Applications of Mathematics

Stanislav Míka; Petr Vaněk

Acceleration of convergence of a two-level algebraic algorithm by aggregation in smoothing process

Applications of Mathematics, Vol. 37 (1992), No. 5, 343–356

Persistent URL: `http://dml.cz/dmlcz/104515`

37 (1992) APPLICATIONS OF MATHEMATICS No. 5, 343–356

ACCELERATION OF CONVERGENCE OF A TWO-LEVEL ALGEBRAIC ALGORITHM BY AGGREGATION IN SMOOTHING PROCESS

STANISLAV MÍKA, PETR VANĚK

(Received January 10, 1990)

Summary. A two-level algebraic algorithm is introduced and its convergence is proved. The restriction as well as prolongation operators are defined with the help of aggregation classes. Moreover, a particular smoothing operator is defined in an analogical way to accelarate the convergence of the algorithm. A model example is presented in conclusion.

Keywords: aggregation class, two-level algorithm, convergence factor, smoothing operator, linear algebraic system

AMS classification: 65F10

1. INTRODUCTION

Let $\mathbf{A} = (a_{ij})_{i,j=1}^{n}$ be a matrix of the order n and let us construct the matrix $\mathbf{A}^G = (a_{ij}^G)_{i,j=1}^{n}$ with elements

$$a_{ij}^G = \begin{cases} 1 & \text{if } a_{ij} \neq 0, \\ 0 & \text{if } a_{ij} = 0. \end{cases} \tag{1.1}$$

The matrix $\mathbf{A}^G$ will be called the structure matrix of $\mathbf{A}$. Further, we construct the undirected graph $G = (U, H)$ determined by $\mathbf{A}$ so that the number of the knots $u_i \in U$ is n, i.e. $i = 1, 2, \ldots, n$, and the set of edges

$$H \subset \big\{\{u_i, u_j\} : u_i, u_j \in U\big\}$$

is determined by defining the structure matrix $\mathbf{A}^G$ of $\mathbf{A}$ to be the incidence matrix of the graph $G = (U, H)$. It means that $\{u_i, u_j\}$, $u_i, u_j \in U$ is an edge if and only if $a_{ij}^G = 1$ (i.e. if $a_{ij} \neq 0$).

The set

$$O(u_i) = \{u_j \in U : \{u_i, u_j\} \in H\}$$

will be called the knot neigbourhood of $u_i \in U$. Let us suppose $a_{ii} \neq 0$, $i = 1, 2, \ldots, n$. Therefore $u_i \in O(u_i)$.

2. Decomposition of a set of knots into aggregation classes

Let $G = (U, H)$ be the graph of a matrix $\mathbf{A}$. We employ the set U to construct the system of aggregation classes $\{T_k\}_{k=1}^m$ which is defined by the following properties:

(2.1)
1. $\forall u_i \in U \quad \exists k : u_i \in T_k$.
2. If $T_j \cap T_k \neq \emptyset$, then $T_j = T_k$.
3. If $u_p, u_q \in T_k$, then there exists $u_i \in U$ such that $u_p, u_q \in O(u_i)$.

Remark. The order n of $\mathbf{A}$ is equal to the number of knots in the set U. The number m of aggregation classes, $m < n$, determines the order of the matrix corresponding to a "coarse" grid.

The decomposition of U into the aggregation classes T_i is performed by the following algorithm (U_i denotes the set of those knots from U which have not been included into any aggregation class in the i-th step).

(2.2) U, H is given

1. $U_1 := U$, $i := 1$.
2. If $X = \{u_j \in U_i : O(u_j) \subset U_i\} \neq \emptyset$, then choose $u_k \in X$, where k is the least integer in X.
3. If $X = \emptyset$, then choose $u_k \in U_i$, where k is the least integer in U_i.
4. $T_i := O(u_k) \cap U_i$.
5. $U_{i+1} := U_i - T_i$.
6. $i := i + 1$.
7. If $U_i \neq \emptyset$, proceed to step 2.
8. END.

Lemma 2.1. *The algorithm* (2.2) *generates a system of aggregation classes satisfying conditions* (2.1).

Proof consists in the verification of (2.1).

If the algorithm ends, then $U_i = \emptyset$ and thus the condition 1 is fulfilled. In step 5 one subtracts a nonempty subset from a finite set. Therefore, the algorithm ends after a finite number of steps. The inclusion $T_i \subset O(u_k)$ is valid for some k [see step 4 of (2.2)], which implies the condition 3 of (2.1). We now verify the property 2 of (2.1). Suppose there is $u \in U$ such that $u \in T_j \cap T_k$ where $j \leqslant k$. Since $U_{j+1} = U_j - T_j$ (step 5 of (2.2)), then $u \notin U_i$ for $i > j$. But $T_i \subset U_i$ and $u \notin U_i$ *imply* $u \notin T_i$, $i > j$. Thus we must have T_k, $k \leqslant j$ such that $u \in T_k$. Therefore $k = j$. □

3. Transfer operators

Let V_n and V_m be linear vector spaces ($n = \dim V_n$ = order of the matrix $\mathbf{A}$, $m = \dim V_m$ = the number of the aggregation classes). We define the matrix $\mathbf{r} = (r_{ij})$ of type (m, n) as follows:

$$r_{ij} = \begin{cases} 1 & \text{for } u_j \in T_i, \\ 0 & \text{for } u_j \notin T_i. \end{cases} \tag{3.1}$$

This matrix represents a linear operator $\mathbf{r}\colon V_n \to V_m$ which will be called a restriction operator. The prolongation operator $\mathbf{p}\colon V_m \to V_n$ is defined by transposing

$$\mathbf{p} = \mathbf{r}^T. \tag{3.2}$$

The matrix $\mathbf{r}$ has the full rank, because there is exactly one unity in each of its columns (the system of aggregation classes is a disjoint decomposition of the set U).

