ELA

ERROR PROPAGATION FORMULA OF MULTI-LEVEL ITERATIVE AGGREGATION-DISAGGREGATION METHODS FOR NON-SYMMETRIC PROBLEMS*

IVANA PULTAROVÁ†

Abstract. Iterative aggregation-disaggregation methods for numerical computing of stationary probability distribution vectors of stochastic matrices are studied. The methods can use arbitrary numbers of levels and of smoothing steps. A formula for the error propagation is derived. Using this formula, some asymptotic convergence properties of these methods for non-symmetric problems are demonstrated.

Key words. Iterative aggregation-disaggregation methods, Multi-level methods, Markov chains, Stationary probability distribution vector.

AMS subject classifications. 65F08, 65F10.

1. Introduction. Multi-level solution methods are considered as an advantageous technique in numerical mathematics. Multi-level and multi-grid methods are popular mainly in numerical solutions of partial differential equations. In this paper we study iterative aggregation-disaggregation (IAD) methods for computing stationary probability distribution vectors of finite discrete Markov chains. A growing interest in this field is motivated by many large scale applications in information technology, data mining, biology, safety and reliability computation, and other theoretical and practical disciplines.

Basics of the IAD methods are found in Stewart's book [9]. Many papers show an efficient performance of the IAD methods, see e.g. [1, 2, 3] and the references therein. However, theoretical results are rare: only some convergence conditions for two-level algorithms [4, 5, 6, 8, 9] and for a special kind of multi-level algorithms [7] are available.

It appears that for non-symmetric problems, the multi-level methods do not possess the appropriate properties as in the case of symmetric problems [8]. Our paper focuses on some rules which could be expected for the IAD methods, but which do

ELA

not hold. As a main tool, we introduce a new formula for an error propagation matrix of the IAD methods with an arbitrary number of levels. While in our recent paper [7] we started this study by the derivation of an error propagation formula for the IAD methods without pre-smoothing steps, in this paper we deal with more general IAD methods with arbitrary numbers of pre-smoothing steps and of post-smoothing steps in every level.

The paper is organized as follows. In Section 2, the notation and description of the IAD method are introduced. Section 3 is the main part of the paper. Since a formula for the error propagation matrix of a general multi-level IAD method is rather complicated, we present it only for a three-level method with an arbitrary number of smoothing steps (Theorem 3.1) and for a method with an arbitrary number of levels but with exactly one pre-smoothing step and one post-smoothing step in every level (Theorem 3.2). Simple examples are presented in Section 4 to confirm the absence of several relations which could seem natural for the multi-level IAD methods. Conclusion and one open question are found in Section 5.

2. Notation and description of the IAD methods. Let us assume that B is an irreducible stochastic matrix. This means that B has non-negative elements, the sum of the elements in each column is equal to unity, and there does not exist any symmetric permutation of the columns and rows of B such that the resulting permuted matrix $\tilde{B}$ admits a block form

$$\tilde{B} = \left[\begin{array}{cc} \tilde{B}_{11} & \tilde{B}_{12} \\ 0 & \tilde{B}_{22} \end{array} \right],$$

where the diagonal blocks $\tilde{B}_{11}$ and $\tilde{B}_{22}$ are square. Let e denote a column vector of ones of an appropriate size. From the Perron-Frobenius theorem [10], B has a unique eigenvalue of modulus 1 and there exists a unique positive eigenvector $\hat{x}$ for which $B\hat{x} = \hat{x}$ and $e^T\hat{x} = 1$. Vector $\hat{x}$ is called a stationary probability distribution vector of the matrix B. Let us define $P = \hat{x}e^T$ and $Z = B - P$. It is obvious that $P^2 = P$ and that $PZ = ZP$ is a null matrix.

Many iterative methods can be exploited for computing the eigenvector $\hat{x}$, for example, the power method, Jacobi method, Gauss-Seidel method or their block modifications. We will study a multi-level approach which is called an iterative aggregation-disaggregation (IAD) method.

Let us denote by L the number of levels used within the method. The original (finest, largest) problem belongs to level $m = 1$ and the coarsest (smallest) problem is solved in level $m = L$. In every level up to the coarsest one, aggregation groups G_{mj} of indices of elements are chosen, where $m = 1, 2, \ldots, L - 1$ and $j = 1, 2, \ldots, N_{m+1}$. Thus N_{m+1} is the number of aggregation groups in level m and at the same time it is

ELA

the number of elements in level $m+1$ for $m = 1, 2, \ldots, L-1$. The size of the original matrix B is N_1 and the size of the coarsest problem solved is N_L.

