American Journal of Computational Mathematics, **2011, 1, 63-71**
doi:10.4236/ajcm.2011.12007 Published Online June 2011 (http://www.scirp.org/journal/ajcm)

The Conditions for the Convergence of Power Scaled Matrices and Applications

Xuzhou Chen[1], Robert E. Hartwig[2]
[1]*Department of Computer Science*, *Fitchburg State University*, *Fitchburg*, *USA*
[2]*Department of Mathematics*, *North Carolina State University*, *Raleigh*, *USA*
E-mail: *xchen@fitchburgstate.edu*, *hartwig@math.ncsu.edu*
Received February 28, 2011; *revised May* 30, 2011; *accepted June* 7, 2011

Abstract

For an invertible diagonal matrix D, the convergence of the power scaled matrix sequence $D^{-N}A_N$ is investigated. As a special case, necessary and sufficient conditions are given for the convergence of $D^{-N}T^N$, where T is triangular. These conditions involve both the spectrum as well as the diagraph of the matrix T. The results are then used to privide a new proof for the convergence of subspace iteration.

Keywords: Convergence, Iterative Method, Triangular Matrix, Gram-Schmidt

1. Introduction

The aim of iterative methods both in theory as well as in numerical settings, is to produce a sequence of matrices $A_0, A_1, \cdots$, that converges to hopefully, something useful. When this sequence diverges, the natural question is how to produce a new converging sequence from this data. One of these convergence producing methods is to diagonally scale the numbers A_N and form the sequence $\{D_N A_N\}$. Examples of this are numerous, such as the Krylov sequence (x, Ax, A^2x ,…), which when divergent can be suitably scaled to yield a dominant eigenvector.

The convergence of power scaled iterative methods and more general power scaled Cesaro sums were studied by Chen and Hartwig [4,6]. In this paper, we continue our investigation of this iteration and derive a formula for the powers of an upper triangular matrix, and use this to investigate the convergence of the sequence $\{D_n^{-N}T^N\}$.

We also investigate the subspace iterations, which has been started by numerous authorss [1,3,10,11,15], and turn our attention to the case of repeated eigenvalues.

The main contributions of this paper are:

•We present the necessary and sufficient conditions for convergence of power scaled triangular matrices $\{D_n^{-N}T^N\}$. We prove that these conditions involve both the spectrum as well as the digraph induced by the matrix T.

•We apply the the convergence of power scaled triangular matrices with the explicit expression for the G-S factors of $D^{-N}T^N$ [3] and present a new proof of the convergence of simultaneous iteration for the case where the eigenvalues of the matrix A satisfy

$$|\lambda_1| \ge |\lambda_2| \ge \cdots \ge |\lambda_r| > |\lambda_{r+1}| \ge \cdots \ge |\lambda_n|$$
$$\text{and } |\lambda_i| = |\lambda_j| \Rightarrow \lambda_i = \lambda_j .$$

Because of the explicit expression for the GS factors, and the exact convergence results, our discussion is more precise than that given previously [12,17].

One of the needed steps in our investigation is the derivation a formula for the powers of a triangular matrix T, which in turn will allow us to analyze the convergence of $D_T^{-N}T^N$.

Throughout this note all our matrices will be complex and, as always, we shall use $\|\cdot\|$ and $\rho(\cdot)$ to denote the Euclidean norm and spectral radius of $(\cdot)$.

This paper is arranged as follows. As a preliminary result, a formula for the power of an upper triangular matrix is presented in Section 2. It is shown in Section 3 that the convergence of $D_T^{-N}T^N$ is closely related to the digraph induced by T. Section 4 is the main section in which convergence of general power scaled sequence $D^{-N}A_N$ is investigated and this, combined with path conditions in Section 3, is then used to discuss the convergence of $D^{-N}T^N$. As an application we analyze the convergence results for subspace iterations, in which the eigenvalues are repeated, but satisfy a peripheral constraint.

2. Preliminary Results

We first need a couple of preliminary results.

Lemma 2.1. If $\rho(A) < 1$ and $0 < \varepsilon_i < 1$, then

$$\sum_{k=0}^{N} A^k \varepsilon_k \tag{1}$$

converges.

Proof. For $f(z) = \sum_{k=0}^{\infty} \varepsilon_k z^k$, we have

$$| f(z) | = | \sum_{k=0}^{\infty} \varepsilon_k z^k | \le \sum_{k=0}^{\infty} | z |^k .$$

As the geometric summation on the right-hand side has radius of convergence 1, $f(z)$ converges for all z such that $|z| < 1$, which in turn tells us that the radius of convergence of $f(z)$ is no less than 1. Therefore, from Theorem 6.2.8. of [8], $f(A)$ converges.

Next consider the triangular matrix

$$\boldsymbol{U} = \begin{bmatrix} \mu_1 & u_{12} & \cdots & u_{1n} \\ 0 & \mu_2 & \ddots & \vdots \\ \vdots & \ddots & \ddots & u_{n-1,n} \\ 0 & \cdots & 0 & \mu_n \end{bmatrix}, \tag{2}$$

which is used in the following characterization of the powers of a trangular matrix.

