Applications of Mathematics

Ivo Marek; Petr Mayer

Aggregation/disaggregation iterative methods applied to Leontev systems and Markov chains

Applications of Mathematics, Vol. 47 (2002), No. 2, 139–156

Persistent URL: http://dml.cz/dmlcz/134491

47 (2002) APPLICATIONS OF MATHEMATICS No. 2, 139–156

AGGREGATION/DISAGGREGATION ITERATIVE METHODS APPLIED TO LEONTEV SYSTEMS AND MARKOV CHAINS*

Ivo Marek, Petr Mayer, Praha

Abstract. The paper surveys some recent results on iterative aggregation/disaggregation methods (IAD) for computing stationary probability vectors of stochastic matrices and solutions of Leontev linear systems. A particular attention is paid to fast IAD methods.

Keywords: Leontev model, Markov chain, stochastic matrix, aggregation, stationary probability vector

MSC 2000: 15A06, 15A51, 65F10

1. Introductory remarks

Our aim is to report some results of computing some quantities arising in mathematical models of real life situations based on classical principles such as mass preservation etc. expressed by Leontev systems [18] and Markov chains [6], [7], [10], [11]. By doing that, special linear systems are to be investigated whose solutions are structured in a particular way. The systems considered are as a rule very large in the sense of the number of degrees of freedom and mostly ill-conditioned. We want to show that in order to overcome the difficulties connected with these circumstances it is necessary to propose adequate and mostly new computational means. Among them the aggregation/disaggregation iterative methods (IAD) are highly efficient. In comparison with other methods the superiority of IAD methods grows with the scale of complexity. We are going to document this statement by a series of experiments coming from our investigations in modeling some structural problems in biology and in particular in physiology on the one hand and in problems of reliable railway safety systems on the other.

* This work was partly supported by the Grant No. 201/02/0595 of the Grant Agency of the Czech Republic, the Grant CEZ J04/98:210000010 and the Grant MSM 113200007.

The paper is organized as follows. Section 1 Introductory Remarks is followed by Section 2, in which some definitions and notation are given. Main objects of our study are described in Section 3 Leontev systems and Section 4 Stationary Probability Vectors of Stochastic Matrices. Section 5 contains the algorithm studied, while convergence results are presented in Section 6 and Section 7. Section 8 is devoted to presenting some numerical experiments and conluding remarks. The paper contains a rich though surely incomplete list of references, especially of the recent ones.

2. Definitions and notation

Our analyses are carried out in finite dimensional Banach spaces. Because of the equivalence of all norms on such spaces we can in principle use any norm. However, in the context of stochastic matrices the l_1-norm is the most adequate. Throughout the whole paper the symbol $\|\cdot\|$ denotes the l_1-norm on the appropriate space without specifying the space explicitly.

Let N and n be positive integers. Objects of our investigation are matrices whose elements are real numbers. Let C denote an $N \times N$ matrix. An $N \times N$ matrix $C = (c_{jk})$ with $c_{jk} \in \mathbb{R}^1$ is called nonnegative if $c_{jk} \geqslant 0$, $j, k = 1, \ldots, N$. In particular, let I denote the $N \times N$ identity matrix. We denote by $\mathbb{R}^N$ the standard arithmetic space of N-tuples of real numbers. Let $[\cdot, \cdot]$ denote the standard inner product on $\mathbb{R}^N$:

$$[x, y] = \sum_{j=1}^{N} x_j y_j, \quad x = (x_1, \ldots, x_N)^T, \quad y = (y_1, \ldots, y_N)^T \in \mathbb{R}^N.$$

For example, we denote, on the one hand, $\|x\| = \sum\limits_{j=1}^{N} |x_j|$, and, on the other hand,

$$\|C\| = \max\left\{ \frac{\|Cx\|}{\|x\|} \colon\ x \in \mathbb{R}^N,\ \ x \neq 0 \right\}.$$

2.1. Definition. Let A be an $N \times N$ matrix. A pair of matrices $\{M, W\}$ is called a splitting of A if $A = M - W$ and the inverse M^{-1} exists. A splitting of the matrix A is said to be of nonnegative type [13], or equivalently, *weak*, if the iteration matrix $T = M^{-1}W$ is nonnegative. If, in particular, the matrices M^{-1} and W are nonnegative, the splitting is called *regular* [24, p. 88]. If M^{-1} and $T = M^{-1}W$ are nonnegative, the splitting is called *weak regular* [19, p. 56]. A splitting $\{M, W\}$ is called *convergent* if $\lim\limits_{k\to\infty} T^k$ exists, T is *zero-convergent* if $T^k \to 0$ as $k \to \infty$.

Let Y denote an $N \times N$ matrix. We call a splitting $\{M, W\}$ *Y-convergent* if

$$\lim_{k \to 0} YT^k = 0.$$

A collection of all distinct eigenvalues of a square matrix A is called the *spectrum* of A and is denoted by $\sigma(A)$. We denote

$$r(A) = \mathrm{Max}\{|\lambda| \colon \lambda \in \sigma(A)\}$$

and call it the *spectral radius* of A.

3. Leontev systems

Let C be an $N \times N$ matrix with nonnegative real elements c_{jk}. A system

$$x - Cx = b, \quad b \in \mathbb{R}^N_+, \tag{3.1}$$

is called a *Leontev system* if C is zero-convergent and the matrix C is a *Leontev matrix* [18, pp. 88–95].

It is obvious that system (3.1) is uniquely solvable and the solution denoted by x^* is componentwise nonnegative.

