THE MYSTERY OF ASYNCHRONOUS ITERATIONS CONVERGENCE WHEN THE SPECTRAL RADIUS IS ONE

Daniel B. Szyld

Report 98-102
October 1998

This report is available in the World Wide Web at `http://www.math.temple.edu/~szyld`
or via anonymous ftp at `ftp.math.temple.edu` in directory `pub/szyld`.

THE MYSTERY OF ASYNCHRONOUS ITERATIONS CONVERGENCE WHEN THE SPECTRAL RADIUS IS ONE

DANIEL B. SZYLD*

Abstract. Chazan and Miranker [*Linear Algebra Appl.* **2** (1969) 199-222], and other authors since, have shown that a necessary and sufficient condition for asynchronous iterations to converge is that the spectral radius of the absolute value of the iteration matrix is strictly less than one. Nevertheless, several authors, including Lubachevsky and Mitra [*J. ACM* **33** (1986) 130–150], show convergence for asynchronous iterations for singular matrices and matrices representing Markov chains when the spectral radius of the nonnegative iteration matrix is exactly one. In this note, this apparent contradiction is resolved. It is shown that in fact, the spectral radius less than one is a sufficient condition. The necessity is a more subtle issue. Under certain conditions, with spectral radius equal to one, convergence can indeed be achieved.

The problem we focus in this note is the solution of $n \times n$ singular linear systems of equations of the form

$$Ax = b \tag{1}$$

by asynchronous iterations. In particular, when $A = I - B$, and B is the stochastic matrix representing a Markov chain, the solution of (1), for $b = 0$, is a stationary probability distribution of the Markov chain (normalized so that $x^T e = 1$ with e having all entries equal to one); see, e.g., [5], [18]. In this case, $\rho(B) = 1$, where $\rho(B)$ denotes the spectral radius of B.

Iterative methods for the solution of (1) based on splittings of the form $A = M - N$, where M is nonsingular, have been successfully used for this problem; see, e.g., [2], [3], [13], [15], [16]. These methods include point and block versions of the classical Jacobi and Gauss-Seidel methods [5], [24], and can be written as the following iteration, starting from an initial vector x^0,

$$x^{k+1} = Tx^k + M^{-1}b. \tag{2}$$

The matrix $T = M^{-1}N$ is called the iteration matrix, and it is generally assumed to be nonnegative (denoted $T \geq O$), e.g., when the splittings are weak regular [5]. Since $A = M(I-T)$ it follows that A singular implies that 1 is an eigenvalue of T, and $\rho(T) = 1$ implies that A is singular. Thus, for Markov chains in particular, we have that for any splitting $A = M - N$, $\rho(T) = 1$. For simplicity, we assume from now on that $b = 0$, and that $A = I - T$, where T is the iteration matrix of the method chosen in (2).

* Department of Mathematics, Temple University (038-16), 1805 N. Broad Street, Philadelphia, Pennsylvania 19122-6094, USA (szyld@math.temple.edu). This work was supported by the National Science Foundation grant DMS-9625865. Results of this study were first presented at the meeting of the International Linear Algebra Society held in Madison, Wisconsin, on June 3–6, 1998.

In parallel asynchronous methods, using p processors, the variables and equations are most often partitioned into p disjoint sets (without overlap), and after a suitable permutation we have $x^T = (x_1^T, x_2^T, \dots, x_p^T)$, where $x_\ell \in \mathbb{R}^{n_\ell}$, $\sum_{\ell=1}^{p} n_\ell = n$, and

$$T = \begin{bmatrix} T_{11} & T_{12} & \cdots & T_{1p} \\ T_{21} & T_{22} & \cdots & T_{2p} \\ \vdots & \vdots & \ddots & \vdots \\ T_{p1} & T_{p2} & \cdots & T_{pp} \end{bmatrix},$$

where $T_{k\ell} \in \mathbb{R}^{n_k \times n_\ell}$.

The following is a general computational model for such a method. It is general in the sense that it can correspond to different types of architectures, including shared or distributed memory machines; see [22] for more details.

Given an initial approximation x to the solution of (1), each processor of a parallel computer executes the following procedure, independently of the others.

1. `determine` ℓ (the block of variables this processor updates at this time).
2. `read` x_k, $k \neq \ell$.
3. `set (or approximate)` $x_\ell = \sum_{k=1}^{p} T_{\ell k} x_k$.
4. `write` x_ℓ.

Steps `1` to `4` are repeated until some termination or stopping criterion is met.

In order to mathematically describe these methods, many authors use an index i to mark a time when one or more variables are updated, and define sets J_i containing the indices of the blocks updated at time i, i.e., if i is the time of completion of step 4, then $\ell \in J_i$, and the block x_ℓ is tagged with the time i, denoted x_ℓ^i. The components of x read at the time of step 2 (x_k, $k \neq \ell$), have tags previous to the 'time' i, we call them $r(k,i)$. With this notation, given an initial approximation x^0 to the solution of (1), the asynchronous method is represented by the following equations, for $i = 1, \dots,$

$$x_\ell^i = \begin{cases} x_\ell^{i-1} & \text{if } \ell \notin J_i \\ x_\ell^i = \sum_{k=1}^{p} T_{\ell k} x_k^{r(k,i)} & \text{if } \ell \in J_i. \end{cases} \tag{3}$$

The basic assumptions used to study this mathematical model are the following.

(i) $r(k,i) \le i$ for all $k = 1, \dots, p$, $i = 1, \dots$.
(ii) The set $\{i \mid \ell \in J_i\}$ is unbounded for all $\ell = 1, \dots, p$.
(iii) $\lim_{i\to\infty} r(k,i) = \infty$ for all $k = 1, \dots, p$.

The representation (3) together with (i)–(iii) can be found (sometimes with some minor variations) in most of the literature on asynchronous methods, including [1], [4], [6]–[12], [14], [19], [21], [23].

Chazan and Miranker [11] proved the following result, where, as before, we assume that T is nonnegative.

THEOREM 1. *If $\rho(T) < 1$ the asynchronous method (3) converges to the solution of (1). If $\rho(T) \ge 1$, an initial vector x^0 and a sequence $\{r(k,i)\}$, $k = 1, \dots, p$, $i = 1, \dots,$ can be constructed for which (3) does not converge.*

Bertsekas and Tsitsiklis [6], Strikwerda [19], and Su et al. [21] suggested different constructions of the non-convergent sequences for the case $\rho(T) \ge 1$.

This Theorem has been widely interpreted as meaning that the condition $\rho(T) < 1$ is necessary and sufficient for the convergence of asynchronous iterations; see, e.g., [7, p. 14].

On the other hand, Lubachevsky and Mitra [14], and more recently Pott [17], studied asynchronous methods for singular matrices, i.e., for the specific case of $\rho(T) = 1$, and gave conditions for the convergence of the asynchronous iteration (3); see also [20]. In other words, these authors show that $\rho(T) < 1$ is not a necessary condition for convergence. This of course would contradict Theorem 1.

Lubachevsky and Mitra [14] are aware of this contradiction, but dismiss it stating that their mathematical model is different from the one discussed by Chazan and Miranker [11]. I claim that this is not truly so, the two models are essentially the same.

So, what is the source of this mysterious discrepancy? A careful inspection of the construction of the non-convergent sequences in [6], [11], and [19], reveals that in all cases a solution of a linear system of the form $(I - T)v = w$ takes place, and the matrix $I - T$ is assumed to be nonsingular. This is not possible if one assumes, as we did, that $\rho(T) = 1$ (or that A is singular). In other words, the case $\rho(T) = 1$ is not covered by any of the constructions of non-converging sequences in the literature. In [21], the case $\rho(T) = 1$ is treated separately, but it should not have been included in the necessity. The authors show that, in this case, there is a sequence of vectors not converging to the zero vector, but the limit is the Perron vector to which the convergence is desired.

One can imagine that the authors of [6], [11], [19], and [21] had in mind a situation in which the system (1), or an associated nonlinear system, has a unique solution, and this is possibly why they overlooked the possibility that $I - T$ be singular.

We should then rewrite Theorem 1, stating it in a way that the contradiction is removed.

THEOREM 2. *If* $\rho(T) < 1$ *the asynchronous method* (3) *converges to the solution of* (1). *If* $\rho(T) = 1$*, under certain conditions, convergence can be achieved. If* $\rho(T) > 1$*, and if 1 is not an eigenvalue of* T*, an initial vector* x^0 *and a sequence* $\{r(k,i)\}$, $k = 1, \ldots, p$, $i = 1, \ldots,$ *can be constructed for which* (3) *does not converge.*

It is hoped that this note will help reassess the commonly, but mistakenly, held belief that the condition $\rho(T) < 1$ is necessary for the convergence of asynchronous iterations.

Acknowledgements. The author thanks Michele Benzi, Amit Bhaya, and Eugenius Kaszkurewicz, who commented on an earlier draft of this note.

REFERENCES

[1] J. Bahi, J.C. Miellou, and K. Rhofir. Asynchronous multisplitting methods for nonlinear fixed point problems. *Numerical Algorithms*, 15:315–345, 1997.

[2] George P. Barker and Robert J. Plemmons. Convergent iterations for computing stationary distributions of Markov chains. *SIAM Journal on Algebraic and Discrete Methods*, 7:390–398, 1986.

[3] Vincent A. Barker. Numerical solution of sparse singular systems of equations arising from ergodic Markov chains. *Communications in Statistics. Stochastic Models*, 5:355–381, 1989.

[4] Gérard M. Baudet. Asynchronous iterative methods for multiprocessors. *Journal of the Association for Computing Machinery*, 25:226–244, 1978.

[5] Abraham Berman and Robert J. Plemmons. *Nonnegative Matrices in the Mathematical Sciences.* Academic Press, New York, third edition, 1979. Reprinted by SIAM, Philadelphia, 1994.

[6] Dimitri P. Bertsekas and John N. Tsitsiklis. *Parallel and Distributed Computation.* Prentice Hall, Englewood Cliffs, New Jersey, 1989.

[7] Dimitri P. Bertsekas and John N. Tsitsiklis. Some aspects of parallel and distributed iterative algorithms–a survey. *Automatica*, 27:3–21, 1991.

[8] Amit Bhaya, Eugenius Kaszkurewicz, and Francisco Mota. Asynchronous block-iterative methods for almost linear equations. *Linear Algebra and its Applications*, 154–156:487–508, 1991.

[9] Rafael Bru, Ludwig Elsner, and Michael Neumann. Models of parallel chaotic iteration methods. *Linear Algebra and its Applications*, 103:175–192, 1988.

[10] Rafael Bru, Violeta Migallón, José Penadés, and Daniel B. Szyld. Parallel, synchronous and asynchronous two-stage multisplitting methods. *Electronic Transactions on Numerical Analysis*, 3:24–38, 1995.

[11] D. Chazan and W. Miranker. Chaotic relaxation. *Linear Algebra and its Applications*, 2:199–222, 1969.

[12] Andreas Frommer, Hartmut Schwandt, and Daniel B. Szyld. Asynchronous weighted additive Schwarz methods. *Electronic Transactions on Numerical Analysis*, 5:48–61, 1997.

[13] Linda Kaufman. Matrix methods for queuing problems. *SIAM Journal on Scientific and Statistical Computing*, 4:525–552, 1983.

[14] Boris Lubachevsky and Debasis Mitra. A chaotic asynchronous algorithm for computing the fixed point of a nonnegative matrix of unit spectral radius. *Journal of the Association for Computing Machinery*, 33:130–150, 1986.

[15] Ivo Marek and Daniel B. Szyld. Iterative and semi-iterative methods for computing stationary

probability vectors of Markov operators. *Mathematics of Computation*, 61:719–731, 1993.

[16] Dianne P. O'Leary. Iterative methods for finding the stationary vector for Markov chains. In Carl D. Meyer and Robert J. Plemmons, editors, *Linear Algebra, Markov Chains and Queuing Models,* IMA Volumes in Mathematics and its Applications, Vol. 48, pages 125–136, New York – Berlin, 1993. Springer Verlag.

[17] Matthias Pott. On the convergence of asynchronous iteration methods for nonlinear paracontractions and consistent linear systems. *Linear Algebra and its Applications*, 283:1–33, 1998.

[18] William J. Stewart. *Introduction to the Numerical Solution of Markov Chains.* Princeton University Press, Princeton, New Jersey, 1994.

[19] John C. Strikwerda. A convergence theorem for chaotic asynchronous relaxation. *Linear Algebra and its Applications*, 253:15–24, 1997.

[20] Yangfeng Su and Amit Bhaya. Convergence of pseudocontractions and applications to two-stage and asynchronous multisplittings. Technical Report SB97/NACAD/05, Laboratory for High Performance Computing, Graduate School of Engineering, Federal University of Rio de Janeiro, 1997.

[21] Yangfeng Su, Amit Bhaya, Eugenius Kaszkurewicz, and Victor S. Kozyakin. Further results on convergence of asynchronous linear iterations. *Linear Algebra and its Applications*, 281:11–24, 1998.

[22] Daniel B. Szyld. Different models of parallel asynchronous iterations with overlapping blocks. *Computational and Applied Mathematics*, 17:101–115, 1998.

[23] Aydin Üresin and Michel Dubois. Sufficient conditions for the convergence of asynchronous iterations. *Parallel Computing*, 10:83–92, 1989.

[24] Richard S. Varga. *Matrix Iterative Analysis.* Prentice-Hall, Englewood Cliffs, New Jersey, 1962.