Jurnal
Teknologi

Full Paper

The Rate of Convergence and Weaker Convergent Condition for the Method for Finding the Largest Singular Value of Rectangular Tensors

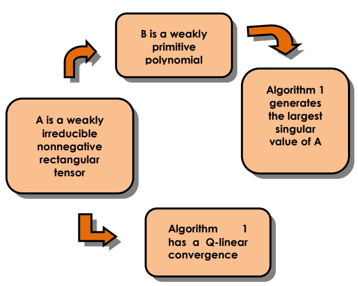
Nur Fadhilah Ibrahim[a]*, Nurul Akmal Mohamed[b]

[a]School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu, 21030 Kuala Terengganu, Malaysia
[b]Mathematics Department, Faculty of Science and Mathematics, 35900 Universiti Pendidikan Sultan Idris, Proton City, Tanjung Malim, Perak, Malaysia

Article history
Received
12 August 2015
Received in revised form
14 December 2015
Accepted
11 January 2016

*Corresponding author
fadhilah@umt.edu.my

Graphical abstract



Abstract

The applications of real rectangular tensors, among others, including the strong ellipticity condition problem within solid mechanics, and the entanglement problem within quantum physics. A method was suggested by Zhou, Caccetta and Qi in 2013, as a means of calculating the largest singular value of a nonnegative rectangular tensor. In this paper, we show that the method converges under weak irreducibility condition, and that it has a Q-linear convergence.

Keywords: Rectangular tensor, iterative method, singular value, convergence

1.0 INTRODUCTION

Tensors can be considered a generalization of matrices. They are represented as a multidimensional array of numbers. The application of real rectangular tensors, amongst others, include part of the strong ellipticity condition problem within solid mechanics and entanglement problem within quantum physics [1-6].

Most properties of tensors are generalized from matrices. We have seen in these past few years that the study of the spectral radius for tensors has developed a great interest. Chang, Qi and Zhou [7] introduced the class of the real rectangular tensor, and have presented a method for calculating the largest singular value of a nonnegative rectangular tensor. This method was originally used in order to find the largest eigenvalue of a nonnegative matrix [8, 9]. Later the method was extended by Ng, Qi and Zhou [10] for square tensors and most recently for rectangular tensors [7]. Zhou, Caccetta and Qi [11] improved the method in [7] and have shown that the algorithm converges for irreducible nonnegative rectangular tensors.

The largest singular value problems can also be solved through the use of metaheuristics optimization algorithms. Some of the novel recent works within the field of metaheuristics include the Enhanced Leader PSO [12,13], the Brainstorm Optimization Algorithm [14], the Chaotic-Based Big Bang-Big Crunch Algorithm [15], and the Chaotic Bat Swarm Optimization [16].

However the convergence of the method presented in [11] is limited to only irreducible nonnegative rectangular tensors. It is unknown whether the method is also convergence for a wider class of tensors.

There are two objectives of this paper. The first objective is to prove that the method presented in the study of Zhou *et al*. [11] for finding the largest singular value is convergent when the tensor is a weakly irreducible nonnegative rectangular tensor, and the second is to prove that the rate of convergence for the method in the study [11] is Q-linear, when the tensor is a weakly irreducible nonnegative rectangular tensor.

In this paper, we introduce the class of the weakly irreducible rectangular tensor, which is a wider class than the irreducible rectangular tensor. We then establish the convergence of the method in the study of Zhou *et al*. [11] for weakly irreducible nonnegative rectangular tensors, and its rate of convergence. This research intends to contribute to the convergence properties for the method of finding the largest singular value of rectangular tensors.

In Section 2, we provide definitions and theorems to be used later. In Section 3, it is proved that the algorithm given in [11] converges for weakly irreducible nonnegative rectangular tensors. In Section 4, it is proven that the algorithm in [11] has Q-linear convergence, and lastly in Section 5 this paper is concluded.

2.0 PRELIMINERIES

Let R be the real field, let R_+ be the set of nonnegative numbers, and let $R_{>0}$ be the set of positive numbers. Let $p,q,m,n\in R_{>0}$ and $m,n\geq 2$. We can say A is a real p,q -th order $m\times n$ dimensional rectangular tensor, where:

$$A=\left(A_{i_1\dots i_p j_1\dots j_q}\right), A_{i_1\dots i_p j_1\dots j_q}\in R, \quad 1\leq i_1,\dots,i_p\leq m,\ 1\leq j_1,\dots,j_q\leq n \tag{1}$$

We call A a nonnegative rectangular tensor if $a_{i_1\dots i_p j_1\dots j_q}\geq 0$. When $m=n,$ A is a square tensor and when $p=1,\ q=1$, matrix A is an $m\times n$ rectangular matrix.

