups and $T = B$, $\nu = 1$: necessary and sufficient cond. for local convergence. Estimate of the asymptotic rate of convergence. [I. Pultarová, 2008].

(Proofs based on relations $\lambda_2(B) \leq \tau(B) = \frac{1}{2} \max_{i,j} ||Be_i - Be_j||_1$ for localization of spectra [E. Seneta, 1984] and on stochastic complement formulation.)

B symmetric or similar to symmetric

B either symmetric (this means B is doubly stochastic and $x = e/N$) or $B = D\,\mathrm{Sym}\,D^{-1}$,

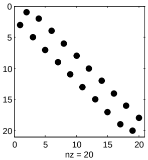
example $B = \begin{pmatrix} 1/2 & 1 \\ 1/2 & 0 \end{pmatrix} = \begin{pmatrix} \sqrt{2} & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1/2 & 1/\sqrt{2} \\ 1/\sqrt{2} & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & 0 \\ 0 & 1 \end{pmatrix}$.

Every 2-level IAD method with $TB = BT$ and $\nu \geq 1$ steps of smoothing converges locally, i.e. $\rho(J(x)) < 1$, [I. Pultarová, I. Marek, 2011].

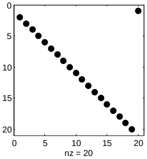
B non-symmetric

For any $K > 0$ there exists (even doubly) stochastic B of size less than $4K$ that $\rho(J(x)) > K$. [I. Pultarová, I. Marek, 2011].

Examples constructed for $T = B$, $\nu = K$, five-diagonal permutation stochastic matrices B of type

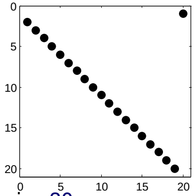


obtained from



by Cuthill-McKee algorithm.

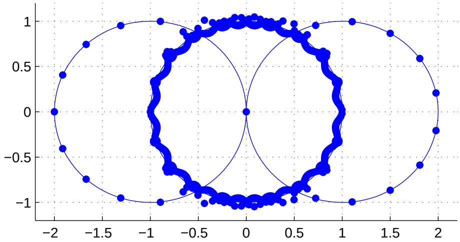
Let B be cyclic defined by the Figure . . .



Let $N = 600$ and number of groups $n = 20$, each of size 30.

Let $T = B^{N/2-1}$.

The spectrum of $J(x)$ is . . .



Multilevel IAD

Reduction and prolongation matrices between levels k and $k+1$ denoted by R_k and $S_k(y)$. Then for the m-levels IAD algorithm we have projections

$$P_{jk}(y) = S_j(y)S_{j+1}(y)\dots S_{k-1}(y)R_{k-1}\dots R_{j+1}R_j, \qquad j<k.$$

Error propagation formula

m-level IAD algorithm with no pre-smoothing and with ν_k steps of the post-smoothing on level k.

The error propagation matrix in step p of the IAD algorithm

$$x^{p+1} - x = J(x^p)(x^p - x),$$

where

$$\begin{aligned} J(y) &= T_1^{\nu_1}\prod_{s=2}^{m-1}(P_{1,s}(y)T_s)^{\nu_s}(I - P_{1,m}(y)Z)^{-1}(I - P_{1,m}(y)) + \\ &\quad + T_1^{\nu_1}\sum_{r=2}^{m-1}\prod_{s=2}^{r-1}(P_{1,s}(y)T_s)^{\nu_s}\sum_{t=0}^{\nu_r-1}(P_{1,r}(y)T_r)^t(I - P_{1,r}(y)). \end{aligned}$$

[I. Pultarová, I. Marek, 2011]