Without loss of generality, we suppose that in every level $m = 1, 2, \ldots, L-1$, the indices are divided into groups in such a manner that from $i_1 \in G_{k_1}$, $i_2 \in G_{k_2}$ and $i_1 < i_2$, it follows that $k_1 < k_2$. In other words, smaller indices fall into groups with smaller indices.

Let us denote by $[A]_{ij}$ an element of a matrix A in row i and column j. Transitions between levels are enabled by reduction and prolongation matrices. Reduction matrices R_m map from $\mathbb{R}^{N_m}$ into $\mathbb{R}^{N_{m+1}}$,

$$[R_m]_{ij} = \begin{cases} 1 & \text{if } j \in G_{mi} \\ 0 & \text{otherwise.} \end{cases}$$

Prolongation matrices $S(x)_m$ map from $\mathbb{R}^{N_{m+1}}$ into $\mathbb{R}^{N_m}$ depending on a positive vector $x \in \mathbb{R}^{N_m}$,

$$[S(x)_m]_{ij} = \begin{cases} \frac{x_i}{\sum_{k \in G_{mj}} x_k} & \text{if } i \in G_{mj}, \\ 0 & \text{otherwise.} \end{cases}$$

In every level $m = 1, 2, \ldots, L-1$, some smoothing steps are performed before and after the solution of the coarser problem of level $m+1$. The smoothing steps, also called basic iterations, are represented by the power method in our algorithm.

A single cycle of the multi-level IAD algorithm can be described as follows. A constant τ is a threshold for the size of problems the solution of which is carried out exactly.

> *Procedure IAD* (input: B, x; output: y)
> if $size(B) < \tau$ solve $By = y$, $e^T y = 1$;
> else
> $\quad x := B^\mu x$;
> $\quad$ build R and $S(x)$ according to aggregation groups in a current level;
> $\quad$ call *Procedure IAD* (input: $RBS(x)$, Rx; output: y);
> $\quad y := B^\nu S(x)y$.

In the main program the procedure is first called with the input parameters B, the original irreducible non-negative matrix, and x as an initial approximation to $\hat{x}$. One cycle of the IAD method is represented by one call of this procedure in the original level. After every cycle, y is assigned to x and next cycle is performed. The procedure is executed until the difference between x and y in the finest level is sufficiently small.

Let μ_m and ν_m denote the numbers of smoothing steps in level m before and after the coarsening of the problem, respectively. Using this notation we can describe

ELA

Ivana Pultarová

a single multi-level cycle of the IAD algorithm. Let x_n be a current approximation to the solution $\hat{x}$. Then the next cycle of the IAD procedure yields x_{n+1} in the form

$$x_{n+1} = B^{\nu_1} S(u_1)_1 v_2,$$

where

$$u_1 = B^{\mu_1} x_n,$$
$$u_m = \left(R_{m-1} \cdots R_1 B S(u_1)_1 \cdots S(u_{m-1})_{m-1} \right)^{\mu_m} R_{m-1} u_{m-1},$$
$$v_m = \left(R_{m-1} \cdots R_1 B S(u_1)_1 \cdots S(u_{m-1})_{m-1} \right)^{\nu_m} S(u_m)_m v_{m+1}$$

for $m = 2, 3, \ldots, L - 1$. Vector v_L is the exact solution of the coarsest problem $T_L x = x$, where T_L is

$$T_L = R_{L-1} \cdots R_1 B S(u_1)_1 \cdots S(u_{L-1})_{L-1}.$$

The vectors u_k are computed during the coarsening process (the first half of a single multi-level IAD cycle) while the vectors v_k are computed during the returning back to finer levels (the second half of a single IAD cycle).

Let the identity matrix be denoted by I. Let $T = A_1^{-1} A_2$ come from some regular splitting [10] of $A_1 - A_2 = I - B$, $A_1^{-1} \geq 0$, $A_2 \geq 0$. Thus $T\hat{x} = \hat{x}$. In addition, let $TB = BT$. Then the iteration matrices in levels $m = 1, 2, \ldots, L - 1$

$$B_m = R_{m-1} \cdots R_1 B S(u_1)_1 \cdots S(u_{m-1})_{m-1}$$

can be substituted by the more general

$$T_m = R_{m-1} \cdots R_1 T S(u_1)_1 \cdots S(u_{m-1})_{m-1}.$$

This generalization will be used in theorems in Section 3. The matrix T can vary during the computation process. However, for simplicity's sake, we will not consider variable T, and, in our numerical examples, only $T = B$ will be used. In the coarsest level, $T \neq B$ could be used, but then the definitions of matrices R_m and $S(x)_m$ would need to be changed. This is why we have used only $T_L = B_L$ in the coarsest level.