The singular value of rectangular tensors is comparable to the eigenvalue of square tensors. Here we use the following definition for the singular value of rectangular tensors. See e.g. [7]. Let $\left(Ax^{p-1}y^q\right)\in R^m$, where:

$$\left(Ax^{p-1}y^q\right)_i=\sum_{i_2,\dots,i_p=1}^{m}\sum_{j_1,\dots,j_p=1}^{n}A_{ii_2\dots i_p j_1\dots j_q}x_{i_2}\dots x_{i_p}y_{j_1\dots j_q},$$

$i=1,\dots,m$, and let $\left(Ax^{p}y^{q-1}\right)\in R^n$, where

$$\left(Ax^{p}y^{q-1}\right)_j=\sum_{i_1,\dots,i_p=1}^{m}\sum_{j_2,\dots,j_p=1}^{n}A_{i_1\dots i_p jj_2\dots j_q}x_{i_1}\dots x_{i_p}y_{j_2\dots j_q},$$

$j=1,\dots,n.$ We set $M=p+q$ and $N=m+n.$ Let

$$Ax^{p-1}y^q=\lambda x^{[M-1]},\ Ax^p y^{q-1}=\lambda y^{[M-1]}. \tag{2}$$

We call $\lambda\in C$ a singular value of $A,$ where C is the set of complex numbers. We can say $x\in C^m\setminus\{0\}$ and $y\in C^n\setminus\{0\}$ are left and right eigenvectors of $A,$ paired with the singular value $\lambda,$ if λ,x and y satisfy the equation (2). The following are some preliminaries:

Theorem 1 (p.20,[17]). An $n\times n$ complex matrix A is irreducible if and only if its directed graph $G(A)$ is strongly connected.

Theorem 2 (p.51,[17]). Let A be an irreducible matrix, with $G(A)$ being the associated directed graph. If the greatest common divisor (gcd) of the lengths of its closed paths is equal to one, then A is primitive.

The converse of Theorem 2 also holds.

Theorem 3 (Chapter 2,[17]). If A is an irreducible nonnegative square matrix, then:

(i) the spectral radius $\rho(A)$ is an eigenvalue;

(ii) there exists a nonnegative vector $x_0>0$, such that $Ax_0=\rho(A)x_0$;

(iii) (uniqueness) if λ is an eigenvalue with a nonnegative eigenvector, then $\lambda=\rho(A)$;

(iv) $\rho(A)$ is a simple eigenvalue of A;

(v) if λ is an eigenvalue of A, then $|\lambda|\leq\rho(A)$.

Furthermore, if a nonnegative matrix A is primitive, then $\rho(A)>|\lambda|,\forall\lambda\in\sigma(A)\setminus\{\rho(A)\},$ where $\sigma(A)$ is the spectrum of A.

Corollary 1 [18]. An irreducible matrix with a nonzero main diagonal is primitive.

Proposition 1 [19,20]. The spectral radius of an $n\times n$ matrix A is characterized by the equality

$$\rho(A)=\inf_{\|\bullet\|\in N}\|A\|$$

where N denotes the set of all possible spectral norms of A. For any $\epsilon>0$, there exists a spectral norm $\|\bullet\|\in N$ such that $\|A\|\leq\rho(A)+\epsilon.$

For any $j=1,2,\dots,n,$ let $A_{\bullet j}=\left(A_{i_1\dots i_p j\dots j}\right)$ be a p-th order, m dimensional square tensor. For any $i=1,2,\dots,m,$ let $A_{i\bullet}=\left(A_{i\dots ij_1\dots j_q}\right)$ be a q-th order n dimensional square tensor. In this paper we consider all polynomials to be monotone and homogeneous.

Definition 1 [7, 11]. A nonnegative rectangular tensor A is irreducible if all the square tensors $A_{\bullet j}=\left(A_{i_1\dots i_p j\dots j}\right),$ $j=1,2,\dots,n,$ and $A_{i\bullet}=\left(A_{i\dots ij_1\dots j_q}\right),$ $i=1,2,\dots,m,$ are irreducible.

Let A be a nonnegative $p,q-$th order $m\times n$ dimensional rectangular tensor. The graph $G(A)=(V,E(A))$ is the associated graph of tensor A. The vertex set is $V=\bigcup_{j=1}^{p}V_j+\bigcup_{j=p+1}^{M}V_j$, with $V_j=\{1,2,\dots,m\}$ for $j=1,2,\dots,p$ and $V_j=\{1,2,\dots,n\}$ for

$j=p+1,\ldots,M$, $M=p+q$. An edge $(i_k,i_l)\in V_k\times V_l$ exists if and only if $A_{i_1\ldots i_p j_1\ldots j_q}>0$ for some $M-2$ indices $\{i_1,\ldots,i_p,j_1,\ldots,j_q\}\setminus\{i_k,i_l\}$. The tensor A is considered weakly irreducible, if the graph $G(A)$ is connected [21].

For a rectangular tensor A, let $\rho>0$, $x\in R_+^m$, $y\in R_+^n$ and:

$$B_x(x,y)=Ax^{p-1}y^q+\rho x^{[M-1]}, \tag{3}$$

$$B_y(x,y)=Ax^p y^{q-1}+\rho y^{[M-1]}. \tag{4}$$

The following Theorem was given in the study of Zhou *et al*. [11].

Theorem 4 [11]. If A is an irreducible nonnegative rectangular tensor of the order p,q and the dimension $m\times n$, then there exists $\mu_0>0$, $x_0\in R_{>0}^m$ and $y_0\in R_{>0}^n$, such that:

$$B_x(x_0,y_0)=\mu_0 x_0^{[M-1]},\ B_y(x_0,y_0)=\mu_0 y_0^{[M-1]}. \tag{5}$$

Moreover, μ_0 satisfies the following equalities:

$$\mu_0=\min_{(x,y)\in(R_+^m\setminus\{0\})\times(R_+^n\setminus\{0\})}\max_{i,j}\left(\frac{B_x(x,y)_i}{x_i^{[M-1]}},\frac{B_y(x,y)_j}{y_j^{[M-1]}}\right)$$

$$=\max_{(x,y)\in(R_+^m\setminus\{0\})\times(R_+^n\setminus\{0\})}\min_{i,j}\left(\frac{B_x(x,y)_i}{x_i^{[M-1]}},\frac{B_y(x,y)_j}{y_j^{[M-1]}}\right)$$

and $\mu_0-\rho$ is the largest singular value of the rectangular tensor A.

