A note on Newton–Noda iteration for computing the Perron pair of a weakly irreducible nonnegative tensor

Chun-Hua Guo[a]

[a] *Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan S4S 0A2, Canada*

Abstract

The Newton–Noda iteration (NNI) can be used to compute the Perron pair of a weakly irreducible nonnegative tensor. The method requires the selection of a positive parameter θ_k in the kth iteration. A practical procedure for determining the parameter θ_k is the halving procedure, starting with $\theta_k = 1$. The NNI has been shown to be globally and quadratically convergent if the sequence $\{\theta_k\}$ is bounded below by a positive constant. In this note, we prove the global convergence of NNI (with θ_k determined by the halving procedure) without assuming that $\{\theta_k\}$ is bounded below by a positive constant. We can then see that $\theta_k = 1$ for all k sufficiently large and the convergence of NNI is quadratic.

Keywords: Newton–Noda iteration, Nonnegative tensor, Perron vector, Spectral radius, Global convergence
2020 MSC: 15A18, 15A69, 65H17

1. Introduction

A real-valued mth-order n-dimensional tensor $\mathcal{A}$ consists of n^m entries in $\mathbb{R}$, and is denoted by

$$\mathcal{A} = (a_{i_1 i_2 \dots i_m}), \quad a_{i_1 i_2 \dots i_m} \in \mathbb{R}, \quad 1 \le i_1, i_2, \dots, i_m \le n.$$

The set of all such tensors is denoted by $\mathbb{R}^{[m,n]}$, and the set of all nonnegative tensors $\mathcal{A} \in \mathbb{R}^{[m,n]}$, for which $a_{i_1 i_2 \dots i_m} \ge 0$ for all $i_1, i_2, \dots, i_m$, is denoted by $\mathbb{R}_+^{[m,n]}$. We use x_i or $(\mathbf{x})_i$ to represent the ith element of a column vector $\mathbf{x}$, and use the 2-norm for vectors. For n-vectors $\mathbf{u}$ and $\mathbf{v}$, with $v_i \ne 0$ for all i, we define $\frac{\mathbf{u}}{\mathbf{v}}$ to be the n-vector whose ith component is $\frac{u_i}{v_i}$. For real vector $\mathbf{u}$, we define $\max(\mathbf{u}) = \max_i u_i$ and $\min(\mathbf{u}) = \min_i u_i$.

Definition 1. *[1] Let $\mathcal{A} \in \mathbb{R}^{[m,n]}$. We say that $(\lambda, \mathbf{x}) \in \mathbb{C} \times (\mathbb{C}^n \backslash \{0\})$ is an eigenpair (eigenvalue-eigenvector) of $\mathcal{A}$ if*

$$\mathcal{A}\mathbf{x}^{m-1} = \lambda \mathbf{x}^{[m-1]}, \tag{1}$$

Email address: chun-hua.guo@uregina.ca (Chun-Hua Guo)

where $(\mathcal{A}\mathbf{x}^{m-1})_i = \sum_{i_2,\ldots,i_m=1}^{n} a_{ii_2\ldots i_m} x_{i_2} \ldots x_{i_m}$ *for* $i = 1, \ldots, n$, *and* $\mathbf{x}^{[m-1]} = [x_1^{m-1}, x_2^{m-1}, \ldots, x_n^{m-1}]^T$.

The following Perron–Frobenius theorem for weakly irreducible nonnegative tensors is from [2]. See also [1, Lemma 3.21 and Theorem 3.26].

Theorem 1. *Let* $\mathcal{A} \in \mathbb{R}_+^{[m,n]}$ *be weakly irreducible. Then there exist* $\lambda_* > 0$ *and a unit vector* $\mathbf{x}_* > 0$ *such that* $\mathcal{A}\mathbf{x}_*^{m-1} = \lambda_* \mathbf{x}_*^{[m-1]}$. *If* λ *is an eigenvalue of* $\mathcal{A}$, *then* $|\lambda| \le \lambda_*$. *If* λ *is an eigenvalue with a positive unit eigenvector* $\mathbf{x}$, *then* $\lambda = \lambda_*$ *and* $\mathbf{x} = \mathbf{x}_*$. *Moreover, for any* $\mathbf{v} > 0$

$$\min\left(\frac{\mathcal{A}\mathbf{v}^{m-1}}{\mathbf{v}^{[m-1]}}\right) \le \lambda_* \le \max\left(\frac{\mathcal{A}\mathbf{v}^{m-1}}{\mathbf{v}^{[m-1]}}\right).$$