3.1. Remark. Let a matrix C with nonnegative real elements be irreducible and satisfy

$$\sum_{j=1}^{N} c_{jk} \leqslant 1$$

and let there be an index k_0, $1 \leqslant k_0 \leqslant N$, such that

$$\sum_{j=1}^{N} c_{jk_0} < 1.$$

Then C is a Leontev matrix.

Problem P

We are going to consider the class of eigenvalue problem characterized by

$$x = Bx, \quad [x, e(N)] = 1, \tag{4.1}$$

under the restriction

$$\sum_{j=1}^{N} b_{jk} = 1, \quad \text{i.e. } B^T e(N) = e(N), \tag{4.2}$$

where $e(N) = (1, \dots, 1)^T \in \mathbb{R}^N$. It follows from (4.1) that

$$r(B) = 1, \quad \operatorname{ind}(I - B) = 1,$$

where $\operatorname{ind}(C)$ denotes the maximal size of the Jordan blocks corresponding to the value 0. We set $\operatorname{ind}(C) = 0$ if 0 is not an eigenvalue of C, i.e. the inverse C^{-1} exists.

Note that any solution to (4.1) is called a *stationary probability vector of* B.

A matrix B satisfying (4.2) is called *(column) stochastic*. There is a permutation matrix H such that [5, p. 341]

$$HBH^T = \begin{pmatrix} G_0 & 0 & \dots & 0 \\ G_1 & F_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ G_p & 0 & \dots & F_p \end{pmatrix},$$

where

$$\lim_{k \to \infty} G_0^k = 0$$

and F_j is an irreducible and stochastic matrix.

It is known [4] that the matrix H can be obtained by the so called *Tarjan algorithm* and that its complexity is almost linear.

Let B be a stochastic matrix. We may always assume that B is convergent. Otherwise we may consider

$$\hat{B} = \frac{1}{2}(I + B).$$

Let $\{M, W\}$ be a splitting of $A = I - B$. Set

$$T = P + Z, \tag{4.3}$$

where

$$P^2 = P, \quad PZ = ZP = 0, \quad r(Z) \leqslant 1, \quad 1 \notin \sigma(Z). \tag{4.4}$$

5. Aggregation/disaggregation algorithms

Let $\mathcal{G}$ be a map of $\{1,\dots,N\}$ onto $\{1,\dots,n\}$. Whenever it is necessary we distinguish the indices from the set $\{1,\dots,n\}$ by bars from those belonging to the set $\{1,\dots,N\}$; otherwise we use a simplified notation.

Let $\mathcal{K}$ denote the permutation of the set $\{1,\dots,N\}$ given by the relations

$$\mathcal{K}(1)=j_1,\dots,\mathcal{K}(N)=j_N$$

and denote the associated permutation matrix by K, and let $\mathcal{G}_{\mathcal{K}}$ be the apropriate map $1,\dots,N\to 1,\dots,n$. We then have

$$\begin{aligned}(Kx)^T &= (x_{j_1},\dots,x_{j_{n_1}},x_{j_{n_1+1}},\dots,x_{j_{n_1+n_2}},\dots,x_{j_N})\\ &= (u^T_{\mathrm{sub}(\overline{1})},\dots,u^T_{\mathrm{sub}(\overline{n})}).\end{aligned}$$

It is easy to see that

$$\begin{aligned}\{j\in\{1,\dots,N\}\colon\ \mathcal{G}(j)=\overline{1}\} &= \{j_1,\dots,j_{n_1}\}\\ \{j\in\{1,\dots,N\}\colon\ \mathcal{G}(j)=\overline{2}\} &= \{j_{n_1+1},\dots,J_{n_1+n_2}\}\\ &\vdots\\ \{j\in\{1,\dots,N\}\colon\ \mathcal{G}(j)=\overline{n}\} &= \{j_{n_1+\dots+n_{p-1}+1},\dots,j_N\}\end{aligned}$$

and

$$\begin{aligned}\{j\in\{1,\dots,N\}\colon\ \mathcal{G}_{\mathcal{K}}(j)=\overline{1}\} &= \{1,\dots,n_1\}\\ \{j\in\{1,\dots,N\}\colon\ \mathcal{G}_{\mathcal{K}}(j)=\overline{2}\} &= \{n_1+1,\dots,n_1+n_2\}\\ &\vdots\\ \{j\in\{1,\dots,N\}\colon\ \mathcal{G}_{\mathcal{K}}(j)=\overline{n}\} &= \{n_1+\dots+n_{n-1}+1,\dots,N\}\end{aligned}$$

where $B=K\tilde{B}K^{-1}$. We see that $\mathcal{G}$ yields an equivalent aggregation scheme for B as does $\mathcal{G}_{\mathcal{K}}$ for the permutation-similar matrix $\tilde{B}$.