3.0 WEAKER CONVERGENT CONDITION

An algorithm for finding the largest singular value of an irreducible nonnegative rectangular tensor was proposed by Chang *et al*. [7]. Later, it was updated by Zhou *et al*. [11]. In this section, we will prove that the algorithm is convergent for weakly irreducible nonnegative rectangular tensors.

Algorithm 1 [11]

Step 0: Choose $\rho>0, x^{(1)}>0$ and $y^{(1)}>0$. Set $k=1$.

Step 1: Calculate $\xi^{(k)}=B_x\left(x^{(k)},y^{(k)}\right)$ and

$\eta^{(k)}=B_y\left(x^{(k)},y^{(k)}\right)$. Let

$$\underline{\mu}_k=\min_{x_i^{(k)}>0,y_j^{(k)}>0}\left(\frac{\xi_i^{(k)}}{\left(x_i^{(k)}\right)^{M-1}},\frac{\eta_j^{(k)}}{\left(y_j^{(k)}\right)^{M-1}}\right),$$

$$\bar{\mu}_k=\max_{x_i^{(k)}>0,y_j^{(k)}>0}\left(\frac{\xi_i^{(k)}}{\left(x_i^{(k)}\right)^{M-1}},\frac{\eta_j^{(k)}}{\left(y_j^{(k)}\right)^{M-1}}\right).$$

Step 2: If $\underline{\mu}_k=\bar{\mu}_k$, then stop. Otherwise, compute

$$x^{(k+1)}=\frac{\left(\xi^{(k)}\right)^{\left[\frac{1}{M-1}\right]}}{\left\|\left(\xi^{(k)},\eta^{(k)}\right)^{\left[\frac{1}{M-1}\right]}\right\|}\quad\text{and}\quad y^{(k+1)}=\frac{\left(\eta^{(k)}\right)^{\left[\frac{1}{M-1}\right]}}{\left\|\left(\xi^{(k)},\eta^{(k)}\right)^{\left[\frac{1}{M-1}\right]}\right\|},$$

replace k with $k+1$ and go to Step 1.

Let $\mu_0=\underline{\mu}_k=\bar{\mu}_k$. The largest singular value of A is $\mu_0-\rho$. Zhou *et al*. [11] have shown that this algorithm is convergent, if A is an irreducible nonnegative rectangular tensor. We will now show that Algorithm 1 is convergent if A is a weakly irreducible nonnegative rectangular tensor.

We define the polynomial map $P=(P_1,\ldots,P_N)^T:R_+^N\to R_+^N$ through:

$$P(z)=\begin{pmatrix}Ax^{p-1}y^q\\ Ax^p y^{q-1}\end{pmatrix},$$

where $N=m+n$, $z=\begin{pmatrix}x\\ y\end{pmatrix}$. Let P_i be a polynomial with degree, $d_i>1$. Suppose that the coefficient of each monomial in P_i is nonnegative. The associated graph of P is the directed graph $G(P)=(V,E(P))$, where the vertices $V=\{1,2,\ldots,N\}$ and the edge $(i,j)\in E(P)$ if the coefficient of variable z_j appears in the expression of P_i.

Definition 2: Let $P=(P_1,\ldots,P_N)^T:R_+^N\to R_+^N$ be a polynomial map, where each P_i is a homogeneous polynomial of the degree $d\geq 1$ with nonnegative coefficients. We call P weakly irreducible if $G(P)$ is strongly connected. If the directed graph $G(P)$ is strongly connected, and the great common divisor (gcd) of the lengths of its circuits is equal to one, then we say P is weakly primitive.

Another way to check the gcd of a graph's lengths of is to observe the diagonal of its associated matrix. An irreducible matrix has a nonzero main diagonal entry if and only if the associated directed graph has a loop, a closed path with length equals to one.

We can show that P is weakly primitive by proving that the associated matrix of its graph is primitive. Let $M(G(P))$ be the associated matrix of graph $G(P)$. We can say that $M(G(P))$ is primitive if the graph is strongly connected, and if the gcd of its lengths is equal to one.

Definition 3: A rectangular tensor A is weakly irreducible if P is weakly irreducible.

Let $B(z)=\begin{pmatrix} Ax^{p-1}y^{q}+\rho x^{[M-1]} \\ Ax^{p}y^{q-1}+\rho y^{[M-1]} \end{pmatrix}$ and let $I(z)=\begin{pmatrix} \rho x^{[M-1]} \\ \rho y^{[M-1]} \end{pmatrix}$.

Hence we have $B(z)=P(z)+I(z)$. Now we prove that Algorithm 1 is convergent, if tensor A is weakly irreducible.

We can now present our results for this section.

Lemma 1: If A is a weakly irreducible nonnegative rectangular tensor with the order p,q and the $m\times n$ dimension, then $B(z)$ is a weakly primitive polynomial.

Proof. Since A is weakly irreducible then $P(z)$ is a weakly irreducible polynomial. By Definition 2, the graph of $P(z)$, $G(P(z))$ is strongly connected. By Theorem 1, the matrix of $G(P(z))$ is irreducible. We know that $G(I(z))$, the graph of $I(z)$, has a self-loop at each vertices. Therefore the matrix of $G(I(z))$ is a diagonal matrix. Hence, by Corollary 1, the matrix of $G(B(z))$ is primitive. By Theorem 2, $G(B(z))$ is strongly connected, and has a gcd that is equal to one. This implies by Definition 2 that $B(z)$ is a weakly primitive polynomial.