Let $\mathbf{x} \in V_n$. The vector $\mathbf{rx} \in V_m$ is given by its components

$$(\mathbf{rx})_k = \sum_{j \in J} x_j, \quad J = \{j\colon u_j \in T_k\}, \quad k = 1, 2, \ldots, m. \tag{3.3}$$

For $\mathbf{y} \in V_m$, the components of $\mathbf{py} \in V_n$ are given by

$$(\mathbf{py})_l = y_r, \quad l = 1, 2, \ldots, r, \tag{3.4}$$

where r is the index of the class that contains the knot u_l.

Given matrices $\mathbf{A}_1, \mathbf{p}, \mathbf{r}$, ($\mathbf{A}_1 = \mathbf{A}$), we define the aggregation matrix $\mathbf{A}_2$ (of order m) by

$$\mathbf{A}_2 = \mathbf{r}\mathbf{A}_1\mathbf{p}. \tag{3.5}$$

From the relations (3.5), (3.1), (3.2) we obtain

$$(a_2)_{ij} = \sum_{s,t} (a_1)_{st}, \quad i, j = 1, 2, \ldots, m, \tag{3.6}$$

where the summation is over those s and t for which $u_s \in T_i$, $u_t \in T_j$.

Remark. Every knot belongs to a single aggregation class if and only if there is one unity in every column of $\mathbf{A}_1$. Every aggregation class contains at least one knot if and only if there is at least one unity in every row of $\mathbf{A}_1$.

4. Two-level algorithm

Suppose that $\mathbf{A}_1$ is symmetric and positive definite. We want to solve the linear system of equations

$$\mathbf{A}_1\mathbf{x} = \mathbf{f}, \quad \mathbf{f} \in V_n. \tag{4.1}$$

One iteration of our algorithm ($\mathbf{x}_i \to \mathbf{x}_{i+1}$) is defined as follows:

(4.2)
1. $\tilde{\mathbf{x}}_i := \mathbf{S}_1\mathbf{x}_i + \mathbf{T}_1\mathbf{f}, \quad \tilde{\mathbf{x}}_i \in V_n$ (pre-smoothing step).
2. $\mathbf{d} := \mathbf{A}_1\tilde{\mathbf{x}}_i - \mathbf{f}, \quad \mathbf{d} \in V_n$.
3. $\mathbf{v} := (\mathbf{A}_2)^{-1}\mathbf{rd}$ (the solution of $\mathbf{A}_2\mathbf{v} = \mathbf{rd}, \quad \mathbf{v} \in V_m$).
4. $\tilde{\tilde{\mathbf{x}}}_i := \tilde{\mathbf{x}}_i - \mathbf{pv}, \quad \tilde{\tilde{\mathbf{x}}}_i \in V_m$ (correction step).
5. $\mathbf{x}_{i+1} := \mathbf{S}_2\tilde{\tilde{\mathbf{x}}}_i + \mathbf{T}_2\mathbf{f}, \quad \mathbf{x}_{i+1} \in V_n$ (post-smoothing step).

We assume that the consistence conditions $\mathbf{I} = \mathbf{T}_1\mathbf{A}_1 + \mathbf{S}_1$, $\mathbf{I} = \mathbf{T}_2\mathbf{A}_1 + \mathbf{S}_2$ are fulfilled, where $\mathbf{S}_1$, $\mathbf{S}_2$ are regular iteration matrices.

5. Properties of errors of the two-level method

Let $(\mathbf{x}, \mathbf{y})_n$ be the Euclidean inner product of the vectors $\mathbf{x}, \mathbf{y} \in V_n$. Let $\mathbf{A}_1$ be a symmetric and positive definite matrix of order n. The energy norm of $\mathbf{x} \in V_n$ is defined as

$$\|\mathbf{x}\|_A = (\mathbf{A}_1\mathbf{x}, \mathbf{x})^{1/2}. \tag{5.1}$$

The matrix norm is then defined by

$$\|\cdot\|_A = \sup_{\substack{\mathbf{x}\in V_n \\ \mathbf{x}\neq 0}} \frac{\|\cdot \mathbf{x}\|_A}{\|\mathbf{x}\|_A}$$

Let $\dot{\mathbf{x}} \in V_m$ be a solution of the system $\mathbf{A}_1\mathbf{x} = f$ and suppose that $\tilde{\mathbf{x}}_i \in V_n$ is the result of the pre-smoothing step and $\tilde{\tilde{\mathbf{x}}}_i \in V_n$ is the result of the correction step.