LEMMA 2.1. *The iteration matrices*

$$T_m = R_{m-1} \cdots R_1 T S(u_1)_1 \cdots S(u_{m-1})_{m-1},$$

$m = 1, 2, \ldots, L$, *are non-negative and irreducible. If $T = B$ then T_m is stochastic.*

Proof. We have $T \geq 0$ and thus $T_m \geq 0$ yields from the definitions of R_m and of $S(x)_m$. The irreducibility follows from the irreducibility of T. From $e^T B = e^T$ we obtain $e^T B_m = e^T$. ☐

ELA

Let us denote

$$P(u_k, u_{k+1}, \ldots, u_{m-1})_{km} = S(u_k)_k \cdots S(u_{m-1})_{m-1} R_{m-1} \cdots R_k,$$

where u_j are positive vectors and $u_j \in \mathbb{R}^{N_j}$, $j = k, \ldots, m-1$. During the proofs in Section 3 we will use the abbreviation $P_m \equiv P(u_1, u_2, \ldots, u_{m-1})_{1m}$ and $P_1 = I$.

LEMMA 2.2. *The following two relations hold.*

1. *For any positive $x \in \mathbb{R}^{N_k}$ we have $S(x)_k R_k x = x$.*
2. *Matrices $P(u_k, u_{k+1}, \ldots, u_{m-1})_{km}$ are projections, for $1 \le k < m \le L$.*

Proof. The first part follows from the definition of R_k and $S(x)_k$. The second part follows from $R_j S(x)_j = I$, $j = 1, \ldots, L-1$. $\square$

3. Error propagation formulae. In this section we introduce our main results. We derive the error propagation matrices of the IAD methods described in Section 2. Due to the complexity of the involved terms, we restrict ourselves to two cases: (i) three levels of hierarchy, $L = 3$, and arbitrary numbers of pre-smoothing steps and of post-smoothing steps, i. e. arbitrary μ_m and ν_m in both levels $m = 1, 2$; and (ii) arbitrary number L of levels and exactly one pre-smoothing step and one post-smoothing step in every level, i. e. $\mu_m = \nu_m = 1$, $m = 1, 2, \ldots, L-1$. Deriving the error propagation matrices for these two kinds of methods within proofs of the following two theorems shows the main ideas and tools which can be used for deducing error propagation formulae for more complicated methods.

THEOREM 3.1. *Let us consider a three-level IAD method, $L = 3$, with $\mu_m + \nu_m \ge 1$, $m = 1, 2$. The error in cycle n is*

$$x_{n+1} - \hat{x} = J(x_n)(x_n - \hat{x}).$$

Let us denote $u_1 = T^{\mu_1} x_n$, $u_2 = \left(R_1 T S(u_1)_1\right)^{\mu_2} R_1 u_1$, $P_1 = I$, $P_2 = S(u_1)_1 R_1$ and $P_3 = S(u_1)_1 S(u_2)_2 R_2 R_1$.

1. *If $\mu_2 \ge 1$ and $\nu_2 \ge 1$ then*

$$J(x_n) = T^{\nu_1} \left(\left(P_2 T\right)^{\nu_2} \left(I - P_3 Z\right)^{-1} \left(\left(P_2 - P_3\right) \sum_{k=0}^{\mu_2-1} \left(T P_2\right)^k (T - I) + I - P_3 \right) \right.$$

$$\left. + \sum_{k=0}^{\nu_2-1} \left(P_2 T\right)^k \left(I - P_2\right) \right) T^{\mu_1}.$$

2. *If $\mu_2 \ge 1$ and $\nu_2 = 0$ then*

$$J(x_n) = T^{\nu_1} \left(I - P_3 Z\right)^{-1} \left(\left(P_2 - P_3\right) \sum_{k=0}^{\mu_2-1} \left(T P_2\right)^k (T - I) + I - P_3 \right) T^{\mu_1}.$$

Ivana Pultarová

3. *If $\mu_2 = 0$ and $\nu_2 \geq 1$ then*

$$J(x_n) = T^{\nu_1}\left(\left(P_2 T\right)^{\nu_2}\left(I - P_3 Z\right)^{-1}\left(I - P_3\right) + \sum_{k=0}^{\nu_2-1}\left(P_2 T\right)^{k}\left(I - P_2\right)\right)T^{\mu_1}.$$

Proof. We start with a proof of the first assertion for $\mu_2 \geq 1$ and $\nu_2 \geq 1$. Let us follow the computation during a single cycle of the considered IAD method. First u_1 is computed, $u_1 = T^{\mu_1} x_n$. Then this vector is reduced and multiplied by $\left(R_1 T S(u_1)_1\right)^{\mu_2}$, and thus

$$u_2 = \left(R_1 T S(u_1)_1\right)^{\mu_2} R_1 u_1 = \left(R_1 T S(u_1)_1\right)^{\mu_2} R_1 T^{\mu_1} x_n.$$