Lemma 2.2. Let $\boldsymbol{T} = \begin{bmatrix} \lambda & a^T & \beta \\ 0 & U & c \\ 0 & 0 & \nu \end{bmatrix}$ where a and c are column vectors and suppose that

$$\boldsymbol{T}^N = \begin{bmatrix} \lambda^N & a_N^T & \beta_N \\ 0 & U^N & c_N \\ 0 & 0 & \nu^N \end{bmatrix}. \tag{3}$$

then

$$\beta_N = \beta\left(\sum_{k=0}^{N-1} \lambda^{N-k-1}\nu^k\right) + \sum_{k=0}^{N-2}\left(a^T U^k c\right)\left(\sum_{i=0}^{N-k-2}\lambda^{N-k-i-2}\nu^i\right). \tag{4}$$

in particular,

1) if $\lambda = 0$, then

$$\beta_N = \beta\nu^{N-1} + \sum_{k=0}^{N-2}\left(a^T U^k c\right)\nu^{N-k-2}, \tag{5}$$

2) if $\nu = 0$, then

$$\beta_N = \beta\lambda^{N-1} + \sum_{k=0}^{N-2}\left(a^T U^k c\right)\lambda^{N-k-2}, \tag{6}$$

3) if $\lambda \ne 0$ and $\lambda \ne \nu$, then

$$\beta_N = \lambda^N \beta\left(\frac{1-(\nu/\lambda)^N}{\lambda - \nu}\right) + \lambda^{N-1}\sum_{k=0}^{N-2} a^T \left(\frac{U}{\lambda}\right)^k c\left(\frac{1-(\nu/\lambda)^{N-k-1}}{\lambda-\nu}\right), \tag{7}$$

4) if $\lambda \ne 0$ and $\lambda = \nu$, then

$$\beta_N = N\beta\lambda^{N-1} + \lambda^{N-2}\sum_{k=0}^{N-2} a^T\left(\frac{U}{\lambda}\right)^k c(N-k-1) \quad \text{and} \tag{8}$$

5) if $\lambda = \nu = 0$, then

$$\beta_N = a^T U^{N-2} c . \tag{9}$$

Proof. It is easily verified by induction that $\boldsymbol{T}^N = \begin{bmatrix} T_1^N & y_N \\ O & \nu^N \end{bmatrix}$, where

$$\boldsymbol{T}_1^k = \begin{bmatrix} \lambda & a^T \\ O & U \end{bmatrix}^k = \begin{bmatrix} \lambda^k & \sum_{j=0}^{k-1}\lambda^{k-j-1} a^T U^j \\ O & U^k \end{bmatrix} \tag{10}$$

and

$$\boldsymbol{y}_N = \begin{bmatrix} \beta_N \\ c_N \end{bmatrix} = \sum_{k=0}^{N-1} T_1^{N-k-1}\begin{bmatrix} \beta \\ c \end{bmatrix}\nu^k . \tag{11}$$

Now

$$\begin{aligned} \boldsymbol{y}_N &= \sum_{k=0}^{N-1}\begin{bmatrix} \lambda^{N-k-1} & \sum_{j=0}^{N-k-2}\lambda^{N-k-j-2}a^T U^k \\ O & U^{N-k-1} \end{bmatrix}\begin{bmatrix} \beta \\ c \end{bmatrix}\nu^k \\ &= \sum_{k=0}^{N-1}\begin{bmatrix} \beta\lambda^{N-k-1} + \sum_{j=0}^{N-k-2}\lambda^{N-k-j-2}a^T U^k c \\ U^{N-k-1}c \end{bmatrix}\nu^k \\ &= \begin{bmatrix} \beta\sum_{k=0}^{N-1}\lambda^{N-k-1}\nu^k + \sum_{k=0}^{N-2}\sum_{j=0}^{N-k-2}\lambda^{N-k-j-2}a^T U^k c\nu^k \\ \sum_{k=0}^{N-1} U^{N-k-1}c\nu^k \end{bmatrix} \end{aligned}.$$

Hence

$$\begin{aligned} \beta_N &= \beta\left(\sum_{k=0}^{N-1}\lambda^{N-k-1}\nu^k\right) + \sum_{k=0}^{N-2}\sum_{j=0}^{N-k-2}\lambda^{N-k-j-2}a^T U^k c\nu^k \\ &= \beta\left(\sum_{k=0}^{N-1}\lambda^{N-k-1}\nu^k\right) + \sum_{k=0}^{N-2}\left(a^T U^k c\right)\left(\sum_{j=0}^{N-k-2}\lambda^{N-k-j-2}\nu^k\right), \\ &= \beta\left(\sum_{k=0}^{N-1}\lambda^{N-k-1}\nu^k\right) + \sum_{j=0}^{N-2}\left(a^T U^j c\right)\left(\sum_{k=0}^{N-j-2}\lambda^{N-k-j-2}\nu^k\right) \end{aligned}$$

completing the proof of (4). The special cases (1) - (5) are easy consequences of (4).

Let us now illustrate how the power of T are related to its digraph.

3. The Digraph of T

Suppose $\boldsymbol{T}=\begin{bmatrix}\lambda & a^T & \beta\\ O & U & c\\ 0 & O & \nu\end{bmatrix}$ is an $(n+2)\times(n+2)$ upper triangular matrix. Correspondingly we select $n+2$ nodes $S_0, S_1,\cdots,S_n, S_{n+1}$, and consider the assignment

$$\begin{array}{c} \\ S_0\\ S_1\\ \vdots\\ \vdots\\ S_n\\ S_{n+1}\end{array}\begin{array}{c}\begin{array}{cccccc}S_0 & S_1 & \cdots & \cdots & S_n & S_{n+1}\end{array}\\ \begin{bmatrix}\lambda & & & a^T & & \beta\\ & \mu_1 & u_{12} & \cdots & u_{1n} & \\ & 0 & \mu_2 & & \vdots & \\ O & \vdots & \ddots & \ddots & u_{n-1,n} & c\\ & 0 & \cdots & 0 & \mu_n & \\ 0 & & & O & & \nu\end{bmatrix}\end{array} \qquad (12)$$

with $a=[a_1, a_2,\cdots,a_n]^T$ and $c=[c_1, c_2,\cdots,c_n]^T$.