If $\mathbf{v} \in V_m$ is the solution of the residual equation $\mathbf{A}_2\mathbf{v} = \mathbf{rd}$, i.e. $\mathbf{v} = (\mathbf{A}_2)^{-1}\mathbf{rd}$, then we have (see step 4 of (4.2))

$$\tilde{\tilde{\mathbf{x}}}_i - \hat{\mathbf{x}} = \left[\mathbf{I} - \mathbf{p}(\mathbf{rA}_1\mathbf{p})^{-1}\mathbf{rA}_1\right](\tilde{\mathbf{x}}_i - \hat{\mathbf{x}}).$$

In this paper we use the following notation:

$$\mathbf{e}_i = \mathbf{x}_i - \hat{\mathbf{x}},\ \mathbf{e} = \tilde{\mathbf{x}}_i - \hat{\mathbf{x}},\ \mathbf{e}(\mathbf{v}) = \tilde{\tilde{\mathbf{x}}}_i - \hat{\mathbf{x}},\ \mathbf{K} = \mathbf{I} - \mathbf{p}(\mathbf{rA}_1\mathbf{p})^{-1}\mathbf{rA}_1.$$

Then we have

$$\mathbf{e} = \mathbf{S}_1\mathbf{e}_i,\ \mathbf{e}(\mathbf{v}) = \mathbf{e} - \mathbf{pv} = \mathbf{Ke},\ \mathbf{d} = \mathbf{A}_1\mathbf{e},\ \mathbf{e}_{i+1} = \mathbf{S}_2\mathbf{KS}_1\mathbf{e}_i.$$

For any $\mathbf{w} \in V_m$ we put $\mathbf{e}(\mathbf{w}) = \mathbf{e} - \mathbf{pw}$.

Lemma 5.1. $\|\mathbf{e}(\mathbf{v})\|_A = \min\limits_{\mathbf{w}\in V_m} \|\mathbf{e}(\mathbf{w})\|_A$ *if and only if* $\mathbf{v} \in V_m$ *solves the residual equation* $\mathbf{A}_2\mathbf{v} = \mathbf{rd}$.

Proof. We want to show that the conditions

$$\|\mathbf{e}(\mathbf{v})\|_A \leqslant \|\mathbf{e}(\mathbf{w})\|_A \quad \forall \mathbf{w} \in V_m, \tag{5.2}$$

and

$$\mathbf{A}_2\mathbf{v} = \mathbf{rd} \tag{5.3}$$

are equivalent.

We write any vector $\mathbf{w} \in V_m$ in the form $\mathbf{w} = \mathbf{v} + t\Delta\mathbf{v}$, where $\mathbf{v} \in V_m$ is fixed and $\Delta\mathbf{v} \in V_m$, $t \in R^1$ are arbitrary. The functional $\varphi\colon V_m \to R^1$ given by

$$\varphi(\mathbf{w}) = \|\mathbf{e}(\mathbf{w})\|_A^2 = \big(\mathbf{A}_1\mathbf{e}(\mathbf{w}), \mathbf{e}(\mathbf{w})\big) \tag{5.4}$$

is strictly convex. The condition

$$\frac{\mathrm{d}}{\mathrm{d}t}\varphi(\mathbf{v} + t\Delta\mathbf{v})\,|_{t=0} = 0 \quad \forall \Delta\mathbf{v} \in V_m$$

is a necessary and sufficient condition for $\mathbf{v}$ to be a unique minimizer of φ. After elementary manipulations this conditions can be written in the form

$$\big(\mathbf{A}_1(\mathbf{e} - \mathbf{pv}), \mathbf{p}\Delta\mathbf{v}\big) = 0 \tag{5.5}$$

or in the form ($\mathbf{r}\colon V_n \to V_m$)

$$\big(\mathbf{rA}_1(\mathbf{e} - \mathbf{pv}), \Delta\mathbf{v}\big) = 0 \quad (\text{in } V_m). \tag{5.6}$$

Hence $\mathbf{rA}_1(\mathbf{e} - \mathbf{pv}) = 0$ in V_m. From this we obtain

$$\mathbf{rA}_1\mathbf{pv} = \mathbf{rA}_1\mathbf{e} = \mathbf{rd}. \tag{5.7}$$

Therefore (5.2) implies (5.3).
On the other hand, (5.3) and (3.5) imply (5.5) and hence (5.2). □

Lemma 5.3. *The matrix* $\mathbf{A}_2 = \mathbf{rA}_1\mathbf{p}$ *is symmetric and positive definite.*

Proof is trivial.

6. Convergence of two-level algorithm

Lemma 6.1. *The iteration method given by algorithm* (4.2) *can be expressed by the iteration formula*

$$\mathbf{x}_{i+1} := \mathbf{M}\mathbf{x}_i + \mathbf{N}\mathbf{f}, \tag{6.1}$$

where

$$\mathbf{M} = \mathbf{S}_2\big[\mathbf{I} - \mathbf{p}(\mathbf{A}_2)^{-1}\mathbf{r}\mathbf{A}_1\big]\mathbf{S}_1 = \mathbf{S}_2\mathbf{K}\mathbf{S}_1,$$
$$\mathbf{N} = \mathbf{S}_2\mathbf{T}_1 + \mathbf{S}_2\mathbf{p}(\mathbf{A}_2)^{-1}\mathbf{r}(\mathbf{I} - \mathbf{A}_1\mathbf{T}_1) + \mathbf{T}_2,$$

and $\mathbf{S}_1$ *is a pre-smoothing matrix,* $\mathbf{S}_2$ *is a post-smoothing matrix.*

Proof is trivial.

Lemma 6.2. *The vector* $\tilde{\mathbf{x}} = (\mathbf{A}_1)^{-1}\mathbf{f}$ *is a fixed point of* (6.1).