In the third level, the computation of the eigenvector v_3 of the reduced matrix is carried out exactly for $T = B$,

$$R_2 R_1 B S(u_1)_1 S(u_2)_2 v_3 = v_3,$$

and $e^T v_3 = 1$. None of the eigenvalues of $R_2 R_1 Z S(u_1)_1 S(u_2)_2$ are equal to unity, and

$$v_3 = \left(I - R_2 R_1 Z S(u_1)_1 S(u_2)_2\right)^{-1} R_2 R_1 \hat{x}.$$

On the way back to the finest level, the vector v_3 is prolonged by a multiplication by $S(u_2)_2$ to level $m = 2$ and then multiplied by the iteration matrix $\left(R_1 T S(u_1)_1\right)^{\nu_2}$ in level $m = 2$,

$$v_2 = \left(R_1 T S(u_1)_1\right)^{\nu_2} S(u_2)_2 v_3.$$

And finally,

$$x_{n+1} = v_1 = T^{\nu_1} S(u_1)_1 v_2.$$

Thus we have

$$\begin{aligned}
x_{n+1} - \hat{x} &= T^{\nu_1} S(u_1)_1 \left(R_1 T S(u_1)_1\right)^{\nu_2} S(u_2)_2 \left(I - R_2 R_1 Z S(u_1)_1 S(u_2)_2\right)^{-1} R_2 R_1 \hat{x} - \hat{x} \\
&= T^{\nu_1}\left(P_2 T\right)^{\nu_2}\left(I - P_3 Z\right)^{-1} P_3 \hat{x} - \hat{x}.
\end{aligned}$$

Now we can rewrite the term $P_3 \hat{x}$ using Lemma 2.2,

$$\begin{aligned}
P_3 \hat{x} &= S(u_1)_1 S(u_2)_2 R_2 \left(R_1 \hat{x} - u_2 + u_2\right) \\
&= S(u_1)_1 u_2 + S(u_1)_1 S(u_2)_2 R_2 \left(R_1 \hat{x} - u_2\right) \\
&= S(u_1)_1 \left(R_1 T S(u_1)_1\right)^{\mu_2} R_1 T^{\mu_1} x_n + S(u_1)_1 S(u_2)_2 R_2 \big(R_1 \hat{x} \\
&\quad - \left(R_1 T S(u_1)_1\right)^{\mu_2} R_1 T^{\mu_1} x_n\big) \\
&= P_2\left(T P_2\right)^{\mu_2} T^{\mu_1} x_n + P_3\left(\hat{x} - \left(T P_2\right)^{\mu_2} T^{\mu_1} x_n\right).
\end{aligned}$$

ELA

Using $P_2 T^{\mu_1} x_n = T^{\mu_1} x_n$ and $T\hat{x} = \hat{x}$, the term $(TP_2)^{\mu_2} T^{\mu_1} x_n$ is

$$(3.1) \quad \begin{aligned} (TP_2)^{\mu_2} T^{\mu_1} x_n &= (TP_2)^{\mu_2-1}\big(TP_2 T^{\mu_1} x_n - TT^{\mu_1}\hat{x} + T^{\mu_1}\hat{x} \\ &\quad - T^{\mu_1} x_n + T^{\mu_1} x_n\big) \\ &= (TP_2)^{\mu_2-1}\big(TT^{\mu_1} x_n - TT^{\mu_1}\hat{x} + T^{\mu_1}\hat{x} - T^{\mu_1} x_n + T^{\mu_1} x_n\big) \\ &= (TP_2)^{\mu_2-1}(T-I)T^{\mu_1}\big(x_n - \hat{x}\big) + (TP_2)^{\mu_2-1} T^{\mu_1} x_n \\ &= \cdots \\ &= \Big((TP_2)^{\mu_2-1} + \cdots + TP_2\Big)(T-I)T^{\mu_1}\big(x_n - \hat{x}\big) \\ &\quad + TT^{\mu_1}\big(x_n - \hat{x}\big) + \hat{x}, \end{aligned}$$

which yields

$$P_3\hat{x} = \big(P_2 - P_3\big)\Big(\big((TP_2)^{\mu_2-1} + \cdots + TP_2 + I\big)(T-I) + I\Big)T^{\mu_1}\big(x_n - \hat{x}\big) + P_2\hat{x}.$$