We next introduce the digraph induced by T, *i.e.* $G=(V, E)$ where $V=\{S_0, S_1,\cdots,S_{n+1}\}$ is the vertex set and $E=\{(S_i, S_j)\,|\,t_{ij}\neq 0\}$ is the edge set. As usual we say $(S_i, S_j)\in E$ if and only if $t_{ij}\neq 0$. A path from S_j to S_k in G is a sequence of vertices $S_j=S_{r_1}$, $S_{r_2},\cdots,S_{r_l}=S_k$ with $(S_{r_i}, S_{r_{i+1}})\in E$, for $i=1,\cdots,l-1$, for some l. If there is a path from S_j to S_k, we say that S_j has *access* to S_k and S_k can be reached from S_j. We write

$$\begin{aligned} &S_i\to S_j && \text{if } (S_i,S_j)\in E,\\ &S_i\to\to S_j && \text{if there is a path from } S_i \text{ to } S_j,\\ &S_i\leftarrow\leftrightarrow\to S_j && \text{if } S_i\to\to S_j \text{ and } S_j\to\to S_i\end{aligned}$$

Let $\pi=\langle S_0, S_{n+1}\rangle=\{S_{p_1},\cdots,S_{p_t}\}$ be the sandwich set of S_0 and S_{n+1}, *i.e.*, $\{S_{p_1},\cdots,S_{p_t}\}$ is the set of all the nodes from $\{S_1,\cdots,S_n\}$ such that $S_0\to\to S_{p_i}\to\to S_{n+1}$, *i.e.*, S_{p_i} can be reached from S_0 and has access to S_{n+1}. Let us now introduce the notation

$$\begin{aligned} a &= [a_{p_1},\cdots,a_{p_t}]^T,\\ \hat{U} &= U\begin{pmatrix}p_1\cdots p_t\\ p_1\cdots p_t\end{pmatrix},\\ c &= [c_{p_1},\cdots,c_{p_t}]^T.\end{aligned}$$

Then we have the following result.

Lemma 3.1. $a^TUc=a^T\hat{U}c$.

Proof. If $a_iu_{ij}c_j\neq 0$, then (S_0, S_i), (S_i, S_j), $(S_j, S_{n+1})\in E$, thus $S_i, S_j\in\pi$, which implies that

$$a^TUc=\sum_{i=1}^n\sum_{j=1}^n a_iu_{ij}c_j=\sum_{i\in\pi}\sum_{j\in\pi}a_iu_{ij}c_j.$$

This completes the proof. □

This following corollaries are the direct consequences of the above lemma.

Corollary 3.2. If $S_0\to\to S_{n+1}$ and there is no intermediate node that lies in $\{S_1,\cdots,S_n\}$ on any path from S_0 to S_{n+1}, then

1) $S_0\to S_{n+1}$, *i.e.* $\beta\neq 0$,
2) $a^TU^ic=a^T\hat{U}^ic=0$

Corollary 3.3. $a^TU^ic=a^T\hat{U}^ic$, for $i=1, 2,\cdots$.

We now turn to the main theorem of this section.

Theorem 3.4. Let $\boldsymbol{T}=\begin{bmatrix}\lambda_1 & t_{12} & \cdots & t_{1n}\\ 0 & \ddots & \ddots & \vdots\\ \vdots & & \ddots & t_{n-1,n}\\ 0 & \cdots & 0 & \lambda_n\end{bmatrix}$ be nonsingular and $\boldsymbol{D_T}=diag(T)=diag(\lambda_1,...,\lambda_n)$. Then, the following statement are equivalent

1) $D_T^{-N}T^N$ converges.
2) if $S_i\to\to S_j$, then $|\lambda_i|>|\lambda_j|$, *i.e.* if there is a path from S_i to S_j, then $|\lambda_i|>|\lambda_j|$.

Proof. We prove the theorem by induction on n. For $n=2$, $\boldsymbol{T}=\begin{bmatrix}\lambda_1 & \beta\\ 0 & \lambda_2\end{bmatrix}$ and $\boldsymbol{M}_2^{(N)}=\boldsymbol{D}_T^{-N}\boldsymbol{T}^N=\begin{bmatrix}1 & \beta_N/\lambda_1^N\\ 0 & 1\end{bmatrix}$ where

$$\frac{\beta_N}{\lambda_1^N}=\begin{cases}\beta[1-(\lambda_2/\lambda_1)^N]/(\lambda_1-\lambda_2) & \text{if } \lambda_1\neq\lambda_2\\ \beta\cdot N/\lambda_1 & \text{if } \lambda_1=\lambda_2\end{cases}.$$

It is easily seen that the convergence of $M^{(2)}$ implies that of $\frac{\beta_N}{\lambda_1^N}$. Hence if $\lambda_1=\lambda_2$, then $\beta=0$. Conversely, if $\beta\neq 0$, then $|\lambda_2/\lambda_1|<1$ which implies that $M^{(2)}$ converges.

Next, assume that the result holds for all triangular matrices of size $n+1$ or less. Let T be defined as in (12) and set $\boldsymbol{D_T}:=diag(\lambda, \mu_1,...,\mu_n, \nu)=diag(\lambda, \Delta, \nu)$ which is nonsingular. Consider the vertex set $V=\{S_0,\cdots,S_{n+1}\}$ and the assignment

$$M_{n+2}^{(N)}=\begin{array}{c} \\ S_0\\ \vdots\\ S_{n+1}\end{array}\begin{array}{c}\begin{array}{cccc}S_0 & S_1\ \cdots\ S_n & S_{n+1}\end{array}\\ \begin{bmatrix}1 & a_N^T/\lambda^N & \beta_N/\lambda^N\\ O & \Delta^{-N}U^N & \Delta^{-N}c^N\\ 0 & O & 1\end{bmatrix}\end{array}. \qquad (13)$$

$1)\Rightarrow 2)$. Assume that $M_{n+2}^{(N)}$ converges. Then by induction, both $\begin{bmatrix}\lambda & a^T\\ O & U\end{bmatrix}$ and $\begin{bmatrix}U & c\\ O & \nu\end{bmatrix}$ obey the theorem. Suppose $S_i\to\to S_j$ in V. If $|i-j|<n+1$ we are done since then both endpoints lie in $\{S_0,\cdots,S_n\}$ or $\{S_1,\cdots,S_{n+1}\}$. So we only need to consider the case where $S_i=S_0$ and $S_j=S_{n+1}$, *i.e.* $S_0\to\to S_{n+1}$.