The following theorem is the main result of this paper.

Theorem 6. Let A be a weakly irreducible rectangular tensor of the p,q-th order and the $m\times n$ dimension. Suppose that (μ_0,x_0,y_0) is the solution of equation (5). Then, Algorithm 1 yields the value of μ_0 through a finite number of steps, or generate two convergent sequences $\{\underline{\mu}_k\}$ and $\{\bar{\mu}_k\}$, both of which converge to μ_0. The largest singular value of A is $\mu_0-\rho$.

Proof. By Lemma 1 and Corollary 5.1 [21], Algorithm 1 converges when the rectangular tensor A is weakly irreducible.

4.0 RATE OF CONVERGENCE

In this section, we will show that Algorithm 1 has Q-linear convergence, when A is a nonnegative weakly irreducible rectangular tensor of p,q-th order and $m\times n$ dimensional. We use the same argument as Zhou, Qi and Wu's study [22].

Define:

$$F(z)=B(z)=\begin{pmatrix} Ax^{p-1}y^{q}+\rho x^{[M-1]} \\ Ax^{p}y^{q-1}+\rho y^{[M-1]} \end{pmatrix},$$

$$D(z)=F(z)^{\left[\frac{1}{M-1}\right]}, \qquad H(z)=\frac{D(z)}{\phi(D(z))},$$

where $\phi:R_+^N\to R_+$ is defined as:

$$\phi(z)=z_1=\sum_{i=1}^{N}z_i,$$

for any nonnegative $z\in R_+^N$. We can see that the sequence $\{z^{(k)}\}$ in Algorithm 1 is generated by

$$z^{(k+1)}=H\left(z^{(k)}\right),\quad k=1,2,\ldots, \tag{6}$$

and $\phi\left(z^{(k)}\right)=1$ for all $k=1,2,\ldots$.

Lemma 2. Let A,μ_0,x_0 and y_0 be as in Theorem 6 and let $H'(z_0)$ be the Jacobian of the function H at z_0. Then, $\rho(H'(z_0))<1$.

Proof. Let μ_0 be the largest singular value of B and z_0 be the corresponding eigenvector. We have $H(z_0)=D(z_0)/\phi(D(z_0))$. We want to show that:

$$\rho(H'(z_0))=\rho\left(\frac{D'(z_0)\phi(D(z_0))-D(z_0)\phi'(D(z_0))}{\phi^2(D(z_0))}\right)<1.$$

We already have $F(z_0)=B(z_0)=\mu_0 z_0^{[M-1]}$ and $\phi(z_0)=1$.

Hence, $D(z_0)=(F(z_0))^{\left[\frac{1}{M-1}\right]}=\mu_0^{\left[\frac{1}{M-1}\right]}z_0$. Let $\mu_1=\mu_0^{\left[\frac{1}{M-1}\right]}$, so we have $D(z_0)=\mu_1 z_0$.

Now we compute $D'(z_0)$, i.e. the Jacobian of D at z_0. Let

$$D(z_0)=(F(z_0))^{\left[\frac{1}{M-1}\right]}=\begin{bmatrix} (F_1(z_0))^{\left[\frac{1}{M-1}\right]} \\ (F_2(z_0))^{\left[\frac{1}{M-1}\right]} \\ \vdots \\ (F_N(z_0))^{\left[\frac{1}{M-1}\right]} \end{bmatrix},$$

$$\nabla\left((F_1(z_0))^{\left[\frac{1}{M-1}\right]}\right)=\frac{1}{M-1}(F_1(z_0))^{\left[\frac{2-M}{M-1}\right]}\nabla F_1(z_0).$$

By the same method, we can get:

$$\nabla\left((F_i(z_0))^{\left[\frac{1}{M-1}\right]}\right)=\frac{1}{M-1}(F_i(z_0))^{\left[\frac{2-M}{M-1}\right]}\nabla F_i(z_0)\ ,\ i=1,\ldots,N.$$

Thus the Jacobian of D at z_0 is given below:

$$D'(z_0)=\nabla\left((F(z_0))^{\left[\frac{1}{M-1}\right]}\right)$$

$$=\begin{bmatrix} \nabla\left((F_1(z_0))^{\left[\frac{1}{M-1}\right]}\right) \\ \nabla\left((F_2(z_0))^{\left[\frac{1}{M-1}\right]}\right) \\ \vdots \\ \nabla\left((F_N(z_0))^{\left[\frac{1}{M-1}\right]}\right) \end{bmatrix}=\begin{bmatrix} \frac{1}{M-1}(F_1(z_0))^{\left[\frac{2-M}{M-1}\right]}\nabla F_1(z_0) \\ \frac{1}{M-1}(F_2(z_0))^{\left[\frac{2-M}{M-1}\right]}\nabla F_2(z_0) \\ \vdots \\ \frac{1}{M-1}(F_N(z_0))^{\left[\frac{2-M}{M-1}\right]}\nabla F_N(z_0) \end{bmatrix}$$

$$= \begin{bmatrix} \frac{1}{M-1}(F_1(z_0))^{\left[\frac{2-M}{M-1}\right]} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{M-1}(F_N(z_0))^{\left[\frac{2-M}{M-1}\right]} \end{bmatrix} \begin{bmatrix} \nabla F_1(z_0) \\ \nabla F_2(z_0) \\ \vdots \\ \nabla F_N(z_0) \end{bmatrix}$$