We also have

$$\begin{aligned} P_2\hat{x} = S(u_1)_1 R_1 \hat{x} = S(u_1)_1 R_1\big(\hat{x} - u_1 + u_1\big) = u_1 + P_2\big(\hat{x} - u_1\big) = \\ = T^{\mu_1} x_n + P_2 T^{\mu_1}\big(\hat{x} - x_n\big) = (I - P_2)T^{\mu_1}\big(x_n - \hat{x}\big) + \hat{x}. \end{aligned}$$

Then

$$(3.2) \quad \begin{aligned} x_{n+1} - \hat{x} &= T^{\nu_1}(P_2 T)^{\nu_2}\big(I - P_3 Z\big)^{-1}\Big(\big(P_2 - P_3\big)\Big(\big((TP_2)^{\mu_2-1} + \cdots \\ &\quad \cdots + TP_2 + I\big)(T-I) + I\Big)T^{\mu_1}\big(x_n - \hat{x}\big) \\ &\quad + (I - P_2)T^{\mu_1}\big(x_n - \hat{x}\big) + \hat{x}\Big) - \hat{x} \\ &= T^{\nu_1}(P_2 T)^{\nu_2}\big(I - P_3 Z\big)^{-1}\Big(\big(P_2 - P_3\big)\Big(\big((TP_2)^{\mu_2-1} + \cdots \\ &\quad \cdots + TP_2 + I\big)(T-I) + I\Big) + (I - P_2)\Big)T^{\mu_1}\big(x_n - \hat{x}\big) \\ &\quad + T^{\nu_1}(P_2 T)^{\nu_2}\hat{x} - \hat{x}. \end{aligned}$$

Exploiting similar tools as in (3.1) we can show that

$$(3.3) \quad T^{\nu_1}(P_2 T)^{\nu_2}\hat{x} - \hat{x} = T^{\nu_1}\big((P_2 T)^{\nu_2-1} + \cdots + I\big)(I - P_2)T^{\mu_1}(x_n - \hat{x}).$$

Combining (3.2) and (3.3) we get the first part of the theorem. The second one can be easily derived using a similar technique. Using either similar tools as for the previous parts or Theorem 2 of paper [7] yields the third part of the theorem. □

THEOREM 3.2. *The error in cycle n of a multi-level IAD method with an arbitrary number $L \geq 2$ of levels and with one pre-smoothing step and one post-smoothing step in every level (that is $\mu_m = \nu_m = 1$ for $m = 1, 2, \ldots, L-1$) is*

$$x_{n+1} - \hat{x} = J\left(x_n\right)\left(x_n - \hat{x}\right),$$

ELA

where

(a)

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

$$+ T \sum_{m=1}^{L-2} \prod_{k=2}^{m}(P_k T) \sum_{k=1}^{m}(P_k - P_{k+1})M_{k-1};$$

(b) $M_0 = T$ *and*

$$M_k = \big(T + \sum_{j=2}^{k} TP_j(T - I)\big)T,$$

for $k = 1, 2, \ldots, L - 2$;

(c) $P_1 = I$ *and*

$$P_k = P(u_1, u_2, \ldots, u_{k-1})_{1k} = S(u_1)_1 \cdots S(u_{k-1})_{k-1} R_{k-1} \cdots R_1,$$

for $k = 2, 3, \ldots, L$;

(d) $u_1 = Tx_n$, $u_2 = R_1 T^2 x_n$, $u_3 = R_2 R_1 T P_2 T^2 x_n$ *and*

$$u_k = R_{k-1} \cdots R_1 T P_{k-1} T P_{k-2} \cdots T P_3 T P_2 T^2 x_n,$$

for $k = 4, \ldots, L - 1$.

Proof. We present the derivation of the formula only for $L = 5$ which is not too much complicated but provides with a sufficiently general insight. We have

$$u_1 = Tx_n,$$
$$u_2 = R_1 T S(u_1) R_1 u_1 = R_1 T u_1 = R_1 T^2 x_n,$$
$$u_3 = R_2 R_1 T S(u_1)_1 S(u_2)_2 R_2 u_2 = R_2 R_1 T S(u_1)_1 u_2 = R_2 R_1 T S(u_1)_1 R_1 T^2 x_n$$
$$\quad = R_2 R_1 T P_2 T^2 x_n,$$
$$u_4 = R_3 R_2 R_1 T S(u_1)_1 S(u_2)_2 S(u_3)_3 R_3 u_3 = \cdots = R_3 R_2 R_1 T P_3 T P_2 T^2 x_n.$$

Since from Lemma 2.2 we have $S(u_k)_k R_k u_k = u_k$, $k = 1, \ldots, 4$, note that