We define communication operators R mapping $\mathbb{R}^N$ into $\mathbb{R}^n$ and $S(x)$ mapping $\mathbb{R}^n$ into $\mathbb{R}^N$ by setting

$$(5.1)\qquad (Ru)_{\overline{j}}=\sum_{\mathcal{G}(j)=\overline{j}}u_j,\quad u\in\mathbb{R}^N,\ \ u^T=(u^T_{\mathrm{sub}(\overline{1})},\dots,u^T_{\mathrm{sub}(\overline{n})}),\ \ u_{\mathrm{sub}(\overline{j})}\in\mathbb{R}^{n_j}$$

and

$$(5.2)\qquad (S(x)z)_j=\frac{x_j}{(Rx)_{\overline{j}}}z_{\overline{j}},\quad z\in\mathbb{R}^n,\ \ z_{\overline{j}}\in\mathbb{R}^1,\ \ \mathcal{G}(j)=\overline{j},\ \ j=1,\dots,N$$

for $x \in \mathcal{D}$, where

$$\mathcal{D} = \{x \in \mathbb{R}^N : \ x^T = (x_1, \ldots, x_N), \ \ x_j > 0, \ \ j = 1, \ldots, N\}.$$

We check immediately that

$$RS(x)z = z, \ \ \forall x \in \mathcal{D}, \ \ z \in \mathbb{R}^p.$$

Therefore,

$$P(x) = S(x)R$$

is a projection,

$$[P(x)]^2 = P(x).$$

Moreover,

$$P(x)x = x, \ \ \forall x \in \mathcal{D}$$

and

$$P(x)^T e = e, \ \ \forall x \in \mathcal{D}.$$

We define a matrix $\mathcal{B}(x) = RBS(x)$, $x \in \mathcal{D}$, and call it the *aggregated matrix* (with respect to B).

To guarantee that the proposed two-level algorithms can be unlimitedly realized we need the following two statements proved in [14].

5.1. Proposition [14]**.** *Let a matrix B be stochastic. Then its aggregated matrix $\mathcal{B}(x)$, $x \in \mathcal{D}$, is stochastic, too.*

5.2. Proposition [14]**.** *Let a stochastic matrix B be irreducible. Then its aggregated matrix $\mathcal{B}(x)$, $x \in \mathcal{D}$, is irreducible, too.*

5.3. Algorithm SPV $(B; M, W; t, s; x^{(0)})$.

Let B be an $N \times N$ irreducible stochastic matrix, let $\{M, W\}$ be a splitting of nonnegative type of a matrix A and let $T = M^{-1}W$ with $r(T) = 1$, $s, t \geqslant 1$ positive integers.

Let $\varepsilon > 0$ be a given tolerance and let $x^{(0)}$ with $(x^{(0)})_j > 0$, $[x^{(0)}, e(N)] = 1$, $j = 1, \ldots, N$, be an arbitrary vector.

Step 1. Set $0 \to k$.

Step 2. Construct the matrix

$$\mathcal{B}(x^{(k)}) = RB^s S(x^{(k)}).$$

Step 3. Find the unique solution vector $\tilde{z}^{(k)}$ to Problem (P) with $\mathcal{B}(y^{(k)})$, i.e. the unique solution to the problem

$$\mathcal{B}(x^{(k)})\tilde{z}^{(k)} = \tilde{z}^{(k)}, \tag{5.3}$$

$$[\tilde{z}^{(k)}, e(n)]_n = 1. \tag{5.4}$$

Step 4. Disaggregate by setting

$$v^{(k+1)} = S(x^{(k)})\tilde{z}^{(k)}.$$

Step 5. Let

$$x^{(k+1,m)} = Tx^{(k+1,m-1)}, \quad x^{(k+1,0)} = v^{(k+1)}, \quad m = 1, \ldots, t,$$
$$x^{(k+1)} = x^{(k+1,t)}, \quad [x^{(k+1)}, e(N)]_N = 1.$$

Step 6. Test whether

$$\|x^{(k+1)} - x^{(k)}\| < \varepsilon.$$

Step 7. If NO in Step 6, then let

$$k + 1 \to k$$

and GO TO Step 2.

Step 8. If YES in Step 6, then set

$$\hat{x} := x^{(k+1)}$$

and STOP.

5.4. Algorithm LM $(C; M, W; t; y^{(0)})$.

Let C be an $N \times N$ aggregation convergent matrix with nonnegative real elements, and let $\{M, W\}$ be a splitting of $A = I - C$ of nonnegative type.

Step 1. Set $0 \to k$.

Step 2. Construct the matrix

$$\mathcal{C}(y^{(k)}) = RCS(y^{(k)}).$$

Step 3. Find the unique solution $\tilde{z}^{(k)}$ to the problem

$$\tilde{z}^{(k)} - \mathcal{C}(y^{(k)})\tilde{z}^{(k)} = Rb. \tag{5.5}$$

Step 4. Disaggregate by setting

$$v^{(k+1)} = S(y^{(k)})\tilde{z}^{(k)}.$$

Step 5. Let

$$My^{(k+1,m)} = Ny^{(k+1,m-1)} + b, \quad y^{(k+1,0)} = v^{(k+1)}, \quad m = 1, \ldots, t,$$

$$y^{(k+1)} = y^{(k+1,t)}.$$

Step 6. Test whether

$$\left\| y^{(k+1)} - y^{(k)} \right\| < \varepsilon.^{1}$$

Step 7. If NO in Step 6, then let

$$k + 1 \to k$$

and GO TO Step 2.

Step 8. If YES in Step 6, then set

$$x^* := y^{(k+1)}$$

and STOP.

[1] Here the symbol $\|\cdot\|$ denotes any norm on $\mathbb{R}^N$. We recommend the l_1-norm.