Subcase (a): There is an intermediate node from

$\{S_1,\cdots,S_n\}$, say $S_0 \to\to S_p \to\to S_{n+1}$ ($1 \le p \le n$). Then by the induction hypothesis $|\lambda|>|\mu_p|>|\nu|$, and we are done.

Subcase (b): There is no intermediate node between S_0 and S_{n+1}. In this case $S_0 \to S_{n+1}$, and by Corollary 3.2., $\beta \ne 0$ and $a^T U^i c = a^T \hat{U}^i c = 0$ for arbitrary i. Since the sandwich set π is empty, we see from Lemma 2.2., that

$$\lambda^{-N}\cdot\beta_N = \begin{cases} \beta[1-(\nu/\lambda)^N]/(\lambda-\nu) & \text{if } \lambda \ne \nu \\ \beta\cdot N/\lambda & \text{if } \lambda = \nu \end{cases}. \quad (14)$$

Now because we are given that $\lambda^{-N}\beta_N$ converges and $\beta \ne 0$, we must have $|\nu/\lambda|<1$.

Conversely, assume that $S_i \to\to S_j \Rightarrow |\lambda_i|>|\lambda_j|$ and assume that the hypothesis holds for matrices of size $n+1$ or less. Since the graph condition also hold for $\{S_0,\cdots,S_n\}$ and $\{S_1,\cdots,S_{n+1}\}$, it follows by the hypothesis that all the entries in $M_{n+2}^{(N)}$ converges, with the possible exception of β_N/λ^N. Consequently, all we have to show is that $\lambda^{-N}\beta_N$ also converges, given the path conditions. Consider

$$\lambda^{-N}\beta_N = \begin{cases} \beta\dfrac{1-(\nu/\lambda)^N}{\lambda-\nu} + \\ \dfrac{1}{\lambda}\sum_{i=0}^{N-2} a^T\left(\dfrac{U}{\lambda}\right)^i c\left(\dfrac{1-(\nu/\lambda)^{N-i-1}}{\lambda-\nu}\right) \\ \quad if\ \lambda \ne \nu \\ \beta N/\lambda + \dfrac{1}{\lambda^2}\sum_{i=0}^{N-2} a^T\left(\dfrac{U}{\lambda}\right)^i c(N-i-1) \\ \quad if\ \lambda = \nu \end{cases} \quad (15)$$

If $S_0 \not\to\to S_{n+1}$, then $S_0 \not\to S_{n+1}$ and therefore $\beta = 0$. Moreover, π is empty and the right hand side of (15) is zero, *i.e.* $\lambda^{-N}\beta_N = 0$ and we are done. So suppose $S_0 \to\to S_{n+1}$ and thus $|\lambda|>|\nu|$. In this case $\beta\dfrac{1-(\nu/\lambda)^N}{\lambda-\nu}$ converges (possibly to 0 when $\beta = 0$). Now if $\pi = \varnothing$ then the second term of (15) vanishes by Lemma 2.2. Lastly suppose $\pi \ne \varnothing$, *i.e.* there are intermediate nodes $S_{p_1},\cdots,S_{p_t}$. From Lemma 2.2., we recall that $a^T U^i c = a^T \hat{U}^i c$ where

$$\hat{\boldsymbol{U}} = \begin{bmatrix} \mu_{p_1} & & * \\ & \ddots & \\ O & & \mu_{p_t} \end{bmatrix} \begin{matrix} S_{p_1} \\ \vdots \\ S_{p_t} \end{matrix}. \quad (16)$$

Since for each i, $S_0 \to\to S_{p_i} \to\to S_{n+1}$, we know that $|\lambda|>|\mu_{p_i}|>|\nu|$ and thus $|\lambda|>\rho(\hat{U})$. Hence $\rho(\hat{U}/\lambda)<1$ which implies that

$$a^T\sum_{i=0}^{N-2}\left(\frac{\hat{U}}{\lambda}\right)^i c \to a^T\left(I-\frac{\hat{U}}{\lambda}\right)^{-1} c.$$

To complete the proof we observe that

$$\sum_{i=0}^{N-2}\left(\frac{\hat{U}}{\lambda}\right)^i\left(\frac{\nu}{\lambda}\right)^{N-i-1}$$

also converges because of Lemma 2.1. with $A = \hat{U}/\lambda$ and $\varepsilon_i = (\nu/\lambda)^{N-i}$.

We at once have, as seen in [3].

Corollary 3.5. Let T be an upper triangular matrix and $\boldsymbol{D_T} = diag(T) = diag(\lambda_1,...,\lambda_n)$. If

$$|\lambda_1|>|\lambda_2|>\cdots>|\lambda_n|,$$

then $D_T^{-N}T^N$ converges to an upper triangular matrix of diagonal 1.

We now turn to the main result in this paper. Our aim is to characterize the convergence of $D^{-N}A_N$ in terms of the GS factorization of A_N.