$$= diag\left(\frac{1}{M-1}(F(z_0))^{\left[\frac{2-M}{M-1}\right]}\right)F'(z_0)$$

$$= \frac{1}{M-1} diag\left((\mu_1 z_0)^{[2-M]}\right)F'(z_0),$$

where $\frac{1}{M-1} diag\left((\mu_1 z_0)^{[2-M]}\right)$ is a constant with $\mu_1 > 0$, and z_0 is a positive vector. Therefore $G(D'(z_0)) = G(F'(z_0))$. For the graph of B, by definition, there exists an edge between i and j, if variable z_j appears in the expression of B_i. Notice that the graph of B is similar to the graph of D', $G(D'(z_0)) = G(F'(z_0)) = G(B(z_0))$. Lemma 1 states that B is weakly primitive, therefore, the graph of B is strongly connected. Hence the graph of D' is also strongly connected, and D' is therefore irreducible. The term $I(z)$ in B ensures that the diagonal is nonzero, and that implies D' be primitive matrix. Since $D'(z_0)$ is a primitive matrix, by Theorem 3, the eigenvalues $v_1, v_2, \ldots, v_N$ of $D'(z_0)$ can be ordered as follows:

$$v_1 = \rho(D'(z_0)) > |v_2| \geq |v_3| \geq \ldots \geq |v_N|.$$

For all $t > 1$, we expand $D(tz_0)$ about z_0 by using Taylor's Series, and obtains:

$$\begin{aligned} t\mu_1 z_0 &= D(tz_0) \\ &= D(z_0) + D'(z_0)(tz_0 - z_0) + o(\|tz_0 - z_0\|) \\ &= \mu_1 z_0 + (t-1)D'(z_0)z_0 + o(t-1) \\ (t-1)\mu_1 z_0 &= (t-1)D'(z_0)z_0 + o(t-1), \end{aligned}$$

which implies that $D'(z_0)z_0 = \mu_1 z_0$. Since $D'(z_0)$ is a primitive matrix, and $z_0 > 0$, by referring to the Theorem 3, z_0 is an eigenvector of $D'(z_0)$ associated with the largest eigenvalue $\mu_1 = v_1$. Therefore, $\phi(D(z_0)) = \phi(\mu_1 z_0) = \mu_1$.

We also have $\phi(D(z_0)) = D_1(z_0) + D_2(z_0) + \ldots + D_N(z_0)$, and $\phi'(D(z_0)) = D_1'(z_0) + D_2'(z_0) + \ldots + D_N'(z_0) = eD'(z_0)$, where e is the row vector of ones with N dimension. From $H(z_0) = D(z_0)/\phi(D(z_0))$, and after some manipulations we attain:

$$\begin{aligned} H'(z_0) &= \frac{D'(z_0)\phi(D(z_0)) - D(z_0)\phi'(D(z_0))}{\phi^2(D(z_0))} \\ &= \frac{D'(z_0) - z_0 eD'(z_0)}{\mu_1}. \end{aligned}$$

Let $S = D'(z_0)$ and $Q = S - z_0 eS$. Therefore the above equation can be written as $H'(z_0) = Q/\mu_1$. Here let it be reminded that we want to prove that $\rho(H'(z_0)) = \rho(Q/\mu_1) < 1$. We can achieve this by showing that the spectral radius of Q is equal to $|v_2|$. We can also show that the spectrum of Q is $\{0, v_2, v_3, \ldots, v_N\}$.

We have $1 = \phi(z_0) = (z_0)_1 + (z_0)_2 + \ldots + (z_0)_N = ez_0$, so $ez_0 = 1$ and $Q = S - z_0 eS$, $Q^T e^T = (S - z_0 eS)^T e^T = 0.$ We can conclude that e^T is an eigenvector of Q^T, associated with the eigenvalue 0.

There are two possible cases of S^T.

Case 1: The matrix $S^T = D'(z_0)^T$ is diagonizable, that is, S^T is semisimple. For $i = 2,3,\ldots,N$, we assume $S^T w^i = v_i w^i$, where w^i is an eigenvector of S^T that is associated with the eigenvalue v_i. Suppose that the set of eigenvector $\{w^1, w^2, \ldots, w^N\}$ is linearly independent.

We can write $v_i z_0^T w^i = z_0^T v_i w^i = z_0^T S^T w^i$, for $i = 2,3,\ldots,N$. We already have $D'(z_0)z_0 = Sz_0 = \mu_1 z_0$. So, $(Sz_0)^T = (\mu_1 z_0)^T$, and

$$z_0^T S^T = \mu_1 z_0^T. \tag{7}$$

Hence, $v_i z_0^T w^i = z_0^T S^T w^i = \mu_1 z_0^T w^i$, $(v_i - \mu_1) z_0^T w^i = 0$. So, for $i = 2,3,\ldots,N$, it is either $v_i = \mu_1$ or $z_0^T w^i = 0$. However $v_i \neq \mu_1$ for $i = 2,3,\ldots,N$. Therefore $z_0^T w^i = 0$.

Now we have $Q^T w^i = (S - z_0 eS)^T w^i = S^T w^i - 0$. Since we assume that $S^T w^i = v_i w^i$, so then $Q^T w^i = v_i w^i$. The vector w^i is an eigenvector of Q^T associated with the eigenvalue v_i for $i = 2,3,\ldots,N$.