(3.4) $$P_2 T x_n = T x_n,$$
(3.5) $$P_3 T^2 x_n = P_2 T^2 x_n,$$
(3.6) $$P_4 T P_2 T^2 x_n = P_3 T P_2 T^2 x_n,$$
(3.7) $$P_5 T P_3 T P_2 T^2 x_n = P_4 T P_3 T P_2 T^2 x_n.$$

ELA

In the fifth (finest) level, the computation of the eigenvector is carried out exactly with $T = B$,

$$R_4 R_3 R_2 R_1 B S(u_1)_1 S(u_2)_2 S(u_3)_3 S(u_4)_4 v_5 = v_5,$$

$e^T v_5 = 1$. Matrix $I - R_4 \cdots R_1 Z S(u_1)_1 \cdots S(u_4)_4$ is invertible, therefore

$$v_5 = \left(I - R_4 \cdots R_1 Z S(u_1)_1 \cdots S(u_4)_4\right)^{-1} R_4 \cdots R_1 \hat{x}.$$

Further computation in finer levels gradually yields

$$v_4 = R_3 R_2 R_1 T S(u_1)_1 S(u_2)_2 S(u_3)_3 S(u_4)_4 v_5,$$
$$v_3 = R_2 R_1 T S(u_1)_1 S(u_2)_2 S(u_3)_3 v_4,$$
$$v_2 = R_1 T S(u_1)_1 S(u_2)_2 v_3,$$
$$x_{n+1} = v_1 = T S(u_1)_1 v_2.$$

Then

(3.8) $$x_{n+1} - \hat{x} = T P_2 T P_3 T P_4 T \left(I - P_5 Z\right)^{-1} P_5 \hat{x} - \hat{x}.$$

Using (3.4)-(3.7) we have

$$P_5 \hat{x} = P_5(\hat{x} - T P_3 T P_2 T^2 x_n) + P_4 T P_3 T P_2 T^2 x_n$$

and

$$T P_3 T P_2 T^2 x_n = \left(T P_3 T P_2 (T - I) T + T P_2 (T - I) T + T^2\right)(x_n - \hat{x}) + \hat{x} = M_3(x_n - \hat{x}) + \hat{x},$$

where $M_3 = T P_3 T P_2 (T - I) T + T P_2 (T - I) T + T^2$. Thus

$$P_5 \hat{x} = (P_4 - P_5) M_3 (x_n - \hat{x}) + P_4 \hat{x}.$$

Computing $P_4 \hat{x}$ and $P_3 \hat{x}$ in a similar fashion, we obtain

(3.9) $$P_4 \hat{x} = (P_3 - P_4) M_2 (x_n - \hat{x}) + P_3 \hat{x},$$
(3.10) $$P_3 \hat{x} = (P_2 - P_3) M_1 (x_n - \hat{x}) + P_2 \hat{x},$$
(3.11) $$P_2 \hat{x} = (I - P_2) M_0 (x_n - \hat{x}) + \hat{x},$$

where $M_2 = T P_2 (T - I) T + T^2$, $M_1 = T^2$ and $M_0 = T$. Thus we have

$$P_5 \hat{x} = \left((P_4 - P_5) M_3 + (P_3 - P_4) M_2 + (P_2 - P_3) M_1 + (I - P_2) M_0\right)(x_n - \hat{x}) + \hat{x}.$$

Substituting this into (3.8), denoting $P_1 = I$ and using $Z \hat{x} = 0$, we get

(3.12) $$x_{n+1} - \hat{x} = T P_2 T P_3 T P_4 T \left(I - P_5 Z\right)^{-1} \left(\left((P_4 - P_5) M_3 + (P_3 - P_4) M_2\right.\right.$$

$$\left.\left. + (P_2 - P_3) M_1 + (P_1 - P_2) M_0\right)(x_n - \hat{x}) + \hat{x}\right) - \hat{x}$$

$$= T P_2 T P_3 T P_4 T \left(I - P_5 Z\right)^{-1} \left(\sum_{k=1}^{4} (P_k - P_{k+1}) M_{k-1}\right)(x_n - \hat{x})$$

$$+ T P_2 T P_3 T P_4 T \hat{x} - \hat{x}.$$

ELA

Finally, using (3.9)-(3.11), we obtain

$$
\begin{aligned}
TP_2TP_3TP_4T\hat{x} - \hat{x} = {}& TP_2TP_3T\Big((P_3 - P_4)M_2 + (P_2 - P_3)M_1 \\
&+ (P_1 - P_2)M_0\Big)(x_n - \hat{x}) \\
&+ TP_2T\Big((P_2 - P_3)M_1 + (P_1 - P_2)M_0\Big)(x_n - \hat{x}) \\
&+ T(P_1 - P_2)M_0(x_n - \hat{x}).
\end{aligned}
$$