6. Convergence results

6.1. Error-vector formulas

6.1. Proposition. *The error-vector formulas for the sequence of approximants* $\{x^{(k)}\}$ *returned either by the* SPV$(B; M, W; t, s; y^{(0)})$*—algorithm* 5.3 *or* $\{y^{(k)}\}$ *returned by the LM*$(C; M, W; t; y^{(0)})$*—algoritm 5.4 read*

$$\begin{cases} x^{(k+1)} - \hat{x} \ = J_t(x^{(k)})(x^{(k)} - \hat{x}), \\ \\ y^{(k+1)} - x^* = J_t(y^{(k)})(y^{(k)} - x^*) \end{cases} \tag{6.1}$$

where

$$J_t(x) = T^t[I - P(x)V]^{-1}(I - P(x)) \tag{6.2}$$

with $V = Z$ *coming from the spectral decomposition of* B (4.3) *for the case of Algorithm* 5.3 *and* $V = C$ *for the case of Algorithm* 5.4. *Furthermore,* $J_t(x) = T^{t-1}J_1(x)$, $t \geqslant 1$, *holds for any* x *with all components positive.*

P r o o f. By the definition of the SPV$(B; M, W; t)$-algorithm,

$$My^{(k+1,1)} = WS(x^{(k)})z^{(k)}, \tag{6.3}$$

where

$$z^{(k)} = \big(I_{\mathcal{F}} - \mathrm{RVS}(x^{(k)})\big)^{-1} R\hat{x}, \quad \hat{x} = (I - V)\hat{x}.$$

It follows that

$$\begin{aligned} My^{(k+1,1)} &= W\{S(x^{(k)})(I_{\mathcal{F}} - \mathrm{RVS}(x^{(k)})^{-1}R\hat{x}\} \\ &= W\big(I - P(x^{(k)})V\big)^{-1}P(x^{(k)})(I - V)\hat{x} \\ &= W\big(I - P(x^{(k)})V\big)^{-1}\big(P(x^{(k)}) - I + I - P(x^{(k)})V\big)\hat{x} \\ &= W\hat{x} - W\big(I - P(x^{(k)})V\big)^{-1}\big(I - P(x^{(k)}\big)\hat{x} \end{aligned}$$

and, since $M\hat{x} = W\hat{x}$,

$$\begin{aligned} My^{(k+1,1)} - W\hat{x} &= M(y^{(k+1,1)} - \hat{x}) \\ &= W\big(I - P(x^{(k)})V\big)^{-1}\big(I - P(x^{(k)})\big)(x^{(k)} - \hat{x}). \end{aligned}$$

Finally,

$$y^{(k+1,1)} - \hat{x} = T\big(I - P(x^{(k)})V\big)^{-1}\big(I - P(x^{(k)})\big)(x^{(k)} - \hat{x}).$$

This is just formula (6.2) for $t = 1$. To obtain (6.2) for arbitrary $t \geqslant 1$ one needs to apply T^{t-1} to $J_1(x^{(k)})$. It is obvious that the algorithm SPV$(B; M, W; t, s; x^{(0)})$ achieves this purpose by applying the iteration procedure determined by the splitting $\{M, W\}$.

It is easy to see that the error-vector formula (6.1) for the sequence returned by Algorithm 5.4 is obtained in the same manner as that of 5.3 because the role of the matrix $V = Z$ in Algorithm 5.3 is played in Algorithm 5.4 by the matrix $V = C$ itself. □

6.2. Proposition. *The spectra of $J_t(\hat{x})$ and $(I-P(\hat{x}))J_t(\hat{x})$ are related as follows:*

$$\sigma(J_t(\hat{x})) \subset \sigma\big((I - P(\hat{x}))J_t(\hat{x})\big) \cup \{0\}.$$

Consequently, $r(J_t(\hat{x})) = r\big((I - P(\hat{x}))J_t(\hat{x})\big)$.

P r o o f. Let $0 \neq \lambda \in \sigma(J_t(\hat{x}))$ and let w be a corresponding eigenvector $J_t(\hat{x})w = \lambda w$, $w \neq 0$. According to the definition of $J_t(x)$ we see that

$$(I - P(\hat{x}))J_t(\hat{x})(I - P(\hat{x}))w = \lambda(I - P(\hat{x}))w.$$

□

6.2. Convergence of the SPV algorithms

Some propositions, lemmas and theorems of this and the next subsection are stated without proof; the corresponding proofs can be found in [15], [1].

In Step 3 of Algorithm 3.5 the existence of a stationary probability vector on the aggregated level is required. To this purpose we have

6.3. Proposition [14]**.** *The aggregated matrix $RBS(x)$, $x \in \mathbb{R}^N$, $x_j > 0$, $j = 1, \ldots, N$, is column stochastic.*

A uniqueness result concerning the stationary probability vector in the above proposition is guaranteed by the following

6.4. Proposition [14]**.** *Let $x \in \operatorname{Int} \mathbb{R}^N$ and let B be an irreducible column stochastic matrix. Then the aggregated matrix $RBS(x)$ is irreducible.*

6.5. Theorem. *Let B be an $N \times N$ irreducible column stochastic matrix. Let $\{M, W\}$ be an aggregation-convergent splitting of nonnegative type of the matrix $I - B$. Denote the iteration matrix corresponding to this splitting by T, i.e. $T = M^{-1}W$.*

Then there is a neighbourhood $\Omega(\hat{x})$ and a positive integer $t \geqslant 1$ such that $\mathrm{SPV}(B; M, W; t, s; x^{(0)})$*—Algoritm* 5.3 *is convergent whenever $x^{(0)} \in \Omega(x^*)$. Moreover, the error estimate*

$$\|x^{(k)} - \hat{x}\| \leqslant \kappa \varrho^k \|x^{(0)} - \hat{x}\|$$

holds where $\|\cdot\|$ denotes any norm on $\mathbb{R}^N$ and κ and $\varrho < 1$ are positive real numbers independent of $k = 0, 1, \ldots$

