MATHEMATICS OF COMPUTATION
Volume 88, Number 316, March 2019, Pages 783–800
http://dx.doi.org/10.1090/mcom/3347
Article electronically published on May 29, 2018

A FAST AND STABLE TEST TO CHECK IF A WEAKLY DIAGONALLY DOMINANT MATRIX IS A NONSINGULAR M-MATRIX

PARSIAD AZIMZADEH

Abstract. We present a test for determining if a substochastic matrix is convergent. By establishing a duality between weakly chained diagonally dominant (w.c.d.d.) L-matrices and convergent substochastic matrices, we show that this test can be trivially extended to determine whether a weakly diagonally dominant (w.d.d.) matrix is a nonsingular M-matrix. The test's runtime is linear in the order of the input matrix if it is sparse, and quadratic if it is dense. This is a partial strengthening of the cubic test in [J. M. Peña., *A stable test to check if a matrix is a nonsingular M-matrix*, Math. Comp., 247, 1385–1392, 2004]. As a by-product of our analysis, we prove that a nonsingular w.d.d. M-matrix is a w.c.d.d. L-matrix, a fact whose converse has been known since at least 1964.

1. Introduction

The *substochastic matrices* are real matrices with nonnegative entries and whose row-sums are at most one. We establish two results relating to this family:

(i) To each substochastic matrix B we associate a possibly infinite *index of contraction* $\widehat{\operatorname{con}}B$ and show that for each nonnegative integer k, B^k is a contraction in the infinity norm (i.e., $\|B^k\|_\infty < 1$) if and only if $k > \widehat{\operatorname{con}}B$.

(ii) We show that the index of contraction of a sparse (resp., dense) square substochastic matrix is computable in time linear (resp., quadratic) in the order of the input matrix.

It follows immediately from (i) that a square substochastic matrix is convergent if and only if its index of contraction is finite.

By establishing a duality between *weakly chained diagonally dominant* (w.c.d.d.) *L-matrices* and convergent substochastic matrices, we use point (ii) to obtain a test to determine whether a weakly diagonally dominant (w.d.d.) matrix is a nonsingular *M-matrix*. Previous work in this regard is the test in [15] to determine if an arbitrary matrix (not necessarily w.d.d.) is a nonsingular M-matrix, which has a cost asymptotically equivalent to Gaussian elimination (i.e., cubic in the order of the input matrix).

W.d.d. M-matrices arise naturally from discretizations of differential operators and appear in the Bellman equation for optimal decision making on a controlled

Received by the editor January 24, 2017, and, in revised form, June 11, 2017, November 14, 2017, and December 14, 2017.

2010 *Mathematics Subject Classification.* Primary 65F30, 15B48, 15B51; Secondary 65F50.

Key words and phrases. M-matrix, substochastic matrix, diagonally dominant, weakly chained diagonally dominant, sparse matrix.

Markov chain [3]. As such, these matrices have attracted a significant amount of attention from the scientific computing and numerical analysis communities.

W.c.d.d. matrices were first studied in a wonderful work by P. N. Shivakumar and K. H. Chew [18] in which they were proven to be nonsingular (see also [1] for a short proof). Various authors have recently studied the family of w.c.d.d. M-matrices, obtaining bounds on the infinity norm of their inverses (i.e., $\|A^{-1}\|_\infty$) [5, 10, 14, 19, 23]. While a w.c.d.d. matrix is w.d.d. by definition, the converse is not necessarily true in general (e.g., $\left(\begin{smallmatrix}+1 & -1\\ -1 & +1\end{smallmatrix}\right)$ is w.d.d. but not w.c.d.d.).

It has long been known (possibly as early as 1964; see the work of J. H. Bramble and B. E. Hubbard [4]) that a w.c.d.d. L-matrix[1] is a nonsingular w.d.d. M-matrix. We obtain a proof of the converse as a by-product of our analysis. In particular, we establish that[2]

$$\begin{aligned} A \text{ is a nonsingular w.d.d. M-matrix} &\iff A \text{ is a nonsingular w.d.d. L-matrix} \\ &\iff A \text{ is a w.c.d.d. L-matrix.} \end{aligned} \tag{1.1}$$

(1.1) is also useful in that it gives a graph-theoretic characterization of nonsingular w.d.d. M-matrices by means of w.c.d.d. matrices. This characterization is often easier to use than the usual characterizations involving, say, inverse-positivity or positive principal minors [16].

We list a few other interesting recent results concerning w.c.d.d. matrices and M-matrices here: [11–13, 20, 22, 24–26].

Section 2 introduces and establishes results on substochastic matrices, M-matrices, and w.c.d.d. matrices. Section 3 gives the procedure to compute the index of contraction. Section 4 presents numerical experiments testing the efficacy of the procedure on randomly sampled matrices.

2. Matrix families

2.1. **Substochastic matrices.**

Definition 2.1. A substochastic matrix is a real matrix $B \coloneqq (b_{ij})$ with nonnegative entries (i.e., $b_{ij} \geq 0$) and row-sums at most one (i.e., $\sum_j b_{ij} \leq 1$). A stochastic (a.k.a. Markov) matrix is a substochastic matrix whose row-sums are exactly one.

Note that in our definition above, we do not require B to be square.

Definition 2.2. Let $A \coloneqq (a_{ij})$ be an $m \times n$ complex matrix.

- (i) The digraph of A, denoted $\operatorname{graph} A$, is defined as follows:
  - (a) If A is square, $\operatorname{graph} A$ is a tuple (V, E) consisting of the vertex set $V \coloneqq \{1, \ldots, m\}$ and edge set $E \subset V \times V$ satisfying $(i, j) \in E$ if and only if $a_{ij} \neq 0$.
  - (b) If A is not square, $\operatorname{graph} A \coloneqq \operatorname{graph} A'$, where A' is the smallest square matrix obtained by appending rows or columns of zeros to A.
- (ii) A walk in $\operatorname{graph} A \equiv (V, E)$ is a nonempty finite sequence of edges (i_1, i_2), (i_2, i_3), $\ldots$, $(i_{\ell-1}, i_\ell)$ in E. The set of all walks in $\operatorname{graph} A$ is denoted $\operatorname{walks} A$.
- (iii) Let $p \in \operatorname{walks} A$. The length of p, denoted $|p|$, is the total number of edges in p. $\operatorname{head} p$ (resp., $\operatorname{last} p$) is the first (resp., last) vertex in p.

[1] In [4], the authors refer to w.c.d.d. L-matrices as *matrices of positive type*.

[2] (1.1) remains true if we replace "L-matrix" by "*Z-matrix* with nonnegative diagonal entries".

To simplify matters, hereafter we denote edges by $i \to j$ instead of (i, j) and walks by $i_1 \to i_2 \to \cdots \to i_\ell$ instead of (i_1, i_2), (i_2, i_3), ..., $(i_{\ell-1}, i_\ell)$. We use the terms "row" and "vertex" interchangeably.

Let $B \coloneqq (b_{ij})$ be an $m \times n$ substochastic matrix. We define the sets

$$\hat{J}(B) \coloneqq \left\{1 \leq i \leq m \colon \textstyle\sum_j b_{ij} < 1\right\},$$

$$\hat{P}_i(B) \coloneqq \left\{p \in \operatorname{walks} B \colon \operatorname{head} p = i \text{ and } \operatorname{last} p \in \hat{J}(B)\right\}.$$

It is understood that when we write $i \notin \hat{J}(B)$, we mean $i \in \hat{J}(B)^{\complement} \coloneqq \{1, \ldots, m\} \setminus \hat{J}(B)$. Note that if $\hat{J}(B)$ is empty, so too is $\hat{P}_i(B)$ for each i. We define the index of contraction associated with B by

$$\widehat{\operatorname{con}} B \coloneqq \max\left(0, \sup_{i \notin \hat{J}(B)} \left\{\inf_{p \in \hat{P}_i(B)} |p|\right\}\right) \tag{2.1}$$

subject to the conventions $\inf \emptyset = \infty$ and $\sup \emptyset = -\infty$. We will see shortly that the matrix B is convergent if and only if $\widehat{\operatorname{con}} B$ is finite.

Example 2.3. The $n \times n$ matrix B in Figure 2.1 satisfies $\hat{J}(B) = \{1\}$ and

$$\min_{p \in \hat{P}_i(B)} |p| = i - 1 \text{ for } i \notin \hat{J}(B).$$

It follows that $\widehat{\operatorname{con}} B = n - 1$.

An immediate consequence of the definition of the index of contraction follows.

Lemma 2.4. *Let B be an $m \times n$ substochastic matrix. If $m \leq n$ (resp., $m > n$) $\widehat{\operatorname{con}} B$ is either infinite or strictly less than m (resp., $n + 1$).*

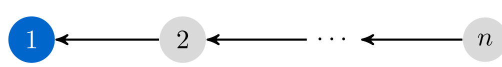
Proof. Suppose $m \leq n$. Let $i_1 \notin \hat{J}(B)$ and $p \coloneqq i_1 \to \cdots \to i_\ell$ be a walk in $\hat{P}_{i_1}(B)$. Since $i_\ell \in \hat{J}(B)$, it follows that $1 \leq i_\ell \leq m$. This implies that $1 \leq i_k \leq m$ for all k since by definition, $\operatorname{graph} B$ has no edges of the form $i \to j$, where $i > m$. Now, suppose $|p| \geq m$. By the pigeonhole principle, we can find integers u and v such that $1 \leq u < v \leq \ell$ and $i_u = i_v$. That is, the walk p contains a cycle (i.e., a subwalk starting and ending at the same vertex). "Removing" the cycle yields the new walk

$$p' \coloneqq i_1 \to i_2 \to \cdots \to i_u \to i_{v+1} \to i_{v+2} \to \cdots \to i_\ell$$

in $\hat{P}_{i_1}(B)$ satisfying $|p'| < |p|$. If $|p'| \geq m$, we can continue removing cycles until we arrive at a walk $p'' \in \hat{P}_{i_1}(B)$ satisfying $|p''| < m$. The case of $m > n$ is handled similarly. $\square$

$$\begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}$$

(A) The matrix B



(B) $\operatorname{graph} B$ (vertices in $\hat{J}(B)$ are highlighted)

FIGURE 2.1. An example of an $n \times n$ substochastic matrix and its graph

We are now ready to present our main result related to substochastic matrices. In the statement below, it is understood that if B is a square matrix, $B^0 = I$.

Theorem 2.5. *Let B be a square substochastic matrix. If $\alpha := \widehat{\operatorname{con}} B$ is finite,*

$$1 = \|B^0\|_\infty = \cdots = \|B^\alpha\|_\infty > \|B^{\alpha+1}\|_\infty \geq \|B^{\alpha+2}\|_\infty \geq \cdots .$$

Otherwise,

$$1 = \|B^1\|_\infty = \|B^2\|_\infty = \cdots .$$

Before giving a proof, it is useful to record some consequences of the above.

Corollary 2.6. *Let B be a square substochastic matrix. Then, its spectral radius is no larger than one. Moreover, the following statements are equivalent:*

(i) *$\widehat{\operatorname{con}} B$ is finite.*
(ii) *B is convergent.*
(iii) *$I - B$ is nonsingular.*

The above can be considered a generalization of the well-known result that a square stochastic (a.k.a. Markov) matrix has spectral radius no larger than one and at least one eigenvalue equal exactly to one (recall that for any matrix M, $I - M$ is singular if and only if $\lambda = 1$ is an eigenvalue of M).

Proof. The claim that the spectral radius of B is no larger than one in magnitude is a direct consequence of the fact that $\|B\|_\infty \leq 1$.

(i) $\implies$ (ii) follows immediately from Theorem 2.5, while (ii) $\implies$ (iii) is true for any matrix. We prove below, by contrapositive, the claim (iii) $\implies$ (i).

Suppose $\widehat{\operatorname{con}} B$ is infinite. Let R be the set of rows $i \notin \hat{J}(B)$ for which $\hat{P}_i(B)$ is empty. Due to our assumptions, there is at least one such row and hence R is nonempty. Without loss of generality, we may assume $R = \{1, \ldots, r\}$ for some $1 \leq r \leq n$, where n is the order of B (otherwise, replace B by $PBP^\intercal$, where P is an appropriately chosen permutation matrix). Let $e \in \mathbb{R}^r$ be the column vector whose entries are all one. If $r = n$, each row-sum of B is one (i.e., $Be = e$ so that $(I - B)e = 0$). Otherwise, B has the block structure

$$B = \left(\begin{array}{c|c} B_1 & 0 \\ \hline B_2 & B_3 \end{array}\right), \text{ where } B_1 \in \mathbb{R}^{r \times r}.$$

The partition above ensures that for each row $i \notin R$, $i \in \hat{J}(B)$ or $\hat{P}_i(B)$ is nonempty. Therefore, $\widehat{\operatorname{con}} B_3$ is finite, and hence the linear system $(I - B_3)x = B_2 e$ has a unique solution x. Moreover, since the row-sums of B_1 are one, $B_1 e = e$. Therefore,

$$(I - B)\begin{pmatrix} e \\ x \end{pmatrix} = \begin{pmatrix} e \\ x \end{pmatrix} - \begin{pmatrix} B_1 e \\ B_2 e + B_3 x \end{pmatrix} = \begin{pmatrix} e \\ x \end{pmatrix} - \begin{pmatrix} e \\ x \end{pmatrix} = 0. \qquad \square$$

Corollary 2.7. *A square irreducible substochastic matrix B is convergent if and only if $\hat{J}(B)$ is nonempty.*

The above result is well known. It can be obtained, for example, by [21, Corollary 1.19 and Lemma 2.8]. We give a short alternate proof using Corollary 2.6:

Proof. Since a square matrix is irreducible if and only if its digraph is strongly connected [21], $\widehat{\operatorname{con}} B$ is finite if and only if $\hat{J}(B)$ is nonempty. The result now follows from Corollary 2.6. $\square$

If B is a square substochastic matrix, we can always find a permutation matrix P and an integer $r \geq 1$ such that $PBP^\intercal$ has the block triangular structure

$$PBP^\intercal = \begin{pmatrix} B_{11} & B_{12} & \cdots & B_{1r} \\ & B_{22} & \cdots & B_{2r} \\ & & \ddots & \vdots \\ & & & B_{rr} \end{pmatrix}, \tag{2.2}$$

where each B_{ii} is a square substochastic matrix that is either irreducible or a 1×1 zero matrix (it is understood that if $r = 1$, then $B = B_{11}$). Following [7, 21], we refer to this as the *normal form* of B (it is shown in [7, p. 90] that the normal form of a matrix is unique up to permutations by blocks). Since $\det(PBP^\intercal - \lambda I) = \prod_i \det(B_{ii} - \lambda I)$, the spectrum of B satisfies

$$\sigma(B) = \sigma(B_{11}) \cup \cdots \cup \sigma(B_{rr}). \tag{2.3}$$

This observation motivates the next result.

Theorem 2.8. *Let B be a square substochastic matrix with normal form* (2.2)*. B is convergent if and only if $\hat{J}(B_{ii})$ is nonempty for each i. Moreover, if B is convergent,*

$$\max_i \{\widehat{\operatorname{con}} B_{ii}\} \leq \widehat{\operatorname{con}} B \leq N + \widehat{\operatorname{con}} B_{rr}, \tag{2.4}$$

where $N \coloneqq \sum_{i=1}^{r-1} n_i$ and n_i is the order of the matrix B_{ii} (it is understood that if $r = 1$, then $N = 0$).

Proof. The first claim is a consequence of Corollary 2.7 and (2.3).

We prove now the leftmost inequality in (2.4). First, note that $\|B^k\|_\infty = \|PB^kP^\intercal\|_\infty = \|(PBP^\intercal)^k\|_\infty$. Moreover, the block diagonal entries of $(PBP^\intercal)^k$ are the matrices $B_{11}^k, \ldots, B_{rr}^k$. Therefore, for each i, $\|B_{ii}^k\|_\infty \leq \|B^k\|_\infty$ and hence $\widehat{\operatorname{con}} B_{ii} \leq \widehat{\operatorname{con}} B$ by Theorem 2.5.

We prove now the rightmost inequality in (2.4). If $\widehat{\operatorname{con}} B \leq N$, the inequality is trivial. As such, we proceed assuming that $N < \widehat{\operatorname{con}} B < \infty$. First, note that $\widehat{\operatorname{con}} B = \widehat{\operatorname{con}}(PBP^\intercal)$. Therefore, $\widehat{\operatorname{con}} B = |p|$, where $p \in \operatorname{walks}(PBP^\intercal)$. Due to the block triangular structure of $PBP^\intercal$, we can write p as

$$p = i_1 \to \cdots \to i_u \to j_1 \to \cdots \to j_v,$$

where $u \leq N$ and $j_k > N$ for all k. Defining $j_k' \coloneqq j_k - N$, it follows that $p' \coloneqq j_1' \to \cdots \to j_v'$ is a walk whose length is no larger than any walk in $\hat{P}_{j_1'}(B_{rr})$, from which we obtain $|p'| \leq \widehat{\operatorname{con}} B_{rr}$. Therefore,

$$\widehat{\operatorname{con}} B = |p| \leq u + |p'| \leq N + \widehat{\operatorname{con}} B_{rr}. \qquad \square$$

Returning to our goal of proving Theorem 2.5, we first establish some lemmata related to substochastic matrices. The first lemma is a consequence of definitions and requires no proof.

Lemma 2.9. *Let B be an $m \times n$ substochastic matrix. Then, $\|B\|_\infty < 1$ if and only if $\hat{J}(B) = \{1, \ldots, m\}$.*

Lemma 2.10. *Let $B \coloneqq (b_{ij})$ and $C \coloneqq (c_{ij})$ be compatible (i.e., the product BC is well-defined) substochastic matrices. Then,*

(i) *BC is a substochastic matrix.*
(ii) *If $i \in \hat{J}(B)$, then $i \in \hat{J}(BC)$.*
(iii) *If $i \notin \hat{J}(B)$, then $i \in \hat{J}(BC)$ if and only if there exists $h \in \hat{J}(C)$ such that $i \to h$ is an edge in* $\operatorname{graph} B$.
(iv) *$i \to j$ is an edge in* $\operatorname{graph}(BC)$ *if and only if there exist edges $i \to h$ and $h \to j$ in* $\operatorname{graph} B$ *and* $\operatorname{graph} C$*, respectively.*

Proof.

(i) BC has nonnegative entries and $\|BCe\|_\infty \leq \|BC\|_\infty \leq \|B\|_\infty \|C\|_\infty \leq 1$.
(ii) Note first that $\sum_j [BC]_{ij} = \sum_j \sum_k b_{ik} c_{kj} = \sum_k b_{ik} \sum_j c_{kj} \leq \sum_k b_{ik}$. If $i \in \hat{J}(B)$, then $\sum_k b_{ik} < 1$ and the desired result follows.
(iii) Suppose $i \notin \hat{J}(B)$. If there exists $h \in \hat{J}(C)$ such that $i \to h$ is an edge in $\operatorname{graph} B$, then $\sum_j c_{hj} < 1$ and $\sum_j [BC]_{ij} = b_{ih} \sum_j c_{hj} + \sum_{k \neq h} b_{ik} \sum_j c_{kj} < \sum_k b_{ik} \leq 1$. Otherwise, $\sum_j c_{kj} = 1$ for all k with $b_{ik} \neq 0$ and hence $\sum_j [BC]_{ij} = \sum_k b_{ik} \sum_j c_{kj} = \sum_k b_{ik} = 1$.
(iv) Suppose $i \to h$ and $h \to j$ are edges in $\operatorname{graph} B$ and $\operatorname{graph} C$, respectively. Then, $[BC]_{ij} = \sum_k b_{ik} c_{kj} \geq b_{ih} c_{hj} > 0$. Otherwise, for each k, at least one of b_{ik} or c_{kj} is zero and hence $[BC]_{ij} = 0$. □

Lemma 2.11. *Let B be a square substochastic matrix, $i \notin \hat{J}(B)$, and k be a positive integer. Then, $i \in \hat{J}(B^k)$ if and only if there is a walk p in $\hat{P}_i(B)$ such that $|p| < k$.*

Proof. To simplify notation, let $i_1 \coloneqq i$.

Suppose there exists a walk $i_1 \to i_2 \to \cdots \to i_\ell$ in $\hat{P}_{i_1}(B)$. We claim that $i_1 \to i_\ell$ is an edge in $\operatorname{graph}(B^{\ell-1})$. If this is the case, Lemma 2.10 (ii) and (iii) guarantee that $i_1 \in \hat{J}(B^{\ell-1}B) = \hat{J}(B^\ell)$. If $\ell \leq k$, $i_1 \in \hat{J}(B^k)$ by Lemma 2.10 (ii), as desired.

We now return to the claim in the previous paragraph. Since the claim is trivial if $\ell = 2$, we proceed assuming $\ell > 2$. Let n be an integer satisfying $2 < n \leq \ell$. If $i_1 \to i_{n-1}$ is an edge in $\operatorname{graph}(B^{n-2})$, then since $i_{n-1} \to i_n$ is an edge in $\operatorname{graph} B$, Lemma 2.10 (iv) implies that $i_1 \to i_n$ is an edge in $\operatorname{graph}(B^{n-2}B) = \operatorname{graph}(B^{n-1})$. Since $i_1 \to i_2$ is an edge in $\operatorname{graph} B$, it follows by induction that $i_1 \to i_\ell$ is an edge in $\operatorname{graph}(B^{\ell-1})$, as desired.

As for the converse, suppose $i_1 \in \hat{J}(B^k)$. Let ℓ be the smallest positive integer such that $i_1 \notin \hat{J}(B^{\ell-1})$ and $i_1 \in \hat{J}(B^\ell)$. Since $i_1 \notin \hat{J}(B)$ and $i_1 \in \hat{J}(B^k)$, it follows that $\ell \leq k$. By Lemma 2.10 (iii), there exists $i_\ell \in \hat{J}(B)$ such that $i_1 \to i_\ell$ is an edge in $\operatorname{graph}(B^{\ell-1})$.

If $\ell = 2$, the trivial walk $i_1 \to i_\ell$ is in $\hat{P}_{i_1}(B)$, and hence we proceed assuming $\ell > 2$. Let n be an integer satisfying $2 < n \leq \ell$. If there exists a positive integer i_n such that $i_1 \to i_n$ is an edge in $\operatorname{graph}(B^{n-1}) = \operatorname{graph}(B^{n-2}B)$, Lemma 2.10 (iv) implies that there exists a positive integer i_{n-1} such that $i_1 \to i_{n-1}$ is an edge in $\operatorname{graph}(B^{n-2})$ and $i_{n-1} \to i_n$ is an edge in $\operatorname{graph} B$. Since $i_1 \to i_\ell$ is an edge in $\operatorname{graph}(B^{\ell-1})$, it follows by induction that $i_{n-1} \to i_n$ is an edge in $\operatorname{graph} B$ for each integer n satisfying $2 \leq n \leq \ell$. Therefore, $i_1 \to i_2 \to \cdots \to i_\ell$ is a walk in $\hat{P}_{i_1}(B)$, as desired. □

We are now ready to prove Theorem 2.5.

Proof of Theorem 2.5. Since $\|B^{k+1}\|_\infty \leq \|B^k\|_\infty \|B\|_\infty \leq \|B^k\|_\infty$, the inequalities $1 \geq \|B^1\|_\infty \geq \|B^2\|_\infty \geq \cdots$ follow trivially.

The remaining inequalities in the theorem statement follow by applying Lemma 2.11 to each row not in $\hat{J}(B)$ and invoking Lemma 2.9. □

2.2. **M-matrices.** In this subsection, we recall some well-known results on M-matrices (see, e.g., [2, Chapter 6]).

Definition 2.12. An M-matrix is a square matrix A that can be expressed in the form $A = sI - B$, where B is a nonnegative matrix and $s \geq \rho(B)$, where $\rho(B)$ is the spectral radius of B.

Definition 2.13. A Z-matrix is a real matrix with nonpositive off-diagonal entries.

Definition 2.14. An L-matrix is a Z-matrix with positive diagonal entries.

Proposition 2.15. *A nonsingular M-matrix is an L-matrix.*

Definition 2.16. Let A be a square real matrix. A is monotone if and only if it is nonsingular and its inverse consists only of nonnegative entries.

Proposition 2.17. *The following are equivalent:*

(i) *A is a nonsingular M-matrix.*
(ii) *A is a monotone Z-matrix.*

We close this subsection by introducing the following enlargement of the family of L-matrices (Definition 2.14), to be used in what follows.

Definition 2.18. An L_0-matrix is a Z-matrix with *nonnegative* diagonal entries.

2.3. **Weakly chained diagonally dominant (w.c.d.d.) matrices.** Before we can define w.c.d.d. matrices, we require some preliminary definitions.

Definition 2.19. Let $A \coloneqq (a_{ij})$ be a complex matrix.

(i) The ith row of A is w.d.d. (resp., s.d.d.) if $|a_{ii}| \geq \sum_{j \neq i} |a_{ij}|$ (resp., $>$).
(ii) A is w.d.d. (resp., s.d.d.) if all of its rows are w.d.d. (resp., s.d.d.).

Let $A \coloneqq (a_{ij})$ be an $m \times n$ complex w.d.d. matrix. We define the sets

$$J(A) \coloneqq \left\{1 \leq i \leq m\colon\ |a_{ii}| > \textstyle\sum_{j \neq i} |a_{ij}|\right\},$$
$$\text{and } P_i(A) \coloneqq \left\{p \in \operatorname{walks} A\colon\ \operatorname{head} p = i \text{ and } \operatorname{last} p \in J(A)\right\}.$$

Note that if $J(A)$ is empty, so too is $P_i(A)$ for each i. We will see shortly that the sets $J(\cdot)$ and $P_i(\cdot)$ are related to $\hat{J}(\cdot)$ and $\hat{P}_i(\cdot)$.

We are now ready to introduce w.c.d.d. matrices:

Definition 2.20. A square complex matrix A is w.c.d.d. if the points below are satisfied:

(i) A is w.d.d.
(ii) $J(A)$ is nonempty.
(iii) For each $i \notin J(A)$, $P_i(A)$ is nonempty.

We now define the index of *connectivity* associated with a square complex w.d.d. matrix A as

$$\operatorname{con} A \coloneqq \max\left(0, \sup_{i \notin J(A)} \left\{\inf_{p \in P_i(A)} |p|\right\}\right)$$

(compare this with the index of *contraction* $\widehat{\operatorname{con}}$ defined in (2.1)). The lemma below is a trivial consequence of the definitions above and as such requires no proof.

Lemma 2.21. *A square complex w.d.d. matrix A is w.c.d.d. if and only if* $\operatorname{con} A$ *is finite.*

We are now able to establish a duality between w.d.d. L-matrices (or more accurately, L_0-matrices) and substochastic matrices that, as we will see, connects the nonsingularity of the former to the convergence of the latter.

Lemma 2.22. *Let $A := (a_{ij})$ be an $n \times n$ w.d.d.* L_0*-matrix and $D := (d_{ij})$ be an $n \times n$ diagonal matrix whose diagonal entries are positive and satisfy $d_{ii} \leq 1/a_{ii}$ for each i such that $a_{ii} \neq 0$. Then, $B := I - DA$ is substochastic and*

$$\operatorname{con} A = \widehat{\operatorname{con}} B. \tag{2.5}$$

Conversely, let B be an $n \times n$ substochastic matrix and D be an $n \times n$ diagonal matrix whose diagonal entries are positive. Then, $A := D(I - B)$ is a w.d.d. L_0*-matrix and* (2.5) *holds.*

Proof. We prove only the first claim, the converse being handled similarly.

Let A and $B := I - DA$ be given as in the lemma statement. To simplify notation, denote by a_{ij} and b_{ij} the elements of A and B. First, note that $b_{ii} = 1 - d_{ii}a_{ii} \geq 0$ and $b_{ij} = -d_{ii}a_{ij} \geq 0$ whenever $i \neq j$. Since

$$\sum_j b_{ij} = 1 - \sum_j d_{ii}a_{ij} = 1 - d_{ii}\left(a_{ii} - \sum_{j \neq i} |a_{ij}|\right) \leq 1,$$

it follows that B is substochastic and $J(A) = \hat{J}(B)$. Letting graph $A \equiv (V, E)$ and graph $B \equiv (V', E')$, note that $V = V'$ and

$$E \setminus \{(i,i)\}_i \subset E' \subset E.$$

More concisely, graph B is simply graph A with zero or more self-loops (i.e., edges of the form $i \to i$) removed. As a result of these facts, (2.5) follows immediately. $\square$

Example 2.23. Let $A := (a_{ij})$ be a square w.d.d. L-matrix of order n and

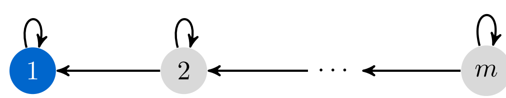
$$B_A := I - \operatorname{diag}(a_{11}, \ldots, a_{nn})^{-1} A$$

denote the point Jacobi matrix associated with A (cf. [21, Chapter 3]). By the previous results, A is w.c.d.d. if and only if $\operatorname{con} A = \widehat{\operatorname{con}} B_A$ is finite.

Note that the substochastic matrix in Figure 2.1 is the point Jacobi matrix associated with the w.d.d. L-matrix in Figure 2.2.

$$\begin{pmatrix} +1 & & & & \\ -1 & +1 & & & \\ & -1 & +1 & & \\ & & \ddots & \ddots & \\ & & & -1 & +1 \end{pmatrix}$$

(A) The matrix A



(B) graph A (vertices in $J(A)$ are highlighted)

FIGURE 2.2. An example of a w.c.d.d. matrix and its graph

We now restate and prove characterization (1.1) from the introduction.

Theorem 2.24. *The following are equivalent:*

(i) *A is a nonsingular w.d.d. M-matrix.*
(ii) *A is a nonsingular w.d.d. L-matrix.*
(iii) *A is a w.c.d.d. L-matrix.*

Since a nonsingular w.d.d. L_0-matrix must be an L-matrix, we can safely replace all occurrences of "L-matrix" with "L_0-matrix" in the above theorem without affecting its validity (recall that any w.c.d.d. matrix is nonsingular [18]).

Proof. (i) $\Longrightarrow$ (ii) follows from Proposition 2.15 while (iii) $\Longrightarrow$ (i) is established in [4, Theorem 2.2]. We prove below the claim (ii) $\Longrightarrow$ (iii).

Let $A := (a_{ij})$ be a nonsingular w.d.d. L-matrix of order n. Then, the associated point Jacobi matrix B_A is substochastic and $I - B_A$ is nonsingular since

$$I - B_A = \operatorname{diag}(a_{11}, \ldots, a_{nn})^{-1} A.$$

Corollary 2.6 and Lemma 2.22 imply that $\operatorname{con} A = \widehat{\operatorname{con}} B_A$ is finite. Therefore, by Lemma 2.21, A is w.c.d.d. $\square$

Remark 2.25. Instead of calling upon the results of [4], it is also possible to prove (iii) $\Longrightarrow$ (i) of Theorem 2.24 directly by using arguments involving the index of contraction. In particular, let A be a w.c.d.d. L-matrix of order n. Then, by Lemma 2.21 and Lemma 2.22, the associated point Jacobi matrix B_A is substochastic with $\widehat{\operatorname{con}} B_A = \operatorname{con} A$ finite. By Corollary 2.6, B_A is convergent and hence the Neumann series $I + B_A + B_A^2 + \cdots$ for the inverse of $I - B_A$ converges to a matrix whose entries are nonnegative. Therefore, A is monotone by Definition 2.16, and hence a nonsingular M-matrix by Proposition 2.17.

An immediate consequence of Theorem 2.24, which can be considered an analogue of Corollary 2.7, is given below.

Corollary 2.26. *A square irreducible w.d.d. L-matrix A is a nonsingular M-matrix if and only if $J(A)$ is nonempty.*

While the reverse direction in the above result is well known [21, Corollary 3.20], we are not aware of a reference for the forward direction.

3. Computing the index of contraction

In this section, we present a procedure to compute the index of contraction $\widehat{\operatorname{con}} B$ of a substochastic matrix B and show that it is robust in the presence of inexact (i.e., floating point) arithmetic.

By the results of the previous section, such a procedure can also be used to determine if an arbitrary w.d.d. matrix A is a nonsingular M-matrix as follows. If A is not a square L-matrix, it is trivially not a nonsingular M-matrix (Proposition 2.15). Otherwise, we can check the finitude of the index of contraction of its associated point Jacobi matrix B_A to determine whether or not A is a nonsingular M-matrix (recall Example 2.23 and Theorem 2.24).

3.1. **The procedure.** Before we can describe the procedure, we require the notion of a vertex contraction (a.k.a. vertex identification), a generalization of the well-known notion of edge contraction from graph theory.

Definition 3.1. Let $G \equiv (V, E)$ be a graph, $W \subset V$, w denote a new vertex (i.e., $w \notin V$), and f be a function which maps every vertex in $V \setminus W$ to itself and every vertex in W to w (i.e., $f|_{V \setminus W} = \mathrm{id}_{V \setminus W}$ and $f|_W(\cdot) = w$). The vertex contraction of G with respect to W is a new graph $G' \equiv (V', E')$, where $V' \coloneqq (V \setminus W) \cup \{w\}$ and $E' \coloneqq \{(f(i), f(j)) \colon (i, j) \in E\}$.

An overview of the procedure for computing the index of contraction for an arbitrary substochastic matrix B is given below:

(1) Obtain the vertex contraction of graph B with respect to $\hat{J}(B)$. Label the new vertex in the contraction $w = 0$ and the new vertex set V'. Note that $V' = \hat{J}(B)^{\complement} \cup \{0\}$ (recall that the superscript $\complement$ denotes complement).
(2) Reverse all edges in the resulting graph (see, e.g., Figure 3.1).
(3) In the resulting graph, find the shortest distances $d(i)$ from the new vertex 0 to all vertices $i \in V'$ by a breadth-first search (BFS) starting at 0. It is understood that $d(0) = 0$ and that if i is unvisited in the BFS, $d(i) = \infty$.
(4) Return $\max_{i \in V'} d(i)$.

That this procedure terminates is trivial (BFS is performed on a graph with finitely many vertices). As for the correctness of the procedure, it is easy to verify that

$$d(i) = \inf_{p \in \hat{P}_i(B)} |p| \text{ for } i \notin \hat{J}(B)$$

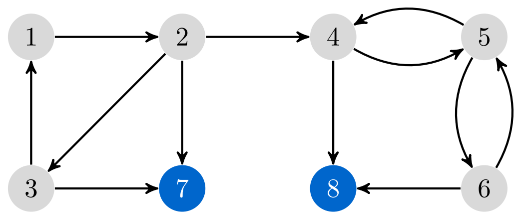
so that $\widehat{\mathrm{con}} B = \max(0, \sup_{i \notin \hat{J}(B)} d(i)) = \max_{i \in V'} d(i)$.

Remark 3.2. Since BFS does not revisit vertices, the correctness of the procedure is unaffected if graph B is preprocessed to remove self-loops (i.e., edges of the form $i \to i$) and edges of the form $i \to j$ with $i \in \hat{J}(B)$.

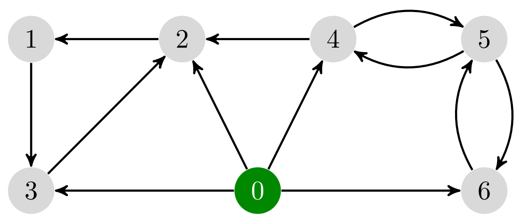
Algorithm 1 gives precise pseudocode for steps (1) to (4). Without loss of generality, it is assumed that the input matrix is square (the rectangular case is obtained by a few trivial additions to the code). The pseudocode makes use of the *list* and *queue* data structures (see, e.g., [6, Chapter 10]). The operation L.add(x) appends the element x to the list L. The operation Q.enqueue(x) adds the element x to the back of the queue Q. The operation Q.dequeue() removes and returns the element at the front of the queue Q.

It is obvious that if the input to Algorithm 1 is a dense matrix of order n, $\Theta(n^2)$ operations are required. Suppose instead that we restrict our inputs to matrices $B \coloneqq (b_{ij})$ that are *sparse* in the sense that $\mathrm{nnz} \coloneqq \max_i |\{j \colon b_{ij} \neq 0\}|$, the maximum number of nonzero entries per row, is bounded independent of n (i.e., $\mathrm{nnz} = \Theta(1)$ as $n \to \infty$). If the matrices are stored in an appropriate format (e.g., compressed sparse row (CSR) format, Ellpack-Itpack, etc. [17]), the loops on lines 6 and 20 require only a constant number of iterations for each fixed i. In this case, $\Theta(n)$ operations are required. An obvious generalization of this fact is that if $\mathrm{nnz} = O(f(n))$, $O(nf(n))$ operations are required.

3.2. **Floating point arithmetic considerations.** The loop on line 6 of Algorithm 1 computes the ith row-sum of the substochastic matrix $B \coloneqq (b_{ij})$. In the



(A) graph B (vertices in $\hat{J}(B)$ are highlighted)



(B) The resulting graph

FIGURE 3.1. Steps (1) and (2) applied to an example

Algorithm 1 Computing the index of contraction of a square substochastic matrix

Input: a square substochastic matrix $B \coloneqq (b_{ij})_{1 \le i,j \le n}$ of order n
Output: $\widehat{\operatorname{con}}B$

// Find all rows in $\hat{J}(B)$
$s \leftarrow 0$
$S[1, \ldots, n] \leftarrow$ **new** array of bools
for all rows i **do**
  $t \leftarrow 0$
  for all cols j s.t. $b_{ij} \neq 0$ **do**
    $t \leftarrow t + b_{ij}$
  end for
  if $t < 1$ **then**
    $s \leftarrow s + 1$
    $S[i] \leftarrow$ **true** // $i \in \hat{J}(B)$
  else
    $S[i] \leftarrow$ **false** // $i \notin \hat{J}(B)$
  end if
end for

// Find neighbours of each vertex (ignoring extraneous edges as per Remark 3.2)
$N[0, \ldots, n] \leftarrow$ **new** array of lists
for all rows i s.t. $S[i] =$ **false** **do**
  for all cols $j \neq i$ s.t. $b_{ij} \neq 0$ **do**
    if $S[j] =$ **true** **then**
      $N[0]$.add(i)
    else
      $N[j]$.add(i)
    end if
  end for
end for

// Perform BFS starting at 0
result $\leftarrow 0$
$Q \leftarrow$ **new** queue
Q.enqueue$((0, 0))$
while Q is not empty **do**
  $(j, d) \leftarrow Q$.dequeue()
  result $\leftarrow \max(\text{result}, d)$
  for all i in $N[j]$ s.t. $S[i] =$ **false** **do**
    $s \leftarrow s + 1$
    $S[i] \leftarrow$ **true**
    Q.enqueue$((i, d + 1))$
  end for
end while

if $s = n$ **then**
  $\widehat{\operatorname{con}}B \leftarrow$ result
else
  $\widehat{\operatorname{con}}B \leftarrow \infty$
end if

presence of floating point arithmetic, the operation $t + b_{ij}$ on line 7 can introduce error into calculations. In order to analyze this error, we take the standard model of floating point arithmetic in which floating point addition introduces error proportional to the size of the result:

$$fl(x+y) = (x+y)\left(1+\delta_{x,y}\right), \text{ where } |\delta_{x,y}| \leq \epsilon. \tag{3.1}$$

$\epsilon > 0$ is a machine-dependent constant (often referred to as machine epsilon) which gives an upper bound on the relative error due to rounding. In performing our analyses, we make the standard assumptions that the order n of the input matrix B satisfies $n\epsilon \leq 1$ [9] and that the entries of B are floating point numbers.

A floating point implementation of the loop on line 6 is represented by the recurrence $S_j := fl(S_{j-1} + b_{ij})$ with initial condition $S_0 := 0$. Letting $\gamma_k := k\epsilon/(1-k\epsilon)$, this direct implementation has an error bound of [9, Eq. (2.6)]

$$\left| S_n - \sum_j b_{ij} \right| \leq \gamma_{\text{nnz}-1} \sum_j b_{ij} \leq \gamma_{\text{nnz}-1}. \tag{3.2}$$

Recall that nnz is the maximum number of nonzero entries per row of the matrix B. If the matrix B is sparse (i.e., $\text{nnz} = \Theta(1)$ as $n \to \infty$), we obtain

$$\gamma_{\text{nnz}-1} = (\text{nnz}-1)\epsilon + O(\epsilon^2) \text{ as } \epsilon \to 0$$

by the power series representation of γ_k. In this case, for each i, the *absolute* error in computing $\sum_j b_{ij}$ is independent of n.

Note that if the *exact* value of $\sum_j b_{ij}$ is close to 1, the comparison $t < 1$ on line 9 may return either a false-positive or a false-negative. Motivated by (3.2), an implementation of Algorithm 1 should use instead the condition $t < 1 - \text{tol}$, where tol is a small constant strictly larger than $\gamma_{\text{nnz}-1}$ to preclude the possibility that the condition evaluates to true when the *exact* value of $\sum_j b_{ij}$ is 1 (for simplicity, we assume $1 - \text{tol}$ has a precise floating point representation). Then, the error bound (3.2) and discussion above yield the accuracy result below.

Lemma 3.3. *Let B be a substochastic matrix with at most* nnz *nonzero entries per row. Denoting by $(\widehat{\text{con}}B)_{fl}$ the quantity computed by Algorithm* 1 *under the standard model of floating point arithmetic* (3.1) *and with condition $t < 1$ replaced by $t < 1 - \text{tol}$, where $\text{tol} > \gamma_{\text{nnz}-1}$, the following results hold:*

(i) *if $\widehat{\text{con}}B = \infty$, then $(\widehat{\text{con}}B)_{fl} = \widehat{\text{con}}B$.*

(ii) *if $\widehat{\text{con}}B \neq \infty$ and $\sum_j b_{ij} \leq 1 - 2\,\text{tol}$ for $i \in \hat{J}(B)$, then $(\widehat{\text{con}}B)_{fl} = \widehat{\text{con}}B$.*

Remark 3.4. If B is not sparse, the error (3.2) depends on n. In this case, one should substitute the naïve summation outlined by the loop on line 6 for a more scalable algorithm, such as Kahan's summation algorithm, whose absolute error in approximating $\sum_j b_{ij}$, is $(2\epsilon + O(n\epsilon^2))\sum_j b_{ij} \leq 2\epsilon + O(n\epsilon^2)$ [9, Eq. (3.11)], which is independent of n due to the assumption $n\epsilon \leq 1$. We can obtain an analogue of Lemma 3.3 under Kahan summation by choosing tol appropriately.

Note that Lemma 3.3 suggests that the value of $(\widehat{\text{con}}B)_{fl}$ and $\widehat{\text{con}}B$ may disagree in certain cases. Fortunately, as demonstrated in the next example, this occurs only if the matrix B is "nearly nonconvergent" (i.e., $\rho(B) = 1 - \epsilon_0$, where $\epsilon_0 > 0$ is close to zero). This error may even be considered desirable behaviour since a nearly nonconvergent matrix may not be convergent in the presence of floating point error.

Example 3.5. Consider the $n \times n$ matrix

$$B_\nu := \begin{pmatrix} & 1 & & & & \\ & & 1 & & & \\ & & & \ddots & & \\ & & & & 1 & \\ & & & & & 1 \\ 1/n - \nu & 1/n & 1/n & \cdots & 1/n & 1/n \end{pmatrix},$$

where $0 < \nu \leq 1/n$. Note that even though $\widehat{\operatorname{con}} B_\nu = n - 1$ independent of the value of ν, $\rho(B_\nu) \to 1$ as $\nu \to 0$. When ν is very close to zero, floating point error may cause Algorithm 1 to erroneously determine that $\hat{J}(B_\nu)$ is empty and thereby mistakenly conclude that the index of contraction is infinite.

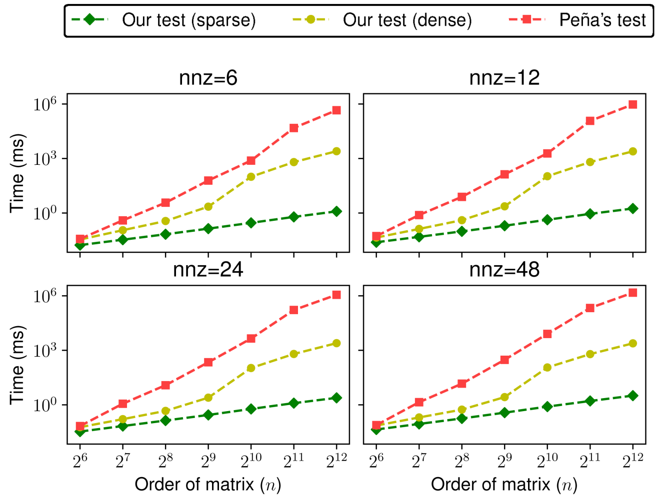
We close this section by discussing stability. The test in [15], which determines if an arbitrary matrix is a nonsingular M-matrix, uses a modified Gaussian elimination procedure. As such, to establish numerical stability, the author proves that the *growth factor* (see the definition in [8]) of the test is bounded by the order of the input matrix [15, Theorem 3.1]. In our case, the floating point error made in computing $\sum_i b_{ij}$ has no bearing on the error made in computing $\sum_{i'} b_{i'j}$ for distinct rows i and i'. That is, floating point errors do not propagate from row to row. Moreover, as demonstrated in the previous paragraphs, the error made in computing each row-sum can be bounded by a constant (without any additional effort in the sparse case, and with, e.g., Kahan summation in the dense case). As such, we conclude that Algorithm 1 is stable in the sense that it does not involve numbers that grow large due to floating point error.

4. Numerical experiments

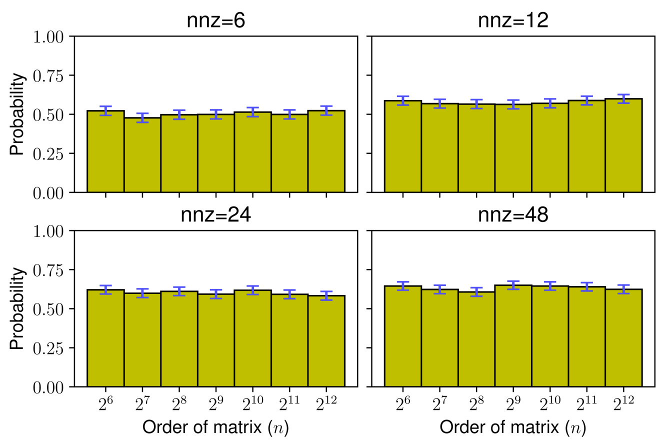
In this section, we compare the efficiency of *our test* described at the beginning of Section 3 to *Peña's test* detailed in [15]. To minimize bias, we run the tests on randomly sampled matrices (sampled according to the procedure in Appendix A).

We run the tests on matrices whose maximum number of nonzeros per row (nnz) are 6, 12, 24, and 48. We employ two versions of our test: a *sparse* version, in which the matrices are stored in compressed sparse row (CSR) format, and a *dense* version, in which the matrices are two-dimensional arrays. All tests are performed on an Intel Xeon E5440 2.83GHz CPU. The average time to process a randomly sampled matrix is shown in Figure 4.1a (error bars are omitted as even the 99% confidence interval is too small to be visible). We mention that in terms of accuracy, the tests produced the same results on all randomly sampled matrices (Figure 4.1b).

Figure 4.1a suggests that our test outperforms Peña's. Even for the experiments involving the 1024×1024 sparse matrices (small by most scientific computing standards), our sparse implementation executes on the order of *tenths of milliseconds* while Peña's test executes on the order of *seconds*.



(A) Timing results ($\log_{10}$ scale on time axis)



(B) Probability that a randomly sampled matrix is a nonsingular M-matrix (99% confidence intervals shown)

Figure 4.1.

Appendix A. Sampling procedure

This appendix details the procedure (employed in the numerical experiments of Section 4) used to randomly sample w.d.d. L_0-matrices. The procedure, for which pseudocode is given in Algorithm 2, works by sampling a matrix $B := (b_{ij})_{1\leq i,j\leq n}$ from the space of substochastic matrices and returning $I - B$, which is a w.d.d. L_0-matrix by Lemma 2.22.

Algorithm 2 Sampling a matrix A from the space of w.d.d. L_0-matrices

Input: positive integers n and $\mathrm{nnz} \leq n$
Output: matrix A

// Initialize zero matrix
$B \equiv (b_{ij}) \leftarrow 0$

for i from 1 to n **do**
    // Determine the number of nonzero entries m in row i
    $m \sim \mathrm{Unif}\{1, \ldots, \mathrm{nnz}\}$

    // Determine the row-sum of row i (less than one with probability $1/n$)
    $u \sim \mathrm{Unif}[0, 1]$
    if $u < 1/n$ **then**
        $s \sim \mathrm{Unif}[0, 1]$
    else
        $s \leftarrow 1$
    end if

    // Determine the indices j_k for which b_{ij_k} is nonzero by uniformly sampling $\{1, \ldots, n\}$ without replacement
    $\mathcal{A} \leftarrow \{1, \ldots, n\}$
    $j_1 \sim \mathrm{Unif}\,\mathcal{A}$
    for k from 2 to m **do**
        $\mathcal{A} \leftarrow \mathcal{A} \setminus \{j_{k-1}\}$
        $j_k \sim \mathrm{Unif}\,\mathcal{A}$
    end for

    // Determine the values of the nonzero entries in row i
    if $m \geq 2$ **then**
        $\alpha \leftarrow (1, \ldots, 1) \in \mathbb{R}^m$
        $(b_{ij_2}, \ldots, b_{ij_m}) \sim \mathrm{Dir}\,\alpha$
    end if
    $b_{ij_1} \leftarrow s$
    for k from 2 to m **do**
        $b_{ij_k} \leftarrow s b_{ij_k}$
        $b_{ij_1} \leftarrow b_{ij_1} - b_{ij_k}$
    end for
end for

// Make a w.d.d. L_0-matrix from the substochastic matrix B
$A \leftarrow I - B$

We use $\mathrm{Unif}\,\Omega$ to denote a uniform distribution on the sample space Ω. For $\alpha \in \mathbb{R}^m$, we use $\mathrm{Dir}\,\alpha$ to denote a Dirichlet distribution of order m with parameter α. It is well known that when α is a vector whose entries are all one, $\mathrm{Dir}\,\alpha$ is a uniform distribution over the unit simplex in $\mathbb{R}^{m-1}$. We use $x \sim \mathcal{D}$ to mean that x is a sample drawn from the distribution $\mathcal{D}$.

The inputs to the procedure are a positive integer n corresponding to the order of the output matrix and a positive integer $\mathrm{nnz} \leq n$ corresponding to the maximum number of nonzero entries per row.

Appendix B. Generalizing Theorem 2.5

This appendix generalizes Theorem 2.5. To present the generalization, we first extend our notion of walks.

Definition B.1. Let $(A_n)_{n\geq 1}$ be a sequence of compatible complex matrices (i.e., the product $A_k A_{k+1}$ is well-defined for each k).

(i) A walk in $(A_n)_n$ is a nonempty finite sequence of edges (i_1, i_2), (i_2, i_3), ..., $(i_{\ell-1}, i_\ell)$ such that each (i_k, i_{k+1}) is an edge in graph A_k. The set of all walks in $(A_n)_n$ is denoted $\mathrm{walks}(A_1, A_2, \ldots)$.

(ii) For $p \in \operatorname{walks}(A_1, A_2, \ldots)$, $\operatorname{head} p$, $\operatorname{last} p$, and $|p|$ are defined in the obvious way.

Note, in particular, that if we fix a square complex matrix A, we are returned to the original definition of a walk given in Section 2 if we take $A_n \coloneqq A$ for all n.

It is also useful to generalize the sets $\hat{P}_i(\cdot)$ of Section 2. In particular, given a sequence $(B_n)_{n\geq 1}$ of compatible substochastic matrices, let

$$\hat{P}_i(B_1, B_2, \ldots) \coloneqq \left\{ p \in \operatorname{walks}(B_1, B_2, \ldots) \colon \operatorname{head} p = i \text{ and } \operatorname{last} p \in \hat{J}(B_{|p|+1}) \right\}.$$

We are now ready to give the generalization.

Theorem B.2. *Let $(B_n)_{n\geq 1}$ be a sequence of compatible substochastic matrices, $(C_n)_{n\geq 0}$ be defined by $C_0 \coloneqq I$ and $C_n \coloneqq B_1 B_2 \cdots B_n$ whenever n is a positive integer, and*

$$\widehat{\operatorname{con}}(B_1, B_2, \ldots) \coloneqq \max \left(0, \sup_{i \notin \hat{J}(B_1)} \left\{ \inf_{p \in \hat{P}_i(B_1, B_2, \ldots)} |p| \right\} \right).$$

If $\alpha \coloneqq \widehat{\operatorname{con}}(B_1, B_2, \ldots)$ is finite, then

$$1 = \|C_0\|_\infty = \cdots = \|C_\alpha\|_\infty > \|C_{\alpha+1}\|_\infty \geq \|C_{\alpha+2}\|_\infty \geq \cdots .$$

Otherwise,

$$1 = \|C_1\|_\infty = \|C_2\|_\infty = \cdots .$$

The proof of the above is nearly identical to that of Theorem 2.5, requiring only a simple generalization of Lemma 2.11. However, in this general case, the finitude of the index of contraction is no longer an indicator of convergence.

Example B.3. Let $(B_n)_{n\geq 1}$ be a sequence of compatible substochastic matrices satisfying $B_n = (1 - 1/2^n)I$ and $(C_n)_{n\geq 0}$ be defined as above. Clearly, each matrix B_n is convergent, but $\|C_n\|_\infty = \prod_{k=1}^n (1 - 1/2^k) \nrightarrow 0$ as $n \to \infty$.

Moreover, even if each B_n is itself convergent, it is still possible that the index of contraction is infinite:

Example B.4. Let $(B_n)_{n\geq 1}$ be given by

$$B_n \coloneqq \frac{1}{2} \begin{pmatrix} 0 & 1 + (-1)^n \\ 1 - (-1)^n & 0 \end{pmatrix}.$$

Defining $(C_n)_{n\geq 0}$ as above, we find that

$$C_n \coloneqq \frac{1}{2} \begin{pmatrix} 0 & 0 \\ 1 - (-1)^n & 1 + (-1)^n \end{pmatrix} \text{ for } n \geq 1.$$

That is, $\|C_n\|_\infty = 1$ independent of n.

It is not hard to find interesting cases in which $\widehat{\operatorname{con}}(B_1, B_2, \ldots)$ is finite.

Example B.5. Let $(B_n)_{n\geq 1}$ be a sequence of square substochastic matrices of order n satisfying the following properties:

(i) B_1 is convergent.
(ii) $\hat{J}(B_1) = \hat{J}(B_n)$ and $\operatorname{graph} B_1 = \operatorname{graph} B_n$ for all n.

Then, $\widehat{\operatorname{con}}(B_1, B_2, \ldots) < n$.

Acknowledgments

The author thanks Edward Cheung (University of Waterloo) for discussions on M-matrices, a careful review of this document, and (more importantly) his unfaltering friendship.

References

[1] P. Azimzadeh and P. A. Forsyth, *Weakly chained matrices, policy iteration, and impulse control*, SIAM J. Numer. Anal. **54** (2016), no. 3, 1341–1364. MR3493959

[2] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*, Classics in Applied Mathematics, vol. 9, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1994. Revised reprint of the 1979 original. MR1298430

[3] O. Bokanowski, S. Maroso, and H. Zidani, *Some convergence results for Howard's algorithm*, SIAM J. Numer. Anal. **47** (2009), no. 4, 3001–3026. MR2551155

[4] J. H. Bramble and B. E. Hubbard, *On a finite difference analogue of an elliptic boundary problem which is neither diagonally dominant nor of non-negative type*, J. Math. and Phys. **43** (1964), 117–132. MR0162367

[5] G.-H. Cheng and T.-Z. Huang, *An upper bound for $\|A^{-1}\|_\infty$ of strictly diagonally dominant M-matrices*, Linear Algebra Appl. **426** (2007), no. 2-3, 667–673. MR2350685

[6] T. H. Cormen, C. E. Leiserson, R. L. Rivest, and C. Stein, *Introduction to Algorithms*, 3rd ed., MIT Press, Cambridge, MA, 2009. MR2572804

[7] F. R. Gantmacher, *Applications of the Theory of Matrices*, Translated by J. L. Brenner, with the assistance of D. W. Bushaw and S. Evanusa, Interscience Publishers, Inc., New York; Interscience Publishers Ltd., London, 1959. MR0107648

[8] G. H. Golub and C. F. Van Loan, *Matrix Computations*, 4th ed., Johns Hopkins Studies in the Mathematical Sciences, Johns Hopkins University Press, Baltimore, MD, 2013. MR3024913

[9] N. J. Higham, *The accuracy of floating point summation*, SIAM J. Sci. Comput. **14** (1993), no. 4, 783–799. MR1223274

[10] T.-Z. Huang and Y. Zhu, *Estimation of $\|A^{-1}\|_\infty$ for weakly chained diagonally dominant M-matrices*, Linear Algebra Appl. **432** (2010), no. 2-3, 670–677. MR2577710

[11] C. Li and Y. Li, *Weakly chained diagonally dominant B-matrices and error bounds for linear complementarity problems*, Numer. Algorithms **73** (2016), no. 4, 985–998. MR3579726

[12] C. Li, Y. Li, and R. Zhao, *New inequalities for the minimum eigenvalue of M-matrices*, Linear Multilinear Algebra **61** (2013), no. 9, 1267–1279. MR3175363

[13] C. Li, R. Ma, Q. Liu, and Y. Li, *Subdirect sums of weakly chained diagonally dominant matrices*, Linear Multilinear Algebra **65** (2017), no. 6, 1220–1231. MR3615539

[14] W. Li, *The infinity norm bound for the inverse of nonsingular diagonal dominant matrices*, Appl. Math. Lett. **21** (2008), no. 3, 258–263. MR2433738

[15] J. M. Peña, *A stable test to check if a matrix is a nonsingular M-matrix*, Math. Comp. **73** (2004), no. 247, 1385–1392. MR2047092

[16] R. J. Plemmons, *M-matrix characterizations. I. Nonsingular M-matrices*, Linear Algebra and Appl. **18** (1977), no. 2, 175–188. MR0444681

[17] Y. Saad, *Iterative Methods for Sparse Linear Systems*, 2nd ed., Society for Industrial and Applied Mathematics, Philadelphia, PA, 2003. MR1990645

[18] P. N. Shivakumar and K. H. Chew, *A sufficient condition for nonvanishing of determinants*, Proc. Amer. Math. Soc. **43** (1974), 63–66. MR0332820

[19] P. N. Shivakumar, J. J. Williams, Q. Ye, and C. A. Marinov, *On two-sided bounds related to weakly diagonally dominant M-matrices with application to digital circuit dynamics*, SIAM J. Matrix Anal. Appl. **17** (1996), no. 2, 298–312. MR1384509

[20] G.-X. Tian and T.-Z. Huang, *Inequalities for the minimum eigenvalue of M-matrices*, Electron. J. Linear Algebra **20** (2010), 291–302. MR2653540

[21] R. S. Varga, *Matrix Iterative Analysis*, Second revised and expanded edition, Springer Series in Computational Mathematics, vol. 27, Springer-Verlag, Berlin, 2000. MR1753713

[22] F. Wang and D.-s. Sun, *Some new inequalities for the minimum eigenvalue of M-matrices*, J. Inequal. Appl. (2015), 2015:195, 7. MR3357017

[23] P. Wang, *An upper bound for $\|A^{-1}\|_\infty$ of strictly diagonally dominant M-matrices*, Linear Algebra Appl. **431** (2009), no. 5-7, 511–517. MR2535528

[24] M. Xu, S. Li, and C. Li, *Inequalities for the minimum eigenvalue of doubly strictly diagonally dominant M-matrices*, J. Appl. Math. (2014), Art. ID 535716, 8. MR3253625

[25] Z. Yang, B. Zheng, and X. Liu, *A new upper bound for $\|A^{-1}\|$ of a strictly α-diagonally dominant M-matrix*, Adv. Numer. Anal. (2013), Art. ID 980615, 6. MR3149442

[26] J. Zhao and C. Sang, *Several new inequalities for the minimum eigenvalue of M-matrices*, J. Inequal. Appl. (2016), Paper No. 119, 9. MR3486018

David R. Cheriton School of Computer Science, University of Waterloo, Waterloo ON, Canada N2L 3G1

Current address: Department of Mathematics, University of Michigan, East Hall, 530 Church Street, Ann Arbor, Michigan 48109

Email address: parsiad@umich.edu