4. Main Theorem

Let us denote the set of increasing sequences of p elements taken from ($1,\ 2,\cdots,m$) by

$$Q_{p,\,m} = \{I = (i_1,\cdots,i_p) \mid 1 \le i_1 < \cdots < i_p \le m\}$$

and assume this set $Q_{p,\,m}$ is ordered lexicographically. Suppose $\langle s:t\rangle := (s,\ s+1,...,t)$ is a subsequence of ($1,\ 2,...,m$) and we define

$$Q_p\langle s:t\rangle = \{U = (u_1,\cdots,u_p) \mid s < u_1 < \cdots < u_p < t\}.$$

Clearly, $Q_{p,\,m} = Q_p\langle 1:m\rangle$.

Suppose B is $m\times n$ matrix of rank r. The determinant of a $p\times p$ submatrix of A ($1 \le p \le \min(m,\ n)$), obtained from A by striking out $m-p$ rows and $n-p$ columns, is called a minor of order p of A. If the rows and columns retained are given by subscripts (see Householder [9]) $I = (i_1,\cdots,i_p) \in Q_{p,\,m}$ and $J = (j_1,\cdots,j_p) \in Q_{p,\,n}$ respectively, then the corresponding $p\times p$ submatrix and minor are respectively denoted by A_J^I and $det(A_J^I)$.

The minors for which $I = J$ are called the principal minors of A of order p, and the minors with $I = J = (1,\ 2,\cdots,p)$ are referred to as the leading principal minors of A.

Let $I = (i_1,\cdots,i_p) \in Q_{p,\,m}$ and $J = (j_1,\cdots,j_q) \in Q_{q,\,m}$. For convenience, we denote by $I[i_k] \in Q_{p-1,m}$ the sequences of $p-1$ elements obtained by striking out the *kth* element i_k; while $I(j)$ denotes the sequences of $p+1$ elements obtained by adding a new element j after i_k, *i.e.*, $I(j) = (i_1,\cdots,i_k,\ j)$. Note that if $i_p > j$, then $I(j)$ is not an element of $Q_{p+1,\,m}$ because it is no longer an increasing sequence. If $p+q \le m$, we denote the concaternation $(i_1,\cdots,i_p,\ j_1,$

$\cdots, j_q)$ of I and J by IJ. It has $p+q$ elements. Again, IJ may not be an element of $Q_{p+q,\,m}$.

Since the natural sequence ($1, 2, \cdots, p$) of p elements will be used frequently, we particularly denote this sequence by $\langle p\rangle = (1, 2, \cdots, p)$; while $\langle p\rangle[t] = (1, \cdots, t-1, t+1, \cdots, p)$ is simply denoted by $\langle p\setminus t\rangle$.

Next recall [2] that the volume $Vol(B)$ of a real matrix B, is defined as the product of all the nonzero singular values of B. It is known [2] that

$$Vol(B) = \sqrt{\sum |\det(B_J^I)|^2}, \tag{17}$$

where B_J^I are all $r\times r$ submatrices of B. In particular, if B has full column rank, then

$$Vol(B) = \sqrt{\det(B^* B)}. \tag{18}$$

Lastly, suppose $A = [a_1, a_2, \cdots, a_r]$ is an $n\times r$ matrix of full column rank and

$$A = YG \tag{19}$$

is its GS factorization so that the columns of $Y = [y_1, y_2, \cdots, y_r]$ are orthogonal and G is $r\times r$ upper triangular matrix of diagonal 1. For $k \le r$, we define $A_k = [a_1, \cdots, a_k]$ and

$$V_k = Vol(A_k). \tag{20}$$

It follows directly that

$$V_k = \sqrt{\sum_{I\in Q_{k,\,m}} |\det(A_{\langle k\rangle}^I)|^2} = \sqrt{\det(A_k^* A_k)}. \tag{21}$$

Theorem 4.1. Let A be an $n\times r$ matrix of rank r and let $A = YG$ be its GS factorization. Then

$$y_{kl} = \sum_{I\in Q_{l-1,\,n}} \det(A_{\langle l\rangle}^{I(k)})\cdot \overline{\det(A_{\langle l-1\rangle}^{I})} / V_{l-1}^2 \tag{22}$$

and

$$g_{jk} = \frac{\det(A^* A)_{\langle j-1\rangle(k)}^{\langle j\rangle}}{V_j^2}. \tag{23}$$

Proof. The result of (22) follows from Theorem 2.1. in [3], while on account of Corollary 2.1. in [3], $G = (Y^* Y)^{-1} Y^* A$. Hence we arrive at

$$g_{jk} = \frac{V_{j-1}^2}{V_j^2}\cdot y_j^* a_k = \frac{V_{j-1}^2}{V_j^2}\sum_{l=1}^{n} \overline{y_{lj}} a_{lk}$$
$$= \sum_{l=1}^{n}\sum_{t=1}^{j} (-1)^{j+t}\, \overline{a_{lt}} a_{lk} \det(A^* A)_{\langle j-1\rangle}^{\langle j\setminus t\rangle} / V_j^2.$$

Because $\sum_{l=1}^{n} \overline{a_{lt}} a_{lk}$ is just the (t, k) element of matrix $A^* A$, we see that

$$g_{jk} = \sum_{t=1}^{j} (-1)^{j+t} (\sum_{l=1}^{n} \overline{a_{lt}} a_{lk}) \det(A^* A)_{\langle j-1\rangle}^{\langle j\setminus t\rangle} / V_j^2,$$

which is the Laplace expansion along column j of $\det(A^* A)_{\langle j-1\rangle(k)}^{\langle j\rangle}$. Thus

$$g_{jk} = \frac{\det(A^* A)_{\langle j-1\rangle(k)}^{\langle j\rangle}}{V_j^2},$$

completing the proof.