6.3. Convergence of the LM algorithms

6.6. Theorem. *Let C be an $N \times N$ matrix with nonnegative real elements c_{jk}, let C be zero-convergent and $b \in \mathbb{R}^N_+$. Let $\{M, W\}$ be an aggregation-convergent splitting of nonnegative type of the matrix $I - C$. Denote the iteration matrix corresponding to this splitting by T, i.e. $T = M^{-1}W$.*

Then there is a neighbourhood $\Omega(x^*)$ and a positive integer $t \geqslant 1$ such that $\mathrm{LM}(C; M, W; t; y^{(0)})$ *Algorithm* 5.4 *is convergent whenever $y^{(0)} \in \Omega(x^*)$. Moreover, the error estimate*

$$\|y^{(k)} - x^*\| \leqslant \kappa \varrho^k \|y^{(0)} - x^*\|$$

holds where $\|\cdot\|$ denotes any norm on $\mathbb{R}^N$ and κ and $\varrho < 1$ are positive real numbers independent of $k = 0, 1, \ldots$

7. Fast IAD methods

7.1. Proposition [16]**.** *Let $\alpha_{\overline{1}}, \ldots, \alpha_{\overline{n}}$ be positive real numbers. Let*

$$x_{\mathrm{sub}(\overline{j})} = \alpha_{\overline{j}} y_{\mathrm{sub}(\overline{j})}, \quad \overline{j} = 1, \ldots, n, \tag{7.1}$$

where $x^T = (x^T_{\mathrm{sub}(\overline{1})}, \ldots, x_{\mathrm{sub}(\overline{n})})$, $y = (y_{\mathrm{sub}(\overline{1})}, \ldots, y_{\mathrm{sub}(\overline{n})})$ are vectors x, y blockwise written in accordance with the block form of B given by the map $\mathcal{G}$. Then

$$P(x)y = y. \tag{7.2}$$

P r o o f. We see that

$$(P(x)y)_j = (S(x)Ry)_j = \frac{x_j}{(Rx)_{\overline{j}}}(Ry)_{\overline{j}}$$

and further, because of (7.1),

$$(P(x)y)_j = \frac{\alpha_{\overline{j}} y_j}{\alpha_{\overline{j}}(Ry)_{\overline{j}}}(Ry)_{\overline{j}} = y_j, \quad j \in \{j\colon\ \mathcal{G}(j) = \overline{j}\},\ \ \overline{j} = 1, \ldots, n.$$

Consequently, (7.2) holds. □

7.2. Corollary. *If* (7.1) *holds for some* $x^{(k)}$*—an element returned by* SPV *Algorithm—instead of* x *and* $\hat{x}$*—the unique stationary probability vector of* B*—instead of* y*, then*

$$\big(I - P(x^{(k)})\big)(x^{(k)} - \hat{x}) = 0$$

and, by (6.1),

$$x^{(k+1)} = \hat{x}.$$

P r o o f. According to (6.1) and Proposition 7.1, we have

$$\big(I - P(x^{(k)})\big)(x^{(k)} - \hat{x}) = (x^{(k)} - \hat{x}) - P(x^{(k)})(x^{(k)} - \hat{x}) = -\hat{x} + P(x^{(k)})\hat{x} = 0.$$

□

7.3. Theorem [16]**.** *Let* B *be an* $N \times N$ *irreducible stochastic matrix such that in the block form*

$$B = \begin{pmatrix} B_{\overline{11}} & B_{\overline{12}} & \dots & B_{\overline{1n}} \\ B_{\overline{21}} & B_{\overline{22}} & \dots & B_{\overline{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ B_{\overline{n1}} & B_{\overline{n2}} & \dots & B_{\overline{nn}} \end{pmatrix}, \tag{7.3}$$

the off-diagonal blocks satisfy the relations $R_{\overline{j}} = \operatorname{range}(B_{\overline{jk}}) = \{\alpha f_{\operatorname{sub}(\overline{j})} \colon \alpha \in \mathbb{R}^1\}$, $\mathcal{G}(j) = \overline{j}$*, where* $f_{\operatorname{sub}(\overline{j})}$ *is a vector in* $\mathbb{R}^{n_j}$*,* n_j *being the size of the block* $B_{\overline{jj}}$*. Let* $\{M, W\}$ *be a splitting of* $A = I - B$ *of positive type such that* $W = M - A = (W_{\overline{jk}})$*,* $W_{\overline{jj}} = 0$*. Let*

$$(Rx)_{\overline{j}} = \sum_{\mathcal{G}(j)=\overline{j}} x_j, \quad \overline{j} = 1, \dots, n,$$

where the index $\overline{j}$ *corresponds to the block* $B_{\overline{jj}}$ *in* (7.3)*.*

Then $x^{(2)} = \hat{x}$*,* $B\hat{x} = \hat{x}$*,* $[\hat{x}, e(N)] = 1$*, where* $x^{(k)}$ *is a return of the* SPV$(B; M, W; t, s; x^{(0)})$ *algorithm in the* k*th step.*

P r o o f. Let $x^{(0)}$ be an initial approximation of $\hat{x}$. By Proposition 7.1 we have

$$(x^{(1)})_{\operatorname{sub}(\overline{j})} \in R_{\overline{j}}, \quad \overline{j} = 1, \dots, n,$$

and consequently,

$$(x^{(2)})_{\operatorname{sub}(\overline{j})} \in R_{\overline{j}}, \quad \overline{j} = 1, \dots, n.$$