P r o o f. From (4.2) and from the consistence conditions $\mathbf{I} = \mathbf{T}_1\mathbf{A}_1 + \mathbf{S}_1$, $\mathbf{I} = \mathbf{T}_2\mathbf{A}_1 + \mathbf{S}_2$ we obtain $\mathbf{I} = \mathbf{M} + \mathbf{N}\mathbf{A}_1$. The sets

$$\mathrm{Im}(\mathbf{p}) = \{\mathbf{z} \in V_n : \mathbf{z} = \mathbf{p}\mathbf{w} \quad \forall \mathbf{w} \in V_m\}$$
$$T = \{\chi \in V_n : (\mathbf{A}_1\chi, \mathbf{p}\mathbf{w}) = 0 \quad \forall \mathbf{w} \in V_m\}$$

are $\mathbf{A}_1$-orthogonal subspaces of V_n, i.e.

$$T = \mathrm{Ker}(\mathbf{r}\mathbf{A}_1) = \big(\mathrm{Im}(\mathbf{p})\big)^{\perp} \tag{6.2}$$

and we can write

$$V_n = T \oplus \mathrm{Im}(\mathbf{p}).$$

It can be seen from (5.5) that $\mathbf{K}\mathbf{e} \in T$, where $\mathbf{K}\mathbf{e} = \mathbf{e} - \mathbf{p}\mathbf{v}$. □

Lemma 6.3. *The following error estimate is valid:*

$$\frac{\|\mathbf{S}_2\mathbf{K}\mathbf{e}\|_A}{\|\mathbf{e}\|_A} \leqslant \sup_{\substack{\chi\in T\\ \chi\neq 0}} \frac{\|\mathbf{S}_2\chi\|_A}{\|\chi\|_A}. \tag{6.3}$$

P r o o f. From (5.2) for $\mathbf{w} = 0$ (i.e. $\mathbf{e}(\mathbf{w}) = \mathbf{e}$) we obtain

$$\|\mathbf{K}\mathbf{e}\|_A \leqslant \|\mathbf{e}\|_A, \quad \mathbf{e} \in V_n. \tag{6.4}$$

For $\mathbf{K}\mathbf{e} \in T$ we have

$$\frac{\|\mathbf{S}_2\mathbf{K}\mathbf{e}\|_A}{\|\mathbf{K}\mathbf{e}\|_A} \leqslant \sup_{\substack{\chi\in T\\ \chi\neq 0}} \frac{\|\mathbf{S}_2\chi\|_A}{\|\chi\|_A}.$$

Using (6.4) we obtain (6.3). □

Theorem 1. *Assume that* $\|\mathbf{S}_1\|_A \leqslant q_1 < 1$, $\|\mathbf{S}_2\|_A \leqslant q_2 < 1$. *Then the two-level iterative process* (4.2) *converges and the rate of convergence can be estimated as*

$$\frac{\|\mathbf{e}_{i+1}\|_A}{\|\mathbf{e}_i\|_A} \leqslant \|\mathbf{S}_1\|_A \sup_{\substack{\chi \in T \\ \chi \neq 0}} \frac{\|\mathbf{S}_2 \chi\|_A}{\|\chi\|_A} \leqslant q_1 q_2 < 1. \tag{6.5}$$

Proof. For an arbitrary $\mathbf{e}_i \in V_n$ and a regular $\mathbf{S}_1$, $\mathbf{e} = \mathbf{S}_1\mathbf{e}_i$ is an arbitrary vector in V_n, too. From (6.4) we obtain

$$\|\mathbf{K}\| \leqslant 1 \quad (\mathbf{K}\colon V_n \to T).$$

From the obvious inequality

$$\|\mathbf{e}_i\|_A \geqslant \frac{\|\mathbf{S}_1 \mathbf{e}_i\|_A}{\|\mathbf{S}_1\|_A}$$

we get

$$\frac{\|\mathbf{e}_{i+1}\|_A}{\|\mathbf{e}_i\|_A} = \frac{\|\mathbf{S}_1 \mathbf{K} \mathbf{e}\|_A}{\|\mathbf{e}_i\|_A} \leqslant \|\mathbf{S}_1\|_A \frac{\|\mathbf{S}_2 \mathbf{K} \mathbf{e}\|_A}{\|\mathbf{S}_1 \mathbf{e}_i\|_A}.$$

We now may use (6.3) to obtain (6.5). □

7. Estimate of the asymptotic convergence factor by Brandt's technique

Lemma 7.1. *The correction operator* $\mathbf{K} = \mathbf{I} - \mathbf{p}(\mathbf{T}\mathbf{A}_1\mathbf{p})^{-1}\mathbf{r}\mathbf{A}_1$ *(see Lemma 6.1) is an identity operator on the space* T, *i.e.* $\mathbf{K}\chi = \chi \ \forall \chi \in T$.

Proof. From (6.2) we obtain

$$(\mathbf{r}\mathbf{A}_1\chi, \mathbf{w}) = 0 \quad \forall \mathbf{w} \in V_m$$

therefore

$$\mathbf{r}\mathbf{A}_1\chi = 0, \quad \chi \in T.$$

Now it can be seen that

$$\mathbf{K}\chi = [\mathbf{I} - \mathbf{p}(\mathbf{r}\mathbf{A}_1\mathbf{p})^{-1}\mathbf{r}\mathbf{A}_1]\chi = \chi - \mathbf{p}(\mathbf{r}\mathbf{A}_1\mathbf{p})^{-1}0 = \chi.$$