Now we prove that the set of eigenvectors of Q, $\{e^T, w^2, w^3, \ldots, w^N\}$ is linearly independent. Suppose that:

$$\alpha_1 e^T + \alpha_2 w^2 + \ldots + \alpha_N w^N = 0, \tag{8}$$

and $v_i \neq 0$ for $i = 2,3,\ldots,p$ and $v_j = 0$ for $j = p+1,\ldots,N$. We know that $Q^T e^T = 0e^T$ and $Q^T w^i = v_i w^i$ for $i = 2,3,\ldots,N$. Adding these two equations on LHS and RHS respectively yield:

$$Q^T e^T + Q^T w^2 + \ldots + Q^T w^N = 0e^T + v_2 w^2 + \ldots + v_p w^p.$$

Now, substitute $e^T, w^2, w^3, \ldots, w^N$ with $\alpha_1 e^T, \alpha_2 w^2, \alpha_3 w^3, \ldots, \alpha_N w^N$ of equation (8) and obtain

$$\begin{aligned} \alpha_1 Q^T e^T + \alpha_2 Q^T w^2 + \ldots + \alpha_N Q^T w^N \\ = \alpha_2 v_2 w^2 + \ldots + \alpha_p v_p w^p \end{aligned} \tag{9}$$

$$Q^T\left(\alpha_1 e^T + \alpha_2 w^2 + \ldots + \alpha_N w^N\right) = \alpha_2 v_2 w^2 + \ldots + \alpha_p v_p w^p = 0,$$

Since we consider the set $\{w^2, w^3, \ldots, w^N\}$ to be linearly independent, we then get $\alpha_2 = \alpha_3 = \ldots = \alpha_p = 0$, and we can now write equation (8) as:

$$\alpha_1 e^T + \alpha_{p+1} w^{p+1} + \ldots + \alpha_N w^N = 0, \tag{10}$$

$$S^T\left(\alpha_1 e^T + \alpha_{p+1}w^{p+1} + \ldots + \alpha_N w^N\right) = 0,$$

$$\alpha_1 S^T e^T + \alpha_{p+1} S^T w^{p+1} + \ldots + \alpha_N S^T w^N = 0.$$

Since $S^T w^j = v_i w^j$ for $j = p+1,\ldots,N,$ we then get

$$\alpha_1 S^T e^T + \alpha_{p+1} v_{p+1} w^{p+1} + \ldots + \alpha_N v_N w^N = 0.$$

Since $v_j = 0$ for $j = p+1,\ldots,N,$ it yields

$$\alpha_1 S^T e^T = 0.$$

We then get $\alpha_1 = 0$ since $S^T e^T > 0$ and S is diagonalizable. From equation (10), we have

$$\alpha_{p+1} w^{p+1} + \ldots + \alpha_N w^N = 0. \tag{11}$$

We know that the set $\left\{w^{p+1}, w^{p+2}, \ldots, w^N\right\}$ is linearly independent, so $\alpha_{p+1} = \ldots = \alpha_N = 0.$ So we get $\alpha_1 = \ldots = \alpha_N = 0.$ This means that the set $\left\{e^T, w^2, w^3, \ldots, w^N\right\}$ is linearly independent and the spectrum of Q is $\{0, v_2, v_3, \ldots, v_N\}.$

Case 2: Consider that S^T is not diagonalizable or defective. We know that a defective matrix has less than N different eigenvalues. Assume that S^T has $q < N$ different eigenvalues where $v_1 = \mu_1, v_2, \ldots, v_q,$ and these eigenvalues can be written as follows:

$$v_1 = \mu_1 > |v_2| \geq |v_3| \geq \ldots \geq |v_q|. \tag{12}$$

So, S^T has the form $S^T = XJX^{-1}$, where the $J = diag\left\{J_1, J_2, \ldots, J_q\right\}$ is in a canonical form. Suppose that the square matrices $J_i,\ i = 1,2,\ldots,q$ be the Jordan blocks with various sizes, in the form of:

$$J_i = \begin{bmatrix} v_i & 1 & 0 & \cdots & 0 \\ 0 & v_i & 1 & \ddots & \vdots \\ \vdots & \ddots & v_i & \ddots & 0 \\ \vdots & 0 & \ddots & \ddots & 1 \\ 0 & \cdots & \cdots & 0 & v_i \end{bmatrix},$$

where v_i is an eigenvalue of S^T. Let $J_1 = [\mu_1],$ and X_i is the ith column vector of $X,\ i = 1,2,\ldots,N.$ Let l_i be the size of J_i of each Jordan block, where $i = 1,2,\ldots,q.$ We now have $S^T = XJX^{-1}$, and therefore

$$S^T X = XJ,$$

$$S^T\begin{bmatrix} X_1 & X_2 & X_3 & X_4 & \cdots \end{bmatrix}$$

$$= \begin{bmatrix} X_1 & X_2 & X_3 & X_4 & \cdots \end{bmatrix} \begin{bmatrix} \mu_1 & 0 & 0 & \cdots & 0 \\ 0 & v_2 & 1 & \ddots & \vdots \\ \vdots & \ddots & v_2 & \ddots & 0 \\ \vdots & 0 & \ddots & v_2 & 1 \\ 0 & \cdots & \cdots & 0 & \ddots \end{bmatrix}.$$

From the above equation, we get:

$$\begin{aligned} S^T X_2 &= v_2 X_2, \\ S^T X_3 &= X_2 + v_2 X_3, \\ S^T X_4 &= X_3 + v_2 X_4, \\ &\vdots \\ S^T X_{l_2+1} &= X_{l_2} + v_2 X_{l_2+1}, \\ S^T X_{l_2+2} &= v_3 X_{l_2+2}, \\ S^T X_{l_2+3} &= X_{l_2+1} + v_3 X_{l_2+2}, \\ &\vdots \end{aligned}$$