Remark: A different proof of (23) was given in [9, § 1.4].

For a diagonal matrix $\boldsymbol{D} = diag(d_1, ..., d_n)$, we say that D is *decreasing*, if

$$|d_1| \ge \cdots \ge |d_n|. \tag{24}$$

Moreover, D is called *locally primitive*, if it is decreasing and

$$|d_i| = |d_j| \Rightarrow d_i = d_j. \tag{25}$$

It is obvious that we can partition a decreasing matrix D as

$$\boldsymbol{D} = diag(D^{(1)}, \cdots, D^{(t)}) \tag{26}$$

where each $\boldsymbol{D}^{(s)} = \delta_s\, diag(e^{i\theta_1^{(s)}}, \cdots, e^{i\theta_{p_s}^{(s)}})$ with $|\delta_1| > |\delta_2| > \cdots > |\delta_t|$. As a special case, if D is locally primitive, then D can be written as

$$\boldsymbol{D} = diag(\delta_1 I_{p_1}, \cdots, \delta_t I_{p_t}). \tag{27}$$

Now let us define $q_s = \sum_{j=0}^{s} p_j$ ($s = 1, \cdots t$, $q_0 = p_0 = 0$) and $Q_u\langle q_{i-1} : q_i\rangle = \{\Omega_u = (\omega_1, \cdots, \omega_u) \mid q_{i-1}\ \omega_1 < \cdots < \omega_u < q_i\}$. Next, suppose $A_N = [a_{ij}^{(N)}]_{n\times r}$ is a sequence of $n\times r$ matrices and let

$$A_N = Y_N G_N = [y_{ij}^{(N)}]_{n\times r} [g_{ij}^{(N)}]_{r\times r} \tag{28}$$

be their GS factorization. Suppose B is a $n\times r$ matrix, we can partition B conformally as D in (26). It is easily verified that the (u, v) element of (i, j) block B_{ij} of B is exactly the $(q_{i-1}+u,\ q_{j-1}+v)$ element of the whole matrix B. B is said to satisfy condition (β) if for each $k = q_{i-1}+u$ there exists $\Omega_u \in Q_u\langle q_{i-1} : q_i\rangle$ such that

$$\det B_{\langle k\rangle}^{q_{i-1}\Omega_u} \ne 0.$$

We now have the following theorem.

Theorem 4.2. Let A_N be a sequence of $n\times r$ matrices of full column rank with GS factor $A_N = Y_N G_N$. Also suppose D is a diagonal matrix and D_r is $r\times r$ leading submatrix of D. Then

1) $D^{-N} A_N$ converges to $\widetilde{B}$ which satisfies condition (β) $\Leftrightarrow$ G_N converges and $D^{-N} Y_N$ converges to Z which satisfies condition (β)

2) If in addition D is decreasing, *i.e.* D satisfies (26), then for $k = q_{i-1}+u$ and $l = q_{j-1}+v$ ($i \le j-1$)

$$\frac{y_{kl}^{(N)}}{d_k^N} = O\left(\left|\frac{\delta_j}{\delta_i}\right|^{2N}\right). \tag{29}$$

Proof. 1) The sufficiency is obvious. So let us turn to the necessary part. For $\boldsymbol{D} = diag(d_1, ..., d_n)$, there exists a permutation Q such that $\hat{D} = Q^* D Q$ is decreasing. Meanwhile, by hypothesis and the fact that $D^{-N} A_N = Q$

$(Q^*D^{-N}Q)(Q^*A_N) = Q\hat{D}^{-N}\hat{A}_N$ with $\hat{A}_N = Q^*A_N$, it follows that $\hat{D}^{-N}\hat{A}_N$ converges. So without loss of generality, we assume that D is decreasing and partition D as (26) and simply consider $D^{-N}A_N$. We shall now, without risk of confusion, abbreviate the set $Q_u\langle q_{i-1}:q_i\rangle = \{\Omega_u = (\omega_1,\cdots,\omega_u) \mid q_{i-1} < \omega_1 < \cdots < \omega_u < q_i\}$ to Q_u and for $I = (i_1,\ldots,i_s)$ set $\pi_I = d_{i_1}\cdots d_{i_s}$. It at once follows that

$$|\pi_{\langle q_{i-1}\rangle\Omega_u}| = |\pi_{\langle k\rangle}|. \tag{30}$$

We now have from (23)

$$\begin{aligned} g_{kl}^{(N)} &= \det(A_N^*A_N)^{\langle k\rangle}_{\langle k-1\rangle(l)}/V_k^2 \\ &= \frac{\sum_{I\in Q_{k,n}} \overline{\det(A_N)^I_{\langle k\rangle}}\cdot\det(A_N)^I_{\langle k-1\rangle(l)}}{\sum_{I\in Q_{k,n}}\left|\det((A_N)^I_{\langle k\rangle})\right|^2} \text{ (from Cauchy-Binet)} \\ &= \frac{\left(\sum_{I\in\langle q_{i-1}\rangle Q_u}+\sum_{I\notin\langle q_{i-1}\rangle Q_u}\right)\overline{\det(A_N)^I_{\langle k\rangle}}\cdot\det(A_N)^I_{\langle k-1\rangle(l)}}{\left(\sum_{I\in\langle q_{i-1}\rangle Q_u}+\sum_{I\notin\langle q_{i-1}\rangle Q_u}\right)\left|\det(A_N)^I_{\langle k\rangle}\right|^2} \\ &= \frac{\sum_{\Omega_u\in Q_u}\overline{\det(A_N)^{\langle q_{i-1}\rangle\Omega_u}_{\langle k\rangle}}\cdot\det(A_N)^{\langle q_{i-1}\rangle\Omega_u}_{\langle k-1\rangle(l)}+\sum_{I\notin\langle q_{i-1}\rangle Q_u}}{\sum_{\Omega_u\in Q_u}\left|\det((A_N)^{\langle q_{i-1}\rangle\Omega_u}_{\langle k\rangle})\right|^2+\sum_{I\notin\langle q_{i-1}\rangle Q_u}} \\ &= \frac{\sum_{\Omega_u\in Q_u}\left(\overline{\frac{\det(A_N)^{\langle q_{i-1}\rangle\Omega_u}_{\langle k\rangle}}{(\pi_{\langle k\rangle})^N}}\right)\cdot\frac{\det(A_N)^{\langle q_{i-1}\rangle\Omega_u}_{\langle k-1\rangle(l)}}{(\pi_{\langle k\rangle})^N}+o\left(\left|\frac{\delta_{i+1}}{\delta_i}\right|^{2N}\right)}{\sum_{\Omega_u\in Q_u}\left|\frac{\det(A_N)^{\langle q_{i-1}\rangle\Omega_u}_{\langle k\rangle}}{(\pi_{\langle k\rangle})^N}\right|^2+o\left(\left|\frac{\delta_{i+1}}{\delta_i}\right|^{2N}\right)}, \end{aligned}$$