Then, by hypothesis,

$$(\hat{x})_{\operatorname{sub}(\overline{j})} = \alpha_{\overline{j}}(x^{(1)})_{\operatorname{sub}(\overline{j})}$$

with some positive reals $\alpha_{\overline{1}}, \dots, \alpha_{\overline{n}}$. Utilizing Corollary 7.2 we deduce $x^{(2)} - \hat{x} = 0$ and this completes the proof. □

Continuing in the spirit of Theorem 6.6 one easily arrives at

7.4. Theorem [16]. *Let $b \in \mathbb{R}^N_+$ be in the range of C_N, where C_N denotes the block matrix C with the diagonal blocks in the expression* (7.4) *replaced by zero-blocks and C is a nonnegative matrix such that in the block form*

$$C = \begin{pmatrix} C_{\overline{11}} & C_{\overline{12}} & \dots & C_{\overline{1n}} \\ C_{\overline{21}} & C_{\overline{22}} & \dots & C_{\overline{2n}} \\ \vdots & \vdots & \ddots & \vdots \\ C_{\overline{n1}} & C_{\overline{n2}} & \dots & C_{\overline{nn}} \end{pmatrix}, \tag{7.4}$$

the off-diagonal blocks satisfy the relations $R_{\overline{j}} = \operatorname{range}(C_{\overline{jk}}) = \{\alpha f_{\operatorname{sub}(\overline{j})} \colon \alpha \in \mathbb{R}^1\}$, $\mathcal{G}(j) = \overline{j}$, where $f_{\operatorname{sub}(\overline{j})}$ is a vector in $\mathbb{R}^{n_j}$, n_j being the size of the block $C_{\overline{jj}}$. Let $\{M, W\}$ be a splitting of $A = I - C$ of nonnegative type such that $W = M - A = (W_{\overline{jk}})$, $W_{\overline{jj}} = 0$.

Then $y^{(2)} = x^*$, $x^* = Cx^* + b$, where $y^{(k)}$ is the return of the LM *algorithm in the kth step.*

7.5. Remark. A particular case, when the ranges of $B_{\overline{jk}}$ consist of a linear hull of a single vector of the standard basis in the appropriate space $\mathbb{R}^{n_j}$, is presented in [7].

7.6. Remark. Note that the iteration processes defined by the splittings described in Theorem 7.3 are in general divergent. As an example of a divergent process can be taken the Jacobi method in the case of a p-cyclic matrix. Nevertheless, the convergence of the SPV algorithm based on such splittings may be fast and terminates after at most two IAD sweeps.

7.7. Remark. To check that a given block matrix satisfies the conditions of Theorem 7.3 might be difficult. This is because the problem to determine the rank of a matrix is non well posed. Fortunately, we have the following modification of Theorem 7.3:

7.8. Theorem. *Let B be an irreducible stochastic matrix such that every column of each of the blocks $B_{\overline{jk}}$, $\overline{j} \neq \overline{k}$, have the form $(1-\tau)\alpha_{\overline{j}} f_{\operatorname{sub}(\overline{j})} + \tau g_{\operatorname{sub}(\overline{j})}$, where $\|g_{\operatorname{sub}(\overline{j})}\| < \alpha$ and α_j, $j = 1, \dots, p$, and α are positive real numbers. Furthermore, let $\{M, W\}$ be a splitting of $A = I - B$ of positive type such that $W = M - A = (W_{\overline{j},\overline{k}})$, and $W_{\overline{j},\overline{j}} = 0$.*

Then there is a $\tau_0 > 0$ such that the rate of convergence of the IAD *algorithm* $\operatorname{SPV}(M; M, W; t, s; x^{(0)})$ *is bounded above by the product $\tau \operatorname{Max}\{\kappa_{\overline{j}} \colon \overline{j} = 1, \dots, p\}$ for all $\tau \leqslant \tau_0$, where $\kappa_{\overline{j}} = \operatorname{cond}(I_{n_j} - B_{\overline{j},\overline{j}})^{-1}$.*

7.9. Remark. It is obvious that a statement analogous to Theorem 7.8 is valid for the case of Algorithm 5.4.

In literature some algorithms are known such as PABLO and TPABLO [2] that transform the original matrix into its block-equivalent permutation-similar form with the goal of conditioning the diagonal blocks optimally. Our analysis shows that it is desirable simultaneously to force the off-diagonal blocks to be as close as possible to rank-one blocks having the same ranges of the row blocks. Obviously, these two requirements are in conflict with each other. Hence, a reasonable criterion needs to be found.

8. Numerical experiments and concluding remarks

We are going to consider the class of $N \times N$ stochastic matrices that are expressed as full $n_j \times n_k$ block-matrices B_{jk}, $1 \leqslant j, k \leqslant n$, $n \geqslant 2$ constructed according to the following rules.

1^0 Each block B_{jk}, $j \neq k$, is given as a combination of the tensor product of vectors $f_{\mathrm{sub}(j)} \otimes g_{\mathrm{sub}(k)}$ and a matrix C_{jk}:

$$B_{jk} = \tau C_{jk} + f_{\mathrm{sub}(j)} \otimes g_{\mathrm{sub}(k)}, \quad 1 \leqslant \tau, \quad j, k = 1, \ldots, n, \quad k \neq j,$$

where the components of the vectors $f_{\mathrm{sub}(j)}$, $g_{\mathrm{sub}(k)}$ and matrices C_{jk} are random numbers from the interval $(0, 1)$.