The eigenvalue λ_* is then the spectral radius of $\mathcal{A}$, denoted by $\rho(\mathcal{A})$. The corresponding eigenvector $\mathbf{x}_*$ is called the Perron vector of $\mathcal{A}$. The pair $(\rho(\mathcal{A}), \mathbf{x}_*)$ is called the Perron pair of $\mathcal{A}$.

The NQZ algorithm [3] is the first algorithm for computing the Perron pair for a nonnegative tensor. Its convergence is proved in [4] for primitive tensors, and linear convergence of the algorithm is proved in [2] for weakly primitive tensors. When $\mathcal{A}$ is weakly irreducible, the tensor $\mathcal{B} = \mathcal{A} + s\mathcal{I}$ is weakly primitive [1, Corollary 3.78], where $s > 0$ is any scalar and $\mathcal{I}$ is the identity tensor. Therefore, the NQZ algorithm can be applied to $\mathcal{B}$ to get the Perron pair of $\mathcal{A}$, as in [5]. The performance of the algorithm will then be dependent on the choice of s.

The Newton–Noda iteration (NNI) [6] is directly applicable to weakly irreducible nonnegative tensors, and is a generalization of the Noda iteration [7] for nonnegative matrices. The method requires the selection of a parameter θ_k in the kth iteration. A practical procedure for determining the parameter θ_k is the halving procedure, starting with $\theta_k = 1$. The NNI has been shown in [6] to be globally and quadratically convergent if the sequence $\{\theta_k\}$ is bounded below (which always means "bounded below by a positive constant" in this note). In this note, we prove the global convergence of NNI (with θ_k determined by the halving procedure) without assuming that $\{\theta_k\}$ is bounded below.

2. The Newton–Noda iteration

The NNI in [6] is reproduced here. In line 3, $\mathbf{J}_\mathbf{x}\mathbf{r}(\mathbf{x},\lambda)$ is the derivative of $\mathbf{r}(\mathbf{x}, \lambda) = \lambda\mathbf{x}^{[m-1]} - \mathcal{A}\mathbf{x}^{m-1}$ with respect to $\mathbf{x}$. It is known [6] that $\mathbf{J}_\mathbf{x}\mathbf{r}(\mathbf{x}_k, \overline{\lambda}_k)$ is an irreducible nonsingular M-matrix if $(\mathbf{x}_k, \overline{\lambda}_k) \ne (\mathbf{x}_*, \rho(\mathcal{A}))$ and $\mathbf{J}_\mathbf{x}\mathbf{r}(\mathbf{x}_*,\rho(\mathcal{A}))$ is an irreducible singular M-matrix. We assume that $\mathbf{x}_k \ne \mathbf{x}_*$ for each k. Otherwise, NNI will terminate with the exact Perron pair in a finite number of iterations. When $(\mathbf{x}_k, \overline{\lambda}_k) \ne (\mathbf{x}_*, \rho(\mathcal{A}))$, the linear system in line 3 can be solved accurately using the GTH algorithm [8].

A number of properties of NNI have been proved in [6]. Here we recall the following result.

Algorithm 1 Newton–Noda iteration (NNI)