Theorem. The error in the cycle n of a multi-level IAD methods with an arbitrary number of levels $L \geq 2$ and with one pre-smoothing step and with one post-smoothing step in every level, $\mu_m = \nu_m = 1$, $m = 1, 2, \ldots, L-1$, is $x^{n+1} - x = J(x^n)(x^n - x)$, where

$$J(x^n) = T \prod_{k=2}^{L-1} (P_k T)(I - P_L Z)^{-1} \sum_{k=1}^{L-1} (P_k - P_{k+1}) M_{k-1} + T \sum_{m=1}^{L-2} \prod_{k=2}^{m} (P_k T) \sum_{k=1}^{m} (P_k - P_{k+1}) M_{k-1},$$

where $M_0 = T$ and

$$M_k = \left(T + \sum_{j=2}^{k} T P_j (T - I)\right) T,$$

for $k = 1, 2, \ldots, L-2$, $P_1 = I$ and

$$P_k = P(u^1, u^2, \ldots, u^{k-1})_{1k} = S(u^1)_1 \ldots S(u^{k-1})_{k-1} R_{k-1} \ldots R_1,$$

for $k = 2, 3, \ldots, L$, where $u^1 = Tx^n$, $u^2 = R_1 T^2 x^n$, $u^3 = R_2 R_1 T P_2 T^2 x^n$ and

$$u^k = R_{k-1} \ldots R_1 T P_{k-1} T P_{k-2} \ldots T P_3 T P_2 T^2 x^n,$$

for $k = 4, \ldots, L-1$. [I. Pultarová, to appear]

Multi-level IAD with $m > 2$ levels

Error propagation formula available.

Consider these hypotheses:

- If the number of basic iteration is increased, then the convergence is faster.
- If the sum of the numbers of basic iterations before the coarse step and after it remains the same, then the rate of convergence is the same.
- If m-level IAD is locally convergent then $(m+1)$-level IAD is locally convergent.
- If $(m+1)$-level IAD is locally convergent then m-level IAD is locally convergent.
- If B symmetric and $TB = BT$ then m-level IAD is locally convergent.

Multi-level IAD with $m > 2$ levels

Error propagation formula available.

Consider these hypotheses:

- If the number of basic iteration is increased, then the convergence is faster.
- If the sum of the numbers of basic iterations before the coarse step and after it remains the same, then the rate of convergence is the same.
- If m-level IAD is locally convergent then $(m+1)$-level IAD is locally convergent.
- If $(m+1)$-level IAD is locally convergent then m-level IAD is locally convergent.
- If B symmetric and $TB = BT$ then m-level IAD is locally convergent.

Multi-level IAD with $m > 2$ levels

Error propagation formula available.

Consider these hypotheses:

- If the number of basic iteration is increased, then the convergence is faster. NO
- If the sum of the numbers of basic iterations before the coarse step and after it remains the same, then the rate of convergence is the same.
- If m-level IAD is locally convergent then $(m+1)$-level IAD is locally convergent.
- If $(m+1)$-level IAD is locally convergent then m-level IAD is locally convergent.
- If B symmetric and $TB = BT$ then m-level IAD is locally convergent.

Multi-level IAD with $m > 2$ levels

Error propagation formula available.

Consider these hypotheses:

- If the number of basic iteration is increased, then the convergence is faster. NO
- If the sum of the numbers of basic iterations before the coarse step and after it remains the same, then the rate of convergence is the same.
- If m-level IAD is locally convergent then $(m+1)$-level IAD is locally convergent.
- If $(m+1)$-level IAD is locally convergent then m-level IAD is locally convergent.
- If B symmetric and $TB = BT$ then m-level IAD is locally convergent.

Multi-level IAD with $m > 2$ levels

Error propagation formula available.

Consider these hypotheses:

- If the number of basic iteration is increased, then the convergence is faster. NO
- If the sum of the numbers of basic iterations before the coarse step and after it remains the same, then the rate of convergence is the same. NO
- If m-level IAD is locally convergent then $(m+1)$-level IAD is locally convergent.
- If $(m+1)$-level IAD is locally convergent then m-level IAD is locally convergent.
- If B symmetric and $TB = BT$ then m-level IAD is locally convergent.

Multi-level IAD with $m > 2$ levels

Error propagation formula available.