2^0 The diagonal blocks of orders $n_j \times n_j$ are formed by n_j^2 random numbers from the interval $(0, 1)$, where $j = 1, \ldots, n$, and then normalized in order to make the resulting matrix B column stochastic.

With matrices from the above class we have tested the behaviour of the SPV algorithm 5.3 by varying the following parameters: τ (the departure of a block from beeing rank one) and ε (the measure for matrix B to be nearly completely decomposable—NCD). We keep $n = 4$ and $n_j = n(\mathrm{block}) = 100$, thus $N = 400$.

The following methods are compared:

the standard power method,
the standard block Jacobi method,
the standard block Gauss-Seidel method (G.-S.),
Marek-Mayer method (M+M) [14],
Vantilborgh's method (Vant.) [23],
Koury-McAllister-Stewart method (KMS) [9], [19], [20].

Positive integers shown in Tables 1–6 denote the amount of iteration sweeps needed to achieve the accuracy 1×10^{-14}.

method	$\tau = 0$	$\tau = 0.001$	$\tau = 0.01$	$\tau = 0.1$	$\tau = 1$
power	1004000	1001000	982000	821000	312000
Jacobi	20	20	20	20	18
G.-S.	9	9	9	9	9
M+M	14	14	14	14	14
Vant.	1	2	2	2	3
KMS	1	2	2	2	3

Table 1. $n = 100$, $n(\text{block}) = 4$, $\varepsilon = 0.00001$.

method	$\tau = 0$	$\tau = 0.001$	$\tau = 0.01$	$\tau = 0.1$	$\tau = 1$
power	113300	113000	110800	92700	35100
Jacobi	23	23	23	22	20
G.-S.	9	9	9	9	10
M+M	14	14	14	14	14
Vant.	1	2	2	3	3
KMS	1	2	2	3	3

Table 2. $n = 100$, $n(\text{block}) = 4$, $\varepsilon = 0.0001$.

method	$\tau = 0$	$\tau = 0.001$	$\tau = 0.01$	$\tau = 0.1$	$\tau = 1$
power	12630	12600	12350	10330	3912
Jacobi	25	25	25	25	23
G.-S.	10	10	10	10	11
M+M	14	14	14	14	14
Vant.	1	2	3	3	4
KMS	1	2	3	3	4

Table 3. $n = 100$, $n(\text{block}) = 4$, $\varepsilon = 0.001$.

method	$\tau = 0$	$\tau = 0.001$	$\tau = 0.01$	$\tau = 0.1$	$\tau = 1$
power	1401	1398	1371	1147	438
Jacobi	28	28	28	27	25
G.-S.	11	11	11	12	12
M+M	14	14	14	14	14
Vant.	1	3	3	4	5
KMS	1	3	3	4	4

Table 4. $n = 100$, $n(\text{block}) = 4$, $\varepsilon = 0.01$.

method	$\tau = 0$	$\tau = 0.001$	$\tau = 0.01$	$\tau = 0.1$	$\tau = 1$
power	164	163	160	135	54
Jacobi	30	30	30	29	26
G.-S.	12	12	12	12	13
M+M	14	14	14	13	12
Vant.	1	3	4	5	7
KMS	1	3	4	4	6

Table 5. $n = 100$, $n(\text{block}) = 4$, $\varepsilon = 0.1$.

method	$\tau = 0$	$\tau = 0.001$	$\tau = 0.01$	$\tau = 0.1$	$\tau = 1$
power	26	26	25	21	13
Jacobi	32	32	32	30	26
G.-S.	12	12	12	13	13
M+M	10	10	10	10	10
Vant.	1	3	5	6	9
KMS	1	3	4	5	7

Table 6. $n = 100$, $n(\text{block}) = 4$, $\varepsilon = 1$.

The results obtained demonstrate very nicely some facts implied by the theory.

First, since all elements of the blocks forming the matrices considered are positive, each of the methods from our list of tested methods is convergent.

Second, the influence of parameters τ and ε can be traced from Tabs. 1–6. While the successive diminishing of the measure of NCD implies a deterioration of convergence of the power method, the IAD methods KMS and Vant. clearly show just an opposite tendency; the M+M method is very moderately dependent on the measure of NCD.

Third, dependence on the parameter τ is absent for the power method, block Jacobi, block Gauss-Seidel and M+M methods. Small changes in speed of these methods are caused by the fact that a decrease of τ implies an increase of the volume of the diagonal blocks and subsequently a possible enlarging of their conditioning.

In general, the speed of convergence of KMS and Vant. is influenced positively by increasing the NCD property and negatively by the condition of the diagonal blocks. Moreover, these methods return an exact solution if the off-diagonal blocks of B are rank one matrices possesssing the same range for all matrices of a fixed row. These facts suggest to build up procedures that would transform a given stochastic block matrix into another suitable form that would equilibrate the condition numbers of the diagonal blocks and simultaneously rearrange the block by permutation similarity to obtain the matrix in the form of the sum of a block diagonal and a matrix whose all off-diagonal blocks are rank one matrices and in every block row possess the same

range. So far, there exist procedures transforming a given block matrix into another one with maximal masses concentrated in the diagonal blocks, e.g. PABLO and TPABLO [2]. The M+M method possesses some more particular properties. First, the speed of convergence is immune of any of the parameters τ and ε. The second is an extremely important property that is absent from all other methods considered in our survey: This method is still applicable even if the individual elements of the matrix B are not available for computations and only the actions of B for any vector are. This property is very useful in some applications such as the modeling of reliable railway safety algorithms [6], [7], [8].