On account of (30), this is equal to

$$\frac{\sum_{\Omega_u\in Q_u}\left(\overline{\frac{\det(A_N)^{\langle q_{i-1}\rangle\Omega_u}_{\langle k\rangle}}{(\pi_{\langle q_{i-1}\rangle\Omega_u})^N}}\right)\cdot\frac{\det(A_N)^{\langle q_{i-1}\rangle\Omega_u}_{\langle k-1\rangle(l)}}{(\pi_{\langle q_{i-1}\rangle\Omega_u})^N}+o\left(\left|\frac{\delta_{i+1}}{\delta_i}\right|^{2N}\right)}{\sum_{\Omega_u\in Q_u}\left|\frac{\det(A_N)^{\langle q_{i-1}\rangle\Omega_u}_{\langle k\rangle}}{(\pi_{\langle q_{i-1}\rangle\Omega_u})^N}\right|^2+o\left(\left|\frac{\delta_{i+1}}{\delta_i}\right|^{2N}\right)}. \tag{31}$$

Since $D^{-N}A_N$ convergence, so does the submatrices $(D^{-N}A_N)^{\langle q_{i-1}\rangle\Omega_u}_I$ and their determinant and hence

$$\begin{aligned} \frac{\det(A_N)^{\langle q_{i-1}\rangle\Omega_u}_I}{(\pi_{\langle q_{i-1}\rangle\Omega_u})^N} &= \det[(D^{\langle q_{i-1}\rangle\Omega_u}_{\langle q_{i-1}\rangle\Omega_u})^{-N}(A_N)^{\langle q_{i-1}\rangle\Omega_u}_I] \\ &= \det(D^{-N}A_N)^{\langle q_{i-1}\rangle\Omega_u}_I \end{aligned}$$

converges, say, to $det(\widetilde{A}_N)^{\langle q_{i-1}\rangle\Omega_u}_I$. We have that consequently (31) converges to

$$\frac{\sum_{\Omega_u\in Q_u}\overline{\det(\tilde{A}^{\langle q_{i-1}\rangle\Omega_u}_{\langle k\rangle})}\cdot\det(\tilde{A}^{\langle q_{i-1}\rangle\Omega_u}_{\langle k-1\rangle(l)})}{\sum_{\Omega_u\in Q_u}\left|\det(\tilde{A}^{\langle q_{i-1}\rangle\Omega_u}_{\langle k\rangle})\right|^2},$$

in which the denominator is nonzero as $\widetilde{A}$ satisfies condition (β). Hence G_N converges and this implies that $D^{-N}Y_N = D^{-N}A_NG_N^{-1}$ also converges.

2) Lastly, what remains is to show that $Y_ND_r^{-N}$ converges if D is decreasing, *i.e.* D satisfies (26). Now for $k = q_{i-1}+u$, $l = q_{j-1}+v$ $(i \le j-1)$, it follows that

$$\begin{aligned} \frac{y_{kl}^{(N)}}{d_k^N} &= \frac{1}{d_k^N}\cdot\frac{\sum_{I\in Q_{l-1,n}}\det(A_N)^{I(k)}_{\langle l\rangle}\cdot\overline{\det(A_N)^I_{\langle l-1\rangle}}}{V_{l-1}^2} \\ &= \frac{1}{d_k^N}\frac{\left(\sum_{I\in\langle q_{j-1}\backslash k\rangle Q_v}+\sum_{I\notin\langle q_{j-1}\backslash k\rangle Q_v}\right)\det(A_N)^{I(k)}_{\langle l\rangle}\cdot\overline{\det(A_N)^I_{\langle l-1\rangle}}}{\left(\sum_{I\in\langle q_{j-1}\rangle Q_{v-1}}+\sum_{I\notin\langle q_{j-1}\rangle Q_{v-1}}\right)\left|\det(A_N)^I_{\langle k-1\rangle}\right|^2} \\ &= \frac{\sum_{\Omega_v\in Q_v}\det(A_N)^{\langle q_{j-1}\backslash k\rangle\Omega_v(k)}_{\langle l\rangle}\overline{\det(A_N)^{\langle q_{j-1}\backslash k\rangle\Omega_v}_{\langle l-1\rangle}}+\sum_{I\notin\langle q_{j-1}\backslash k\rangle Q_v}}{d_l^N\left(\sum_{\Omega_{v-1}\in Q_{v-1}}\left|\det((A_N)^{\langle q_{j-1}\rangle\Omega_{v-1}}_{\langle l-1\rangle})\right|^2+\sum_{I\notin\langle q_{j-1}\rangle Q_{v-1}}\right)} \\ &= \frac{1}{d_k^N}\cdot\frac{\sum_{\Omega_v\in Q_v}\frac{\det(A_N)^{\langle q_{j-1}\backslash k\rangle\Omega_v(k)}_{\langle l\rangle}}{(\pi_{\langle l-1\rangle})^N}\left(\overline{\frac{\det(A_N)^{\langle q_{j-1}\backslash k\rangle\Omega_v}_{\langle l-1\rangle}}{(\pi_{\langle l-1\rangle})^N}}\right)+\sum_{I\notin\langle q_{j-1}\rangle Q_v}}{\sum_{\Omega_{v-1}\in Q_{v-1}}\left|\frac{\det((A_N)^{\langle q_{j-1}\rangle\Omega_{v-1}}_{\langle l-1\rangle})}{\pi_{\langle l-1\rangle})^N}\right|^2+\sum_{I\notin\langle q_{j-1}\rangle Q_{v-1}}} \end{aligned}$$