1. Given $\mathbf{x}_0 > 0$ with $\|\mathbf{x}_0\| = 1$, $\overline{\lambda}_0 = \max\left(\frac{\mathcal{A}\mathbf{x}_0^{m-1}}{\mathbf{x}_0^{[m-1]}}\right)$, and $\mathsf{tol} > 0$.
2. **for** $k = 0, 1, 2, \ldots$
3. Solve the linear system $\mathbf{J}_{\mathbf{x}}\mathbf{r}(\mathbf{x}_k, \overline{\lambda}_k)\mathbf{w}_k = \mathbf{x}_k^{[m-1]}$.
4. Normalize the vector $\mathbf{w}_k$: $\mathbf{y}_k = \mathbf{w}_k/\|\mathbf{w}_k\|$.
5. Choose a scalar $\theta_k > 0$.
6. Compute the vector $\widetilde{\mathbf{x}}_{k+1} = (m-2)\mathbf{x}_k + \theta_k \mathbf{y}_k$.
7. Normalize the vector $\widetilde{\mathbf{x}}_{k+1}$: $\mathbf{x}_{k+1} = \widetilde{\mathbf{x}}_{k+1}/\|\widetilde{\mathbf{x}}_{k+1}\|$.
8. Compute $\overline{\lambda}_{k+1} = \max\left(\frac{\mathcal{A}\mathbf{x}_{k+1}^{m-1}}{\mathbf{x}_{k+1}^{[m-1]}}\right)$ and $\underline{\lambda}_{k+1} = \min\left(\frac{\mathcal{A}\mathbf{x}_{k+1}^{m-1}}{\mathbf{x}_{k+1}^{[m-1]}}\right)$.
9. **until** convergence: $|\overline{\lambda}_{k+1} - \underline{\lambda}_{k+1}|/\overline{\lambda}_{k+1} < \mathsf{tol}$.

Lemma 2. *[6] Let $\mathcal{A} \in \mathbb{R}_+^{[m,n]}$ be weakly irreducible and assume that the sequence $\{\overline{\lambda}_k, \mathbf{x}_k\}$ is generated by Algorithm 1, with $\{\overline{\lambda}_k\}$ bounded. Let $\{\mathbf{x}_{k_j}\}$ be any subsequence of $\{\mathbf{x}_k\}$. If $\mathbf{x}_{k_j} \to \mathbf{v}$ as $j \to \infty$, then $\mathbf{v} > 0$.*

In NNI, we would like to choose $\theta_k > 0$ such that the sequence $\{\overline{\lambda}_k\}$ is strictly decreasing (and hence bounded). Note that

$$\overline{\lambda}_k - \overline{\lambda}_{k+1} = \overline{\lambda}_k - \max\left(\frac{\mathcal{A}\widetilde{\mathbf{x}}_{k+1}^{m-1}}{\widetilde{\mathbf{x}}_{k+1}^{[m-1]}}\right) = \min\left(\frac{\mathbf{h}_k(\theta_k)}{\widetilde{\mathbf{x}}_{k+1}^{[m-1]}}\right), \tag{2}$$

where $\mathbf{h}_k(\theta) = \mathbf{r}((m-2)\mathbf{x}_k + \theta\mathbf{y}_k, \overline{\lambda}_k)$. The strategy in [6] is to find $\theta_k > 0$ such that

$$\mathbf{h}_k(\theta_k) \geq \frac{\theta_k \mathbf{x}_k^{[m-1]}}{(1+\eta)\,\|\mathbf{w}_k\|}, \tag{3}$$

for any given constant $\eta > 0$. It has been explained in [6] that $\theta_k = 1$ is desirable if condition (3) is satisfied and a smaller θ_k is used only when it is necessary for (3) to hold.

For $m = 2$, (3) holds for $\theta_k = 1$ [6]. So we assume $m \geq 3$ and let

$$\mathbf{g}_k(\theta) = \mathbf{h}_k(\theta) - \frac{\theta \mathbf{x}_k^{[m-1]}}{(1+\eta)\,\|\mathbf{w}_k\|}. \tag{4}$$

Then [6] $\mathbf{g}_k(0) = \mathbf{r}((m-2)\mathbf{x}_k, \overline{\lambda}_k) \geq 0$ and

$$\mathbf{g}_k'(0) = \left[(m-2)^{(m-2)} - \frac{1}{(1+\eta)}\right] \frac{\mathbf{x}_k^{[m-1]}}{\|\mathbf{w}_k\|} > 0. \tag{5}$$

So we can indeed choose $\theta_k \in (0, 1]$ in NNI such that (3) holds (so the sequence $\{\overline{\lambda}_k\}$ is strictly decreasing and bounded below by $\rho(\mathcal{A})$).