Substituting this last formula into (3.12) completes the proof. $\square$

EXAMPLE 3.3. Assume a three-level IAD method, $L = 3$, and three simple choices of parameters μ_k and ν_k. According to both Theorems 3.1 and 3.2, for $\mu_1 = \mu_2 = \nu_1 = \nu_2 = 1$ the error propagation matrix J is

$$
(3.13) \qquad J = TP_2T(I - P_3Z)^{-1}\big((P_2 - P_3)T + I - P_2\big)T + T(I - P_2)T.
$$

When we chose $\mu_1 = \mu_2 = 0$ and $\nu_1 = \nu_2 = 2$, from Theorem 3.1 we obtain

$$
(3.14) \qquad J = T^2\big((P_2T)^2(I - P_3Z)^{-1}(I - P_3) + (I + P_2T)(I - P_2)\big).
$$

When we chose $\mu_1 = \mu_2 = 2$ and $\nu_1 = \nu_2 = 0$, from Theorem 3.1 we have

$$
(3.15) \qquad J = (I - P_3Z)^{-1}\big((P_2 - P_3)(I + TP_2)(T - I) + I - P_3\big)T^2.
$$

4. Symmetric and non-symmetric matrices. In this part we will exploit the 8×8 non-symmetric irreducible stochastic matrix

$$
(4.1) \qquad B =
\begin{bmatrix}
\epsilon & 1 & 0 & 0 & 0 & 0 & 0 & \beta \\
0 & 0 & \delta + \zeta & 1 & \alpha & \delta & 0 & 0 \\
1 & 0 & 0 & 0 & 0 & 0 & 0 & \alpha + \gamma \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\
\zeta & 0 & \beta & 0 & 1 & \gamma & 0 & 0 \\
\delta & 0 & 0 & 0 & 0 & 0 & 1 & \epsilon
\end{bmatrix}.
$$

We apply the multi-level IAD methods to the matrix B for several choices of parameters α, β, γ, δ, ϵ and ζ. After setting the values of these parameters, the columns are normalized in such a way that the resulting matrix is stochastic.

EXAMPLE 4.1. Consider the matrix B given by (4.1), $L = 3$, $N_1 = 8$, $N_2 = 4$, $N_3 = 2$ and the dimensions of aggregation groups equal to N_m/N_{m+1} in every level m.

ELA

TABLE 4.1

Spectral radii $\rho(J(\hat{x}))$ for $L = 3$, $N_1 = 8$, $N_2 = 4$, $N_3 = 2$ and for matrix B defined by (4.1) with different choices of α, β, γ.

$(\delta, \epsilon, \zeta) = (0,0,0)$, (α, β, γ):	$(10,0,0)$	$(0,10,0)$	$(0,0,10)$
$\mu_k = 1$, $\nu_k = 1$	0.2447	1.1144	0.3414
$\mu_k = 0$, $\nu_k = 2$	1.1419	0.2916	0.3513
$\mu_k = 2$, $\nu_k = 0$	0.5069	0.3676	1.0729

TABLE 4.2

Spectral radii $\rho(J(\hat{x}))$ for either $L = 2$ or $L = 3$ and $\mu_k = \nu_k = 1$ for matrix B defined by (4.1) with different choices of δ, ϵ, ζ.

$(\alpha, \beta, \gamma) = (0,0,0)$, $(\delta, \epsilon, \zeta)$:	$(10,0,0)$	$(0,10,0)$	$(0,0,10)$
$L = 2$, $N_2 = 4$	0.0789	2.0463	0.1289
$L = 2$, $N_2 = 2$	0.9090	0.4092	1.5757
$L = 3$, $N_2 = 4$, $N_3 = 2$	1.3390	0.3448	0.6861

We examine three choices of parameters μ_k, ν_k, $k = 1, 2$: (i) $\mu_k = \nu_k = 1$; (ii) $\mu_k = 0$, $\nu_k = 2$; (iii) $\mu_k = 2$, $\nu_k = 0$. See Table 4.1 for the approximate asymptotic spectral radii $\rho(J(\hat{x}))$ of the error propagation matrices for $\delta = \epsilon = \zeta = 0$ and for different choices of α, β, γ. The spectral radii are calculated according to (3.13)-(3.15). In our paper [7] the hypothesis appeared, that keeping the number of smoothing steps constant in all levels, i.e. $\mu_k + \nu_k$ constant for every k, leads to the same asymptotic convergence rate. It can be seen from the table that this hypothesis is false. Moreover, the convergence itself is also not kept for different μ_k, ν_k, but constant $\mu_k + \nu_k$.