Just like in Case 1, $S^T X_2 = v_2 X_2$ and based on the equation (7), $v_2 z_0^T X_2 = z_0^T v_2 X_2 = z_0^T S^T X_2 = \mu_1 z_0^T X_2,$ $(v_2 - \mu_1) z_0^T X_2 = 0.$ From equation (12), $v_2 \neq \mu_1.$ So $z_0^T X_2 = 0.$ Hence,

$$Q^T = \left(S - z_0 e S\right)^T,$$

$$Q^T X_2 = \left(S - z_0 e S\right)^T X_2 = S^T X_2 - S^T e^T z_0^T X_2 = S^T X_2 - 0,$$

which means that $Q^T X_2 = v_2 X_2.$ This implies that X_2 is an eigenvector of Q^T associated with the eigenvalue $v_2.$

From the equation $S^T X_3 = X_2 + v_2 X_3,$ we get

$$v_2 z_0^T X_3 = z_0^T v_2 X_3 = z_0^T (S^T X_3 - X_2) = z_0^T S^T X_3 - 0.$$

By equation (7), we get $v_2 z_0^T X_3 = \left(\mu_1 z_0^T\right) X_3.$ Consequently, $(v_2 - \mu_1) z_0^T X_3 = 0.$ By equation (12), and since $v_2 \neq \mu_1,$ we obtain $z_0^T X_3 = 0.$ Therefore,

$$Q^T = \left(S - z_0 e S\right)^T,$$

$$\begin{aligned} Q^T X_3 &= \left(S - z_0 e S\right)^T X_3 = S^T X_3 - S^T e^T z_0^T X_3 \\ &= S^T X_3 = X_2 + v_2 X_3. \end{aligned}$$

Likewise, we obtain:

$$\begin{aligned} Q^T X_2 &= v_2 X_2 \\ Q^T X_3 &= X_2 + v_2 X_3 \\ Q^T X_4 &= X_3 + v_2 X_4 \\ &\vdots \\ Q^T X_{l_2+1} &= X_{l_2} + v_2 X_{l_2+1} \\ Q^T X_{l_2+2} &= v_3 X_{l_2+2} \\ Q^T X_{l_2+3} &= X_{l_2+1} + v_3 X_{l_2+2} \\ &\vdots \end{aligned}$$

Like in Case 1, we want to show that the set $\left\{e^T, X_i, i = 2,3,\ldots,N\right\}$ is linearly independent. Let $Y = \left[e^T, X_i, i = 2,3,\ldots,N\right]$. Therefore, $Q^T Y = Y diag\left\{[0], J_2, \quad \ldots, J_q\right\}$. We now have the spectrum of $Q,$ $\left\{0, v_2, v_3, \ldots, v_q\right\}$ which is similar to the spectrum of $Q^T.$ The spectral radius of Q is $|v_2|.$ Therefore we get the following result:

$$\rho\left(H'(z_0)\right) = \rho\left(\frac{Q}{\mu_1}\right) = \frac{|v_2|}{\mu_1} < 1,$$

since $\mu_1 > |v_2|$.

Now we can determine the convergence rate of Algorithm 1.

Theorem 7. Let A and $\left\{z_0^{(k)}\right\}$ be as in Theorem 6. Then the convergence rate of the sequence $\left\{z_0^{(k)}\right\}$ is Q-linear, which means, there exists a vector norm $\|\bullet\|$ such that

$$\limsup_{k\to\infty}\frac{\left\|z^{(k+1)}-z_0\right\|}{\left\|z^{(k)}-z_0\right\|}<1.$$

Proof. By Proposition 1, there exist an $\epsilon>0$ and a spectral norm $\|\bullet\|$ such that $\|H'(z_0)\|_\epsilon \le \rho(H'(z_0))+\epsilon$.
By Lemma 2:

$$\|H'(z_0)\|_\epsilon \le \rho(H'(z_0))+\epsilon<1. \tag{13}$$

Hence, by equation (6), we have $z^{(k+1)}=H\left(z^{(k)}\right)$, $k=1,2,\ldots,$ and $z_0=H(z_0)$. Therefore, $z^{(k+1)}-z_0=H\left(z^{(k)}\right)-H(z_0)$. Expand $z^{(k)}$ at z_0 by using the Taylor expansion, we get:

$$H\left(z^{(k)}\right)=H(z_0)+H'(z_0)\left(z^{(k)}-z_0\right)+o\left(\left\|z^{(k)}-z_0\right\|_\epsilon\right),$$

$$z^{(k+1)}-z_0=H'(z_0)\left(z^{(k)}-z_0\right)+o\left(\left\|z^{(k)}-z_0\right\|_\epsilon\right),$$

$$\frac{\left\|z^{(k+1)}-z_0\right\|_\epsilon}{\left\|\left(z^{(k)}-z_0\right)\right\|_\epsilon}=\|H'(z_0)\|_\epsilon\ ,$$

From equation (13), we can get

$$\limsup_{k\to\infty}\frac{\left\|z^{(k+1)}-z_0\right\|}{\left\|z^{(k)}-z_0\right\|}<1.$$

Therefore Algorithm 1 is Q-linear convergence.