Acknowledgement. Several comments and alterations have been suggested by P. Přikryl, the authors express their sincere thanks to him.

References

[1] *W. L. Cao, W. J. Stewart*: Iterative aggregation/disaggregation techniques for nearly uncoupled Markov chains. J. Assoc. Comput. Mach. *32* (1985), 702–719.

[2] *H. Choi, D. Szyld*: Application of threshold partitioning of sparse matrices to Markov chains. In: Proceedings of the IEEE International Computer Performance and Dependability Symposium IPDS'96. Urbana, 1996, pp. 158–165.

[3] *T. Dayar, W. J. Stewart*: Comparison of partitioning techniques for two-level iterative solvers on large, sparse Markov chains. Tech. Rep. BU-CEIS-9805. Department of Computer Engineering and Information Science, Bilkent University, Ankara, 1998.

[4] *I. S. Duff, J. K. Reid*: An implementation of Tarjan's algorithm for the block triangularization of a matrix. ACM Trans. Math. Software *4* (1978), 337–147.

[5] *F. R. Gantmacher*: The Theory of Matrices. Gos. Izd. Lit., Moscow, 1954 (In Russian.); English translation: Applications of the Theory of Matrices. Interscience, New York, 1959.

[6] *Š. Klapka, P. Mayer*: Reliability modelling of safety equipments. In: Proceedings of Programs and Algorithms of Numerical Mathematics, Libverda, June 2000. Mathematical Institute of the Academy of Sciences of the Czech Republic, Prague, 2000, pp. 78–84. (In Czech.)

[7] *Š. Klapka, P. Mayer*: Some aspects of modelling railway safety. Proc. SANM'99, Nečtiny (I. Marek, ed.). University of West Bohemia, Plzeň, 1999, pp. 135–140.

[8] *Š. Klapka, P. Mayer*: Aggregation/disaggregation method for safety models. Appl. Math. *47* (2002).

[9] *R. Koury, D. F. McAllister and W. J. Stewart*: Methods for computing stationary distributions of nearly completely decomposable Markov chains. SIAM J. Alg. Disc. Meth. *5* (1984), 164–186.

[10] *U. Krieger*: On iterative aggregation/disaggregation methods for finite Markov chains. Preprint Deutsche Bundespost Telekom. Research and Technology Centre, 1990.

[11] *U. Krieger*: Numerical solution methods for large finite Markov chains. In: Performance and Reliability Evaluation (K. Hare et al., eds.). R. Oldenbourg, Wien, 1994, pp. 267–318.

[12] *S. T. Leutenegger, G. H. Horton*: On the utility of the multi-level algorithm for the solution of nearly completely decomposable Markov chains. In: Proceedings of the Second

Internationl Workshop on the Numerical Solution of Markov Chains (W. J. Stewart, ed.). Kluwer, Boston, 1995, pp. 425–442.
[13] *I. Marek*: Frobenius theory of positive operators. Comparison theorems and applications. SIAM J. Appl. Math. *19* (1970), 607–628.
[14] *I. Marek, P. Mayer*: Convergence analysis of an aggregation/disaggregation iterative method for computation stationary probability vectors of stochastic matrices. Numer. Linear Algebra Appl. *5* (1998), 253–274.
[15] *I. Marek, P. Mayer*: Convergence theory of a class of iterative aggregation/disaggregation methods for computing stationary probability vectors of stochastic matrices. Linear Algebra Appl. (2001). Submitted.
[16] *I. Marek, P. Mayer*: Iterative aggregation/disaggregation methods for computing stationary probability vectors of stochastic matrices can be finitely terminating. International Journal of Differential Equations *3* (2001), 301–313.
[17] Matrix Market. A repository of test matrices of the National Institute of Standards and Technology. http://www math.nist.gov/MatrixMarket.
[18] *H. Nikaido*: Convex Structures and Economic Theory. Academic Press, New York-London, 1968.
[19] *J. Ortega, W. Rheinboldt*: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York, 1970.
[20] *G. W. Stewart, W. J. Stewart and D. F. McAllister*: A two stage iteration for solving nearly uncoupled Markov chains. In: IMA Volumes in Mathematics and Applications 60: Recent Advances in Iterative Methods (G. H. Golub, A. Greenbaum and M. Luskin, eds.). Springer Verlag, New York, 1994, pp. 201–216.
[21] *W. J. Stewart*: Introduction to the Numerical Solution of Markov Chains. Princeton University Press, Princeton, 1994.
[22] *Y. Takahashi*: A lumping method for numerical calculations of stationary distributions of Markov chains. Res. Rep. B-18. Dept. of Inf. Sci. Tokyo Inst. of Tech., Tokyo, Japan, June 1975.
[23] *H. Vantilborgh*: The error aggregation. A contribution to the theory of decomposable systems and applications. PhD Thesis. Faculty of Applied Sciences, Louvain Catholic University, Louvain-la Neuve, Belgium, 1981.
[24] *R. S. Varga*: Matrix Iterative Analysis. Prentice-Hall, Englewood Cliffs, 1962.

Authors' addresses: *I. Marek*, Department of Mathematics, Faculty of Civil Engineering, Czech Technical University, 166 29 Praha 6, Czech Republic, e-mail: `marek@ms.mff.cuni.cz`; *P. Mayer*, Department of Computational Mathematics, Charles University, Sokolovská 83, 186 00 Praha 8, Czech Republic, e-mail: `pmayer@ms.mff.cuni.cz`.