EXAMPLE 4.2. In this example we examine the IAD methods for the matrix B given by (4.1) where $\alpha = \beta = \gamma = 0$. We compare three types of methods: (i) a two-level method with $N_2 = 4$; (ii) a two-level method with $N_2 = 2$; (iii) a three-level method with $N_2 = 4$; $N_3 = 2$. The numbers of smoothing steps are "symmetric": $\mu_k = \nu_k = 1$. Three choices of parameters δ, ϵ, ζ lead to three different situations, when one of the methods diverges and the other ones converge locally. Thus there is no apparent relation between local convergence of the IAD algorithms of the same kind with different numbers of levels.

REMARK 4.3. When a stochastic matrix B is symmetric or satisfies $B = DSD^{-1}$ where S is symmetric and D is diagonal, then all of our numerical experiments have reflected that $\rho(J(\hat{x})) \leq 1$. Although the proof of this is known [8] for two-level IAD methods, the statement for general multi-level IAD methods has not been proved yet.

5. Conclusion. New convergence characteristics of the multi-level IAD methods are introduced in this paper. Especially, we present an error propagation formula which is a generalization of our recent result [7]. Using the formula it can be confirmed

ELA

20   Ivana Pultarová

that the local convergence of the multi-level IAD methods is not guaranteed for non-symmetric problems in general.

The derivation of the error propagation formula can serve as a model: the tools of the proofs of Theorems 3.1 and 3.2 can be exploited to obtain error propagation formulae for more complicated multi-level IAD methods. For example, the formulae for arbitrary number of levels L and for arbitrary numbers of smoothing steps μ_k and ν_k in every level $k = 1, 2, \ldots, L-1$ can be derived. This can lead to the estimation of local convergence rates of particular methods for particular matrices which can help in searching general conditions under which the local convergence of the IAD methods for non-symmetric problems is guaranteed.

In paper [7], a new hypothesis was announced, that the asymptotic rate of convergence does not depend on the numbers of pre- and post-smoothing steps but depends solely on the sum of these two quantities. Example 4.1 disproves it. In Example 4.2 we give a counter-example of another hypothesis, which states that for $\mu_k = \nu_k$, $k = 1, 2, \ldots, L-1$, there exists some relation between the local convergence of IAD with L levels and with $L+1$ levels.

Our numerical experiments support a new hypothesis: if a stochastic matrix B is symmetric or similar to a symmetric matrix S such that $B = DSD^{-1}$ and D is diagonal, then the asymptotic spectral radius of the error propagation matrix is bounded from above by unity for an arbitrary number of levels and of smoothing steps and for any choice of aggregation groups. This hypothesis was proved for two-level IAD methods only [8], but for general multi-level IAD methods it has not been proved yet.

REFERENCES

[1] P. Buchholz and T. Dayar. On the convergence of a class of multilevel methods for large sparse Markov chains. *SIAM Journal on Matrix Analysis and Applications*, 29:1025–1049, 2007.

[2] T. Dayar. State space orderings for Gauss-Seidel in Markov chains revisited. *SIAM Journal on Scientific Computing*, 19:148–154, 1998.

[3] H. De Sterck, K. Miller, E. Treister and I. Yavneh. Fast multilevel methods for Markov chains. To appear in *Numerical Linear Algebra with Applications*.

[4] I. C. F. Ipsen and S. Kirkland. Convergence analysis of a PageRank updating algorithm by Langville and Meyer. *SIAM Journal on Matrix Analysis and Applications*, 27:952–967, 2006.

[5] I. Marek and P. Mayer. Convergence theory of some classes of iterative aggregation-

ELA

disaggregation methods for computing stationary probability vectors of stochastic matrices. *Linear Algebra and Its Applications*, 363:177–200, 2003.

[6] I. Pultarová. Necessary and sufficient local convergence condition of one class of iterative aggregation - disaggregation methods. *Numerical Linear Algebra with Applications*, 15:339–354, 2008.

[7] I. Pultarová and I. Marek. Convergence of multi-level iterative aggregation-disaggregation methods. *Journal of Computational and Applied Mathematics*, 236:354-363, 2011.

[8] I. Pultarová and I. Marek. Physilogy and and pathology of iterative aggregation-disaggregation methods. *Numerical Linear Algebra with Applications*, 18:1051-1065, 2011.

[9] W. J. Stewart. *Introduction to the Numerical Solution of Markov Chains*. Princeton University Press, Princeton, 1994.

[10] R. S. Varga. *Matrix Iterative Analysis*. 2nd ed., Springer-Verlag, New York, 2000.