$$=\frac{|d_l|^{2N}}{|d_k|^{2N}}\cdot\frac{\sum_{\Omega_v\in Q_v}\frac{\det(A_N)_{\langle l\rangle}^{\langle q_{j-1}\backslash k\rangle\Omega_v(k)}}{(\pi_{\langle l\rangle})^N}\overline{\left(\frac{\det(A_N)_{\langle l-1\rangle}^{\langle q_{j-1}\backslash k\rangle\Omega_v}}{(\pi_{\langle l\backslash k\rangle})^N}\right)}+o\left(\left(\frac{\delta_{j+1}}{\delta_{i+1}}\right)^{2N}\right)}{\sum_{\Omega_{v-1}\in Q_{v-1}}\left|\frac{\det((A_N)_{\langle l-1\rangle}^{\langle q_{j-1}\rangle\Omega_{v-1}})}{(\pi_{\langle q_{j-1}\rangle\Omega_{v-1}})^N}\right|^2+o\left(\left(\frac{\delta_{j+1}}{\delta_j}\right)^{2N}\right)}$$

$$=\frac{|\delta_j|^{2N}}{|\delta_i|^{2N}}\cdot\frac{\sum_{\Omega_v\in Q_v}\frac{det(A_N)_{\langle l\rangle}^{\langle q_{j-1}\backslash k\rangle\Omega_v(k)}}{(\pi_{\langle q_{j-1}\backslash k\rangle\Omega_v(k)})^N}\overline{\left(\frac{\det(A_N)_{\langle l-1\rangle}^{\langle q_{j-1}\backslash k\rangle\Omega_v}}{(\pi_{\langle q_{j-1}\backslash k\rangle\Omega_v})^N}\right)}+o\left(\left(\frac{\delta_{j+1}}{\delta_{i+1}}\right)^{2N}\right)}{\sum_{\Omega_{v-1}\in Q_{v-1}}\left|\frac{\det((A_N)_{\langle l-1\rangle}^{\langle q_{j-1}\rangle\Omega_{v-1}})}{(\pi_{\langle q_{j-1}\rangle\Omega_{v-1}})^N}\right|^2+o\left(\left(\frac{\delta_{j+1}}{\delta_j}\right)^{2N}\right)}=O\left(\left|\frac{\delta_j}{\delta_i}\right|^{2N}\right)$$

This completes the proof of 2).

As a consequence of the above theorem we have

Corollary 4.3. Suppose D is decreasing and A_N's have orthogonal columns. If $D^{-N}A_N$ converges to $\tilde{B}$ which satisfies condition (β), then for $k=q_{i-1}+u$ and $l=q_{j-1}+v$ $(i\le j-1)$

$$\frac{a_{kl}^{(N)}}{d_k^N}=O\left(\left|\frac{\delta_j}{\delta_i}\right|^{2N}\right)$$

Proof. In this case the GS factorization of A_N are $A_N=A_NI_r$. So the result is the direct consequnce of Theorem 4.2.

Lemma 4.4. Suppose D is decreasing and A_N's are of full column rank. If $B_N=[B_{pq}^{(N)}]=D^{-N}A_N$ converges, say, to $\tilde{B}$, then $A_ND_r^{-N}$ converges iff

1) $(\tilde{B}_{jj})_{u,v}=\tilde{B}_{q_{j-1}+u,\,q_{j-1}+v}\ne 0\Rightarrow\theta_u^{(j)}=\theta_v^{(j)}$ $(j=1,\cdots,t)$

2) If $i<j$, then

$$\left(\frac{\delta_i}{\delta_j}\right)^N B_{q_{i-1}+u,\,q_{j-1}+v}^{(N)}e^{iN(\theta_u^{(i)}-\theta_v^{(j)})}$$

converges.

Proof. It is not difficult to see that the $(q_{i-1}+u,\ q_{j-1}+v)$ element of $A_ND_r^{-N}$ is

$$(A_ND_r^{-N})_{q_{i-1}+u,\,q_{j-1}+v}=\left(\frac{\delta_i}{\delta_j}\right)^N B_{q_{i-1}+u,\,q_{j-1}+v}^{(N)}e^{iN(\theta_u^{(i)}-\theta_v^{(j)})}. \tag{32}$$

As $B_{q_{i-1}+u,\,q_{j-1}+v}^{(N)}$ converges and $\left|\frac{\delta_i}{\delta_j}\right|<1$ for $i>j$, it follows that (32) converges to zero in this case. Hence $A_ND_r^{-N}$ converges iff i) and ii) hold.

Suppose B is an $n\times n$ matrix and correspondingly there are n nodes $S_1,S_2,\cdots,S_n$. We say that B is *indecomposable* if for every i and