Global convergence of NNI is proved in [6] under the assumption that $\{\theta_k\}$ is bounded below. One way for obtaining such a sequence has been given in [6]. Indeed, one can define a sequence $\{\theta_k\}$ by

$$\theta_k = \begin{cases} 1 & \text{if } \mathbf{h}_k(1) \geq \frac{\mathbf{x}_k^{[m-1]}}{(1+\eta)\|\mathbf{w}_k\|}, \\ \eta_k & \text{otherwise,} \end{cases} \tag{6}$$

where $\eta_k = \sup\{\xi_k : \ \mathbf{g}_k'(\theta) \geq 0 \text{ on } [0, \xi_k]\}$ with $\mathbf{g}_k(\theta)$ given by (4). It is clear that $0 < \eta_k < 1$ in (6) and the condition (3) holds for $\{\theta_k\}$ defined by (6). It is shown in [6] that this sequence $\{\theta_k\}$ is bounded below. However, this sequence is very difficult to obtain and is thus not used in actual computation.

In practice, for each k we can take $\theta_k = 1$ first and check whether (3) holds. If not, then we update θ_k using $\theta_k \leftarrow \frac{1}{2}\theta_k$ and check again, until we get a θ_k for which (3) holds. This process is the so-called halving procedure proposed in [6]. Of course, we could also have used the update $\theta_k \leftarrow r\theta_k$ for any $r \in (0, 1)$. The analysis for this more general strategy is similar to that for the special case $r = \frac{1}{2}$ and using $r \neq \frac{1}{2}$ does not yield any particular advantage. So we will stick with the halving procedure.

It is important to recognize that we cannot conclude that the sequence $\{\theta_k\}$ from the halving procedure is bounded below, from the fact that the sequence $\{\theta_k\}$ defined by (6) is bounded below. Whether the sequence $\{\theta_k\}$ from the halving procedure is always bounded below was left as an open problem in [6]. The global convergence of NNI with a practical procedure for choosing θ_k has been proved only for $m = 3$ in [9], where θ_k is determined by a different practical procedure.

In the next section, we prove the global convergence of the NNI (with θ_k determined by the halving procedure) without first proving that θ_k is bounded below. After the global convergence is proved, we can then conclude that θ_k is actually bounded below.

3. Global convergence

We start with a k-dependent lower bound for θ_k.

Lemma 3. *Let θ_k in NNI be determined by the halving procedure. Then $\theta_k = 1$ or $\frac{1}{2} \geq \theta_k > \xi \frac{\min \mathbf{x}_k^{[m-1]}}{\|\mathbf{w}_k\|}$ for a positive constant ξ.*

Proof. For each $k \geq 0$, we have $\|\mathbf{x}_k\| = \|\mathbf{y}_k\| = 1$ and $\overline{\lambda}_k \in [\rho(\mathcal{A}), \overline{\lambda}_0]$. Then, for all $\theta \in [0, 1]$, we have by Taylor's theorem that

$$\mathbf{g}_k(\theta) \geq \mathbf{g}_k(0) + \mathbf{g}_k'(0)\theta - M\theta^2\mathbf{e} \tag{7}$$

for some constant $M > 0$ (independent of k and θ), where $\mathbf{e} = [1, 1, \ldots, 1]^T$.

Recall that $\mathbf{g}_k(0) \geq 0$ and $\mathbf{g}_k'(0)$ is given by (5). If $M \leq \min \mathbf{g}_k'(0)$, then $\mathbf{g}_k(1) \geq 0$ by (7). In this case, the halving procedure always returns $\theta_k = 1$. If $M > \min \mathbf{g}_k'(0)$, then we see from (7) that $\mathbf{g}_k(\theta) \geq 0$ when $\theta \leq \frac{1}{M} \min \mathbf{g}_k'(0)$.