5.0 CONCLUSION

Within this paper, we proved that the algorithm for finding the largest singular value of nonnegative rectangular tensors, as proposed by Zhou *et al*. [11], is convergent under weak irreducibility condition and has a Q-linear rate of convergence. This paper only presents the convergence properties of Algorithm 1. In regards to numerical tests, the reader can refer to the referenced studies [7, 11].

The study of rectangular tensors is relatively new. Another method for determining the largest singular value of rectangular tensors can be found in Zhang's study [23], and it has been proven to be convergent under some assumptions. Algorithm 1 has also been generalised to nonnegative polynomials, as presented in Ibrahim's study [24]. The method is also convergent.

Acknowledgement

The second author is thankful to the Universiti Pendidikan Sultan Idris (UPSI) and the Ministry of Higher Education (Malaysia) for financial support throughout this study, under the RAGS grant Vot 2014-0122-101-72.

References

[1] Knowles, J. K. and Sternberg, E. 1975. On The Ellipticity of The Equations of Nonlinear Elastostatics for A Specialmaterial. *J. Elasticity*. 5: 341-361.
[2] Knowles, J. K. and Sternberg, E. 1977. On The Failure of Ellipticity of The Equations for Finite Elastostatic Plane Strain. *Arch. Ration. Mech. Anal.* 63: 321-336.
[3] Rosakis, P. 1990. Ellipticity and Deformations with Discontinuous Deformation Gradients in Finite Elastostatics. *Arch. Ration. Mech. Anal.* 109: 1-37.
[4] Wang, Y. and Aron, M. 1996. A Reformulation of The Strong Ellipticity Conditions for Unconstrained Hyperelasticmedia. *J. Elasticity*. 44: 89-96.
[5] Dahl, D., Leinass, J. M. Myrheim, J. and Ovrum, E. 2007. A Tensor Product Matrix Approximation Problem in Quantum Physics. *Linear Algebra Appl.* 420: 711-725.
[6] Einstein, A., Podolsky, B. and Rosen, N. 1935. Can Quantum-Mechanical Description of Physical Reality be Considered Complete? *Phys. Rev.* 47: 777-780.
[7] Chang, K., Qi, L. and Zhou, G. 2010. Singular Values of a Real Rectangular Tensor. *Journal of Mathematical Analysis and Applications*. 370: 284-294.
[8] Wood, R. J. and O'Neill, M. J. 2007. Finding The Spectral Radius of A Large Sparse Non-Negative Matrix. *ANZIAM J.* 48: C330-C345.
[9] Wood, R. J. and O'Neill, M. J. 2007. An Always Convergent Method for Finding The Spectral Radius of An Irreducible Non-Negative Matrix. *ANZIAM J*. 45: C474-C485.
[10] Ng, M., Qi, L. and Zhou, G. 2009. Finding The Largest Eigenvalue of A Nonnegative Tensor. *SIAM J. Matrix Anal. Appl*. 31: 1090-1099.
[11] Zhou, G., Caccetta, L. and Qi, L. 2013. Convergence of An Algorithm for The Largest Singular Value of A Nonnegative Rectangular Tensor. *Linear Algebra and its Applications*. 438: 959-968.
[12] Jordehi, A.R. 2015. Enhanced Leader PSO (ELPSO): A New PSO Variant for Solving Global Optimisation Problems. *Applied Soft Computing.* 26: 401-417.
[13] Jordehi, A. R., Jasni, J., Wahab, N. A., Kadir M. Z., and Javadi, M. S. 2015. Enhanced Leader PSO (ELPSO): A New Algorithm for Allocating Distributed TCSC's in Power Systems. *International Journal of Electrical Power & Energy Systems.* 64: 771-784.
[14] Jordehi, A.R. 2015. Brainstorm Optimisation Algorithm (BSOA): An Efficient Algorithm for Finding Optimal Location and Setting of FACTS Devices in Electric Power Systems. *International Journal of Electrical Power & Energy Systems*. 69: 48-57.
[15] Jordehi, A. R. 2014. A Chaotic-Based Big Bang-Big Crunch Algorithm For Solving Global Optimisation Problems. *Neural Computing and Applications*. 25: 1329-1335.
[16] Jordehi, A. R. 2015. Chaotic Bat Swarm Optimisation (CBSO). *Applied Soft Computing.* 26: 523-530.
[17] Varga, R. 1965. *Matrix Iterative Analysis.* Englewood Cliffs, New Jersey: Prentice-Hall, Inc.,
[18] Minc, H. 1988. *Nonnegative Matrices*. New York: John Wiley and Sons.
[19] Golub, G. H. and Loan, C. F. V. 1996. *Matrix Computations.* Baltimore: Johns Hopkins University Press.
[20] Horn, R. A. and Johnson, C. R. 1985. *Matrix Analysis*. Cambridge University Press.
[21] Friedland, S., Gaubert, S. and Han, L. 2011. Perron-Frobenius Theorem for Nonnegative Multilinear Forms and Extensions. *Linear Algebra Appl.* 438: 738-749.
[22] Zhou, G., Qi, L., and Wu, S. 2013. Efficient Algorithms for Computing The Largest Eigenvalue of A Nonnegative Tensor. *Front. Math. China*. 8: 155-168.
[23] Zhang, L. 2013. Linear Convergence of An Algorithm for Largest Singular Value of A Nonnegative Rectangular Tensor. *Front. Math. China*. 8: 141-153.
[24] Ibrahim, N. F. 2014. An Algorithm for The Largest Eigenvalue of Nonhomogeneous Nonnegative

Polynomials. *Numerical Algebra, Control and Optimization.* 4: 75-91.