□

Lemma 7.2. *Let* $\mathbf{S}_2 = \mathbf{I} - \omega \mathbf{A}_1$ *(the damped Jacobi method) and suppose that* $0 < \omega \leqslant \frac{2}{\|\mathbf{A}_1\|_2}$. *Further suppose that for each* $\chi \in T$ *there is* $\mathbf{w} \in V_m$ *such that the condition*

$$\|\chi\|_A \geqslant C\|\chi - \mathbf{p}\mathbf{w}\|_2 \tag{7.2}$$

with $C > 0$ *is satisfied. Then the following error estimate is valid:*

$$\frac{\|\mathbf{S}_2\chi\|_A^2}{\|\chi\|_A^2} \leqslant 1 - \omega C^2(2 - \omega\|\mathbf{A}_1\|_2). \tag{7.3}$$

Here, we denote

$$\|\mathbf{x}\|_2 = (\mathbf{x}, \mathbf{x})^{1/2} \qquad \textit{for } \mathbf{x} \in V_n, \ \|\mathbf{A}_1\|_2 = \sup_{\substack{\mathbf{x}\in V_n \\ \mathbf{x}\neq 0}} \frac{\|\mathbf{A}\mathbf{x}\|_2}{\|\mathbf{x}\|_2}.$$

Proof. Using the equality

$$\|\mathbf{S}_2\chi\|_A^2 = (\mathbf{A}_1(\mathbf{I} - \omega\mathbf{A}_1)\chi, \ (\mathbf{I} - \omega\mathbf{A}_1)\chi) = \|\chi\|_A^2 + \omega^2\|\mathbf{A}_1\chi\|_A^2 - 2\omega\|\mathbf{A}_1\chi\|_2^2$$

and the inequality

$$\|\mathbf{A}_1\chi\|_A^2 \leqslant \|\mathbf{A}_1\|_2 \cdot \|\mathbf{A}_1\chi\|_2^2$$

we obtain

$$\|\mathbf{S}_2\chi\|_A^2 \leqslant \|\chi\|_A^2 - \omega\|\mathbf{A}_1\chi\|_2^2(2 - \omega\|\mathbf{A}_1\|_2). \tag{7.4}$$

Every vector $\chi \in T$ is $\mathbf{A}_1$-orthogonal to $\mathrm{Im}(p)$, i.e. $(\mathbf{A}_1\chi, \mathbf{p}\mathbf{w}) = 0 \ \forall \mathbf{w} \in V_m$. Thus $(\mathbf{A}_1\chi, \chi) = (\mathbf{A}_1\chi, \chi - \mathbf{p}\mathbf{w})$ (Brandt's technique [1]) and therefore

$$\|\chi\|_A^2 \leqslant \|\mathbf{A}_1\chi\|_2 \cdot \|\chi - \mathbf{p}\mathbf{w}\|_2.$$

Multiplying by C and using (7.2) we obtain the estimate $C\|\chi\|_A \leqslant \|\mathbf{A}_1\chi\|_2$. From the assumption $0 < \omega \leqslant \frac{2}{\|\mathbf{A}_1\|_2}$ and from (7.4) we then obtain

$$\|\mathbf{S}_2\chi\|_A^2 \leqslant \|\chi\|_A^2 - \omega C^2\|\chi\|_A^2(2 - \omega\|\mathbf{A}_1\|_2).$$

□

Theorem 2. *Let the assumptions of Lemma 7.2 be satisfied. Then the rate of convergence of the two-level method with* $\mathbf{S}_2 = \mathbf{I} - \omega\mathbf{A}_1$ *can be estimated as*

$$\frac{\|\mathbf{e}_{i+1}\|_A^2}{\|\mathbf{e}_i\|_A^2} \leqslant \|\mathbf{S}_1\|_A^2 \left[1 - \omega C^2(2 - \omega\|\mathbf{A}_1\|_2)\right]. \tag{7.5}$$

Proof follows immediately from Lemma 7.2 and Theorem 1.

We will now generalize the results of Theorem 2 to a two-level method with the smoothing operator $\mathbf{S}_2 = \mathbf{I} - \omega\mathbf{D}^{-1}\mathbf{A}_1$, where $\mathbf{D}$ is a symmetric positive definite matrix.

Lemma 7.4. *Let* $\mathbf{S}_2 = \mathbf{I} - \omega\mathbf{D}^{-1}\mathbf{A}_1$, $\omega > 0$ *(the generalized Jacobi method) and suppose that for each* $\chi \in T$ *there exists* $\mathbf{w} \in V_m$ *such that the condition*

$$\|\chi\|_A \geqslant C_1\|\chi - \mathbf{p}\mathbf{w}\|_D \tag{7.6}$$

with $C_1 > 0$ *is satisfied and the condition*

$$\|\mathbf{D}^{-1}\mathbf{A}_1\chi\|_A \leqslant C_2\|\mathbf{D}^{-1}\mathbf{A}_1\chi\|_D, \qquad \omega C_2^2 < 2 \tag{7.7}$$

with $C_2 > 0$ *is satisfied. Then*

$$\frac{\|\mathbf{S}_2\chi\|_A^2}{\|\chi\|_A^2} \leqslant 1 - \omega C_1^2(2 - \omega C_2^2). \tag{7.8}$$

Proof. Using the evident relation $(\mathbf{A}_1\chi, \mathbf{D}^{-1}\mathbf{A}_1\chi) = (\mathbf{D}\mathbf{D}^{-1}\mathbf{A}_1\chi, \mathbf{D}^{-1}\mathbf{A}_1\chi) = \|\mathbf{D}^{-1}\mathbf{A}_1\chi\|_D^2$, $\chi \in T$ we get $\|\mathbf{S}_2\chi\|_A^2 = \|\chi - \omega\mathbf{D}^{-1}\mathbf{A}_1\chi\|_A^2 = \|\chi\|_A^2 + \omega^2\|\mathbf{D}^{-1}\mathbf{A}_1\chi\|_A^2 - 2\omega\|\mathbf{D}^{-1}\mathbf{A}_1\chi\|_D^2$. Using now (7.7) we have

$$\|\mathbf{S}_2\chi\|_A^2 \leqslant \|\chi\|_A^2 + \omega^2 C_2^2\|\mathbf{D}^{-1}\mathbf{A}_1\chi\|_D^2 - 2\omega\|\mathbf{D}^{-1}\mathbf{A}_1\chi\|_D^2. \tag{7.9}$$