Consider these hypotheses:

- If the number of basic iteration is increased, then the convergence is faster. NO
- If the sum of the numbers of basic iterations before the coarse step and after it remains the same, then the rate of convergence is the same. NO
- If m-level IAD is locally convergent then $(m+1)$-level IAD is locally convergent.
- If $(m+1)$-level IAD is locally convergent then m-level IAD is locally convergent.
- If B symmetric and $TB = BT$ then m-level IAD is locally convergent.

Multi-level IAD with $m > 2$ levels

Error propagation formula available.

Consider these hypotheses:

- If the number of basic iteration is increased, then the convergence is faster. NO
- If the sum of the numbers of basic iterations before the coarse step and after it remains the same, then the rate of convergence is the same. NO
- If m-level IAD is locally convergent then $(m+1)$-level IAD is locally convergent. NO
- If $(m+1)$-level IAD is locally convergent then m-level IAD is locally convergent.
- If B symmetric and $TB = BT$ then m-level IAD is locally convergent.

Multi-level IAD with $m > 2$ levels

Error propagation formula available.

Consider these hypotheses:

- If the number of basic iteration is increased, then the convergence is faster. NO
- If the sum of the numbers of basic iterations before the coarse step and after it remains the same, then the rate of convergence is the same. NO
- If m-level IAD is locally convergent then $(m+1)$-level IAD is locally convergent. NO
- If $(m+1)$-level IAD is locally convergent then m-level IAD is locally convergent.
- If B symmetric and $TB = BT$ then m-level IAD is locally convergent.

Multi-level IAD with $m > 2$ levels

Error propagation formula available.

Consider these hypotheses:

- If the number of basic iteration is increased, then the convergence is faster. NO
- If the sum of the numbers of basic iterations before the coarse step and after it remains the same, then the rate of convergence is the same. NO
- If m-level IAD is locally convergent then $(m+1)$-level IAD is locally convergent. NO
- If $(m+1)$-level IAD is locally convergent then m-level IAD is locally convergent. NO
- If B symmetric and $TB = BT$ then m-level IAD is locally convergent.

Multi-level IAD with $m > 2$ levels

Error propagation formula available.

Consider these hypotheses:

- If the number of basic iteration is increased, then the convergence is faster. NO
- If the sum of the numbers of basic iterations before the coarse step and after it remains the same, then the rate of convergence is the same. NO
- If m-level IAD is locally convergent then $(m+1)$-level IAD is locally convergent. NO
- If $(m+1)$-level IAD is locally convergent then m-level IAD is locally convergent. NO
- If B symmetric and $TB = BT$ then m-level IAD is locally convergent.

Multi-level IAD with $m > 2$ levels

Error propagation formula available.

Consider these hypotheses:

- If the number of basic iteration is increased, then the convergence is faster. NO
- If the sum of the numbers of basic iterations before the coarse step and after it remains the same, then the rate of convergence is the same. NO
- If m-level IAD is locally convergent then $(m+1)$-level IAD is locally convergent. NO
- If $(m+1)$-level IAD is locally convergent then m-level IAD is locally convergent. NO
- If B symmetric and $TB = BT$ then m-level IAD is locally convergent. ???

Conclusions

- Necessary and sufficient conditions for local convergence of IAD for $T = B$, $\nu = 1$.
- Local convergence depends ONLY on non-zero pattern of B for $T = \alpha B + (1 - \alpha)I$ and $\nu = 1$, **but not** in the case of $\nu > 1$,
  and **not** in the case of three or more levels.
- Symmetric B, T leads to local convergence if $TB = BT$ and $\nu \geq 1$, i.e. $\rho(J(x)) < 1$.
  Non-symmetric $B \in \mathcal{R}^{N \times N}$ can cause $\rho(J(x)) > N/4$.
- No relation between convergence of m-level and $(m + 1)$-level IAD methods.

Open questions

- Does B, T symmetric, $TB = BT$ yield local convergence also for multi-level IAD?
- How many cumulative points can a sequence of x^k have? Finite number of them?
- **Main question.** Does local convergence always imply global convergence?

”. . . The improvement of AMG schemes is a hot research topic. . . . ”

. . . Yvan Notay, An aggregation-based multigrid method, ETNA, 2010