In this case, the halving procedure may still return $\theta_k = 1$ and we always have $\theta_k > \frac{1}{2M} \min \mathbf{g}_k'(0) = \xi \frac{\min \mathbf{x}_k^{[m-1]}}{\|\mathbf{w}_k\|}$, where $\xi = \frac{1}{2M}\left[(m-2)^{(m-2)} - \frac{1}{(1+\eta)}\right]$. □

Even though the above lower bound for θ_k is k-dependent, we can still use it to show that the sequence $\left\{\overline{\lambda}_k, \mathbf{x}_k, \mathbf{y}_k\right\}$ generated by Algorithm 1 always converges to $(\rho(\mathcal{A}), \mathbf{x}_*, \mathbf{x}_*)$. We then know that $\theta_k = 1$ for k sufficiently large. This proves in a roundabout way that $\{\theta_k\}$ is bounded below. Our previous efforts aim to show $\{\theta_k\}$ is bounded below *before* we set out to show the convergence of $\left\{\overline{\lambda}_k, \mathbf{x}_k\right\}$. Those efforts all failed unfortunately.

Theorem 4. *Let $\mathcal{A} \in \mathbb{R}_+^{[m,n]}$ be weakly irreducible and assume that the sequence $\{\theta_k\}$ in Algorithm 1 is determined by the halving procedure. Then the monotonically decreasing sequence $\left\{\overline{\lambda}_k\right\}$ converges to $\rho(\mathcal{A})$, and $\{\mathbf{x}_k\}$ from Algorithm 1 converges to $\mathbf{x}_*$. Moreover, $\{\mathbf{y}_k\}$ from Algorithm 1 converges to $\mathbf{x}_*$ as well.*

Proof. By (2) and (3) we have

$$\begin{aligned} \overline{\lambda}_k - \overline{\lambda}_{k+1} &= \min\left(\frac{\mathbf{h}_k(\theta_k)}{\widetilde{\mathbf{x}}_{k+1}^{[m-1]}}\right) \geq \min\left(\frac{\theta_k \mathbf{x}_k^{[m-1]}}{(1+\eta)\|\mathbf{w}_k\|\widetilde{\mathbf{x}}_{k+1}^{[m-1]}}\right) \\ &\geq \frac{1}{(1+\eta)(m-1)^{m-1}} \theta_k \frac{\min \mathbf{x}_k^{[m-1]}}{\|\mathbf{w}_k\|}, \end{aligned}$$

since $\|\widetilde{\mathbf{x}}_{k+1}\| = \|(m-2)\mathbf{x}_k + \theta_k \mathbf{y}_k\| \leq m-1$ and $\left(\widetilde{\mathbf{x}}_{k+1}^{[m-1]}\right)_i \leq \|\widetilde{\mathbf{x}}_{k+1}\|^{m-1}$ for each i. Let $\gamma = (1+\eta)(m-1)^{m-1}$.

By Lemma 3 we have $\theta_k = 1$ or $\frac{1}{2} \geq \theta_k \geq \xi \frac{\min \mathbf{x}_k^{[m-1]}}{\|\mathbf{w}_k\|}$. In the first case, we have $\frac{\min \mathbf{x}_k^{[m-1]}}{\|\mathbf{w}_k\|} \leq \gamma(\overline{\lambda}_k - \overline{\lambda}_{k+1})$. In the second case, we have

$$\overline{\lambda}_k - \overline{\lambda}_{k+1} \geq \frac{\xi}{\gamma}\left(\frac{\min \mathbf{x}_k^{[m-1]}}{\|\mathbf{w}_k\|}\right)^2.$$

Therefore, for all $k \geq 0$,

$$\frac{\min \mathbf{x}_k^{[m-1]}}{\|\mathbf{w}_k\|} \leq \max\left\{\gamma(\overline{\lambda}_k - \overline{\lambda}_{k+1}), \left((\gamma/\xi)(\overline{\lambda}_k - \overline{\lambda}_{k+1})\right)^{1/2}\right\}. \tag{8}$$

Since the sequence $\left\{\overline{\lambda}_k\right\}$ converges, we obtain from (8) that

$$\lim_{k\to\infty} \|\mathbf{w}_k\|^{-1} \min(\mathbf{x}_k^{[m-1]}) = 0.$$

As in the proof of [6, Theorem 3], Lemma 2 implies that $\min(\mathbf{x}_k^{[m-1]})$ is bounded below by a positive constant. Therefore, $\lim_{k\to\infty} \|\mathbf{w}_k\|^{-1} = 0$. Then, as in the proof of [6, Theorem 3], we have $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}_*$, $\lim_{k\to\infty} \mathbf{y}_k = \mathbf{x}_*$, and $\lim_{k\to\infty} \overline{\lambda}_k = \rho(\mathcal{A})$. □