Since $\chi \in T$, i.e. $(\mathbf{A}_1\chi, \mathbf{p}\mathbf{w}) = 0$ $\forall \mathbf{w} \in V_m$, and since the condition (7.6) is valid, we successively obtain $\|\chi\|_A^2 = (\mathbf{A}_1\chi, \chi) = (\mathbf{A}_1\chi, \chi - \mathbf{p}\mathbf{w}) = (\mathbf{D}^{-1}\mathbf{A}_1\chi, \mathbf{D}(\chi - \mathbf{p}\mathbf{w})) = (\mathbf{D}^{1/2}\mathbf{D}^{-1}\mathbf{A}_1\chi, \mathbf{D}^{1/2}(\chi - \mathbf{p}\mathbf{w})) \leqslant \|\mathbf{D}^{-1}\mathbf{A}_1\chi\|_D\|\chi - \mathbf{p}\mathbf{w}\|_D \leqslant \|\mathbf{D}^{-1}\mathbf{A}_1\chi\|_D\frac{1}{C_1}\|\chi\|_A$. Therefore

$$C_1\|\chi\|_A \leqslant \|\mathbf{D}^{-1}\mathbf{A}_1\chi\|_D. \tag{7.10}$$

Combining (7.9) and (7.10) we obtain (7.8). □

Theorem 3. *Let the assumptions of Lemma 7.4 be satisfied. Then the rate of convergence of the two-level method with* $\mathbf{S}_2 = \mathbf{I} - \omega\mathbf{D}^{-1}\mathbf{A}_1$ *can be estimated as*

$$\frac{\|\mathbf{e}_{i+1}\|_A^2}{\|\mathbf{e}_i\|_A^2} \leqslant \|\mathbf{S}_1\|_A^2 \left[1 - \omega C_1^2(2 - \omega C_2^2)\right]. \tag{7.11}$$

Proof follows from Lemma 7.4 and Theorem 1.

Remark. The assumptions $\|\mathbf{S}_1\|_A \leqslant q_1 < 1$, $\|\mathbf{S}_2\|_A \leqslant q_2 < 1$ of Theorem 1 are not necessary to prove Theorems 2, 3. Hence, the smoothing need not be a convergent process for a two-level method to converge.

8. Generalized Jacobi method determined by the system of aggregation classes

We shall briefly describe the construction of the positive definite matrix $\mathbf{D}$ in the generalized Jacobi method with the iteration matrix $\mathbf{I} - \omega\mathbf{D}^{-1}\mathbf{A}_1$ by the system of aggregation classes.

Let us have the system $\{T_k\}_{k=1}^m$ of the aggregation classes generated by the algorithm (2.2). For each class T_k we construct a diagonal matrix $\mathbf{V}_k$ of order n with elements $v_{ij}^{(k)}$ such that

$$v_{ij}^{(k)} = 0 \text{ for } i \neq j, \qquad v_{ii}^{(k)} = \begin{cases} 1 & \text{if } u_i \in T_k, \\ 0 & \text{if } u_i \notin T_k, \end{cases} \tag{8.1}$$

where u_i, $i = 1, 2, \ldots, n$ is the i-th knot of the graph G introduced in Sec. 1.

We define matrices $\mathbf{D}_k$, $k = 1, 2, \ldots, m$ of order n by

$$\mathbf{D}_k = \mathbf{V}_k\mathbf{A}_1\mathbf{V}_k \tag{8.2}$$

and we put

$$\mathbf{D} = \sum_{k=1}^{m} \mathbf{D}_k. \tag{8.3}$$

Now we want to determine the constant C_2 in the condition (7.7) to get the particular estimate of the convergence factor (7.11) for this matrix $\mathbf{D}$.

Remark. From (8.2) it is clear that $\mathbf{D}$ is a symmetric and positive definite matrix.

Lemma 8.1. *Let* $\mathbf{B}$ *be a matrix of order* n. *Then for the matrix*

$$\tilde{\mathbf{D}} = \mathbf{D}^{-1} + \mathbf{Bpr} \tag{8.4}$$

and for any $\chi \in T$ *we have*

$$\tilde{\mathbf{D}}\mathbf{A}_1\chi = \mathbf{D}^{-1}\mathbf{A}_1\chi. \tag{8.5}$$

Proof. $\tilde{\mathbf{D}}\mathbf{A}_1\chi = \mathbf{D}^{-1}\mathbf{A}_1\chi + \mathbf{BprA}_1\chi$. From the definition of T we have $(\mathbf{A}_1\chi, \mathbf{pw}) = 0\ \forall \mathbf{w} \in V_n$, i.e. $(\mathbf{rA}_1\chi, \mathbf{w}) = 0$. Therefore $\mathbf{rA}_1\chi = 0$ on T. □

Theorem 4. *Let* $\tilde{\mathbf{D}}_1$ *and* $\tilde{\mathbf{D}}_2$ *be matrices which can be expressed in the form* (8.4) *with matrices* $\mathbf{B}_1$ *and* $\mathbf{B}_2$, *respectively, and let all the eigenvalues* $\lambda(\tilde{\mathbf{D}}_1)$ *and* $\lambda(\tilde{\mathbf{D}}_2)$ *be real. If*

$$\|\mathbf{A}_1\|_2 \cdot \lambda^2_{\max}(\tilde{\mathbf{D}}_1) \leqslant C_2^2 \cdot \lambda_{\min}(\tilde{\mathbf{D}}_2) \tag{8.6}$$

then the inequality (7.7) *holds.*

Proof. For $\chi \in T$, (8.5) yields

$$\begin{aligned}
\|\mathbf{D}^{-1}\mathbf{A}_1\chi\|_A^2 &= (\mathbf{A}_1\mathbf{D}^{-1}\mathbf{A}_1\chi, \mathbf{D}^{-1}\mathbf{A}_1\chi) = (\mathbf{A}_1\tilde{\mathbf{D}}_1\mathbf{A}_1\chi, \tilde{\mathbf{D}}_1\mathbf{A}_1\chi)\\
&\leqslant \|\mathbf{A}_1\|_2(\tilde{\mathbf{D}}_1\mathbf{A}_1\chi, \tilde{\mathbf{D}}_1\mathbf{A}_1\chi) \leqslant \|\mathbf{A}_1\|_2 \cdot \lambda^2_{\max}(\tilde{\mathbf{D}}_1)\|\mathbf{A}_1\chi\|_2^2,\\
\|\mathbf{D}^{-1}\mathbf{A}_1\chi\|_D^2 &= (\mathbf{D}\mathbf{D}^{-1}\mathbf{A}_1\chi, \mathbf{D}^{-1}\mathbf{A}_1\chi) = (\tilde{\mathbf{D}}_2\mathbf{A}_1\chi, \mathbf{A}_1\chi)\\
&\geqslant \lambda_{\min}(\tilde{\mathbf{D}}_2)\|\mathbf{A}_1\chi\|_2^2.
\end{aligned}$$

Using the inequality (8.6) we obtain

$$C_2^2\lambda_{\min}(\tilde{\mathbf{D}}_2)\|\mathbf{A}_1\chi\|_2^2 \geqslant \|\mathbf{A}_1\|_2\lambda^2_{\max}(\tilde{\mathbf{D}}_1)\|\mathbf{A}_1\chi\|_2^2,$$

which implies (7.7) □

9. A MODEL EXAMPLE

We consider the two-level algorithm (4.2) with $\mathbf{S}_1 = \mathbf{I}$, $\mathbf{S}_2 = \mathbf{I} - \omega\mathbf{D}^{-1}\mathbf{A}_1$, where the matrix $\mathbf{A}_1$ was obtained by applying the usual finite-difference method to the boundary value problem

$$\begin{aligned} -u'' &= f \qquad \text{on } (0,1),\\ u(0) &= a, \quad u(1) = b. \end{aligned}$$

We use a uniform grid of $2m$ interior points with a step h.

The matrix $\mathbf{A}_1$ of order $2m$ is then of the well-known form

$$\mathbf{A}_1 = \frac{1}{h^2}\begin{bmatrix} 2 & -1 & & & 0\\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1\\ 0 & & & -1 & 2 \end{bmatrix}.$$

The system of the aggregation classes is defined as follows:

$$T_k = \{u_{2k-1}, u_{2k}\}, \quad k = 1, 2, \ldots, m,$$

where the knots u_i are defined by the interior points of the grid. The matrix $\mathbf{D}$ will be block-diagonal

$$\mathbf{D} = \frac{1}{h^2}\cdot\left[\begin{array}{cc|cc|c|cc} 2 & -1 & & & & & \\ -1 & 2 & & & & & \\ \hline & & 2 & -1 & & & \\ & & -1 & 2 & & & \\ \hline & & & & \ddots & & \\ \hline & & & & & 2 & -1\\ & & & & & -1 & 2 \end{array}\right].$$

Analogously, the $(2,2)$-blocks of the block-diagonal matrix $\mathbf{D}^{-1}$ are

$$\frac{h^2}{3}\begin{bmatrix} 2 & 1\\ 1 & 2 \end{bmatrix}.$$

The matrix **pr** has the same structure, its $(2,2)$-blocks being $\begin{bmatrix} 1 & 1\\ 1 & 1 \end{bmatrix}$.

We now show that a special construction of the matrices $\tilde{\mathbf{D}}_1$, $\tilde{\mathbf{D}}_2$ from Sec. 8 allows us to give an estimate of the rate of convergence (7.11) with particular values of C_1, C_2.

Put $\mathbf{B} = -\frac{h^2}{3}\mathbf{I}$, i.e. $\tilde{\mathbf{D}} = \mathbf{D}^{-1} - \frac{h^2}{3}\cdot\mathbf{pr} = \frac{h^2}{3}\mathbf{I}$, and put $\tilde{\mathbf{D}}_1 = \tilde{\mathbf{D}}_2 = \tilde{\mathbf{D}}$. Hence we have $\lambda_{\min}(\tilde{\mathbf{D}}) = \lambda_{\max}(\tilde{\mathbf{D}}) = \frac{h^2}{3}$. Because $\lambda_{\max}(\mathbf{A}_1) = \|\mathbf{A}_1\|_2$, the inequality (8.6) can be written in the form

$$\lambda_{\max}(\mathbf{A}_1)\left(\frac{h^2}{3}\right)^2 \leqslant C_2^2\,\frac{h^2}{3},$$

or in the form

$$\frac{h^2}{3}\lambda_{\max}(\mathbf{A}_1) \leqslant C_2^2.$$

From Gershgorin's theorem we have $\lambda_{\max}(\mathbf{A}_1) \leqslant \frac{4}{h^2}$ and hence we can put $C_2^2 = \frac{4}{3}$.

Our first attempt is to find an estimate with the constant C_1 (see (7.6) and (7.11)). We can write (7.6) in the form

$$C_1^2(\mathbf{D}(\chi - p\mathbf{w}), \chi - \mathbf{pw}) \leqslant (\mathbf{A}_1\chi, \chi). \tag{9.1}$$

We want to find C_1 such that for every $\chi \in T$ there exists $\mathbf{w} \in V_m$ such that (9.1) holds.

With respect to (7.11) it is desirable that C_1 be as large as possible (then the error after the correction step will be much less then the error after the smoothing step alone).

Therefore, we will minimize $\big(\mathbf{D}(\chi - \mathbf{pw}), \chi - \mathbf{pw}\big)$. This inner product will be evidently minimal, if

$$(\mathbf{D}(\chi - \mathbf{pw}), \mathbf{p}\Delta\mathbf{w}) = 0 \quad \forall \Delta\mathbf{w} \in V_m. \tag{9.2}$$

It can be easily verified that this condition on the matrix $\mathbf{D}$ is satisfied if we take

$$\mathbf{w} = \tfrac{1}{2}\mathbf{r}\chi.$$

Then we have $\big(\chi = (\chi_1, \chi_2, \ldots, \chi_n)\big)$

$$\big(\mathbf{D}(\chi - \tfrac{1}{2}\mathbf{pr}\chi), \chi - \tfrac{1}{2}\mathbf{pr}\chi\big) = \frac{1}{h^2}\Big[\frac{3}{2}\sum_{i=1}^{m}(\chi_{2i} - \chi_{2i-1})^2\Big],$$

$$(\mathbf{A}_1\chi, \chi) = \frac{1}{h^2}\Big[\chi_1^2 + \chi_{2m}^2 + \sum_{i=1}^{2m-1}(\chi_{i+1} - \chi_i)^2\Big].$$

Therefore the inequality (9.1) (with $\mathbf{w} = \frac{1}{2}\mathbf{r}\chi$), i.e.

$$C_1^2(\mathbf{D}(\chi - \tfrac{1}{2}\mathbf{pr}\chi), \chi - \tfrac{1}{2}\mathbf{pr}\chi) \leqslant (\mathbf{A}_1\chi, \chi)$$

will be fulfilled if $\frac{3}{2}C_1^2 \leqslant 1$. The biggest C_1^2 satisfying the last inequality is $C_1^2 = \frac{2}{3}$.

Substituting $C_1^2 = \frac{2}{3}$, $C_2^2 = \frac{4}{3}$ into (7.11) we obtain

$$\frac{\|\mathbf{e}_{i+1}\|_A^2}{\|\mathbf{e}_i\|_A^2} \leqslant 1 - \frac{2}{3}\omega\Big(2 - \frac{4}{3}\omega\Big).$$

We can now readily find that the optimum value of the damping factor ω is $\omega = \frac{3}{4}$ (the minimizer of the quadratic function $1 - \frac{2}{3}\omega(2 - \frac{4}{3}\omega)$). For this ω we arrive at the estimate

$$\frac{\|\mathbf{e}_{i+1}\|_A^2}{\|\mathbf{e}_i\|_A^2} \leqslant \frac{1}{2}.$$

Remark. Numerical experiments performed with a TGM algorithm using the aggregation according to (2.2) have shown an interesting property of the generalized Jacobi method of Sec. 8: the rate of convergence of the two-level method used almost did not decrease with an increasing number of knots in the aggregation classes. In other words, to realize the correction on the coarse grid it would be possible to employ aggregation matrices of an essentially smaller order. This property could be useful when solving large scale, in particular 3D-problems. A heuristic exploitation of this phenomenon may be the fact that the more knots the aggregation classes contain the more "similar" to $\mathbf{A}_1$ the matrix $\mathbf{D}$ is, and thus the more effective the iterative method with $\mathbf{I} - \omega\mathbf{D}^{-1}\mathbf{A}$ is.

References

[1] *Blaheta, R.*: Iteration methods for numerical solution of boundary elasticity problems, VÚB, Ostrava, 1987, Dissertation. (In Czech.)

[2] *Blaheta, R.*: A multi-level method with correction by aggregation for solving discrete elliptic problems, Aplikace matematiky 5 no. 31 (1986), 365–378.

[3] *Brandt, A.*: Algebraic Multigrid Theory: The Symmetric Case, Preliminary Proceedings of the International Multigrid Conference, Copper Mountain, Colorado, April 6–8 1983.

[4] *Ruge, J. W., Stüben, K.*: Algebraic Multigrid, in [5].

[5] Multigrid Methods. Frontiers in Applied Mathematics (Mc Cormick, S. F., ed.), Society for industrial and applied mathematics, Philadelphia, Pennsylvania, 1987.

[6] *Míka, S., Vaněk, P.*: On the convergence of a two-level algebraic algorithm, Sborník referátů VIII. letní školy Software a algoritmy numerické matematiky (Sušice 1989), JČMF, 1990.

Authors' address: Doc. RNDr. *Stanislav Míka*, CSc., ing. *Petr Vaněk*, Západočeská universita, Americká 42, 306 14 Plzeň.