From the proof, we see that $\lim_{k\to\infty} \|\mathbf{w}_k\|^{-1} = 0$. Therefore, the inequality $\theta_k \geq \xi \frac{\min \mathbf{x}_k^{[m-1]}}{\|\mathbf{w}_k\|}$ by itself does not show $\{\theta_k\}$ is bounded below (by a positive constant). However, after Theorem 4 has been proved, we know from [6] that $\theta_k = 1$ for all k sufficiently large. This means that $\{\theta_k\}$ is indeed bounded below. We also know from [6] that the convergence of NNI is quadratic.

Our approach here can also be used to improve a convergence result in [10], where a generalized $\mathcal{M}$-tensor pair $(\mathcal{A}, \mathcal{B})$ is considered. It is shown in [10] that there exists a unique positive eigenpair (s, x_*) for the tensor pair $(\mathcal{A}, \mathcal{B} - \mathcal{A})$, up to a multiplicative constant for x_*. This eigenpair can be computed by the generalized Newton–Noda iteration (GNNI) in [10]. The GNNI is analyzed in [10] in the same way as NNI is analyzed in [6]. In particular, the parameters θ_k in GNNI can also be determined by a halving procedure. The global convergence of GNNI is proved in [10, Theorem 4.5] under the assumption that $\{\theta_k\}$ is bounded below, as in [6]. With the approach in this note, that assumption can also be removed.

4. Conclusion

We have re-examined NNI with the practical halving procedure for choosing the parameters θ_k, for computing the Perron pair of a weakly irreducible nonnegative tensor. We have found a simple way to prove the global convergence of NNI without any further assumptions. We then know that $\theta_k = 1$ for all k sufficiently large and the convergence of NNI is quadratic. The convergence theory of NNI is now complete.

Acknowledgments

This research was supported in part by an NSERC Discovery Grant.

Declarations of interest: none.

References

[1] L. Qi, Z. Luo, Tensor Analysis: Spectral Theory and Special Tensors, SIAM, Philadephia, 2017.

[2] S. Friedland, S. Gaubert, L. Han, Perron–Frobenius theorem for nonnegative multilinear forms and extensions, Linear Algebra Appl. 438 (2013) 738–749.

[3] M. Ng, L. Qi, G. Zhou, Finding the largest eigenvalue of a nonnegative tensor, SIAM J. Matrix Anal. Appl. 31 (2009) 1090–1099.

[4] K.-C. Chang, K. J. Pearson, T. Zhang, Primitivity, the convergence of the NQZ method, and the largest eigenvalue for nonnegative tensors, SIAM J. Matrix Anal. Appl. 32 (2011) 806–819.

[5] Y. Liu, G. Zhou, N. F. Ibrahim, An always convergent algorithm for the largest eigenvalue of an irreducible nonnegative tensor, J. Comput. Appl. Math. 235 (2010) 286–292.

[6] C.-S. Liu, C.-H. Guo, W.-W. Lin, Newton–Noda iteration for finding the Perron pair of a weakly irreducible nonnegative tensor, Numer. Math. 137 (2017) 63–90.

[7] T. Noda, Note on the computation of the maximal eigenvalue of a non-negative irreducible matrix, Numer. Math. 17 (1971) 382–386.

[8] W. K. Grassmann, M. J. Taksar, D. P. Heyman, Regenerative analysis and steady-state distributions for Markov chains, Oper. Res. 33 (1985) 1107–1116.

[9] C.-S. Liu, C.-H. Guo, W.-W. Lin, A positivity preserving inverse iteration for finding the Perron pair of an irreducible nonnegative third order tensor, SIAM J. Matrix Anal. Appl. 37 (2016) 911–932.

[10] W. Ma, W. Ding, Y. Wei, Noda iteration for computing generalized tensor eigenpairs, J. Comput. Appl. Math. 432 (2023), Paper No. 115284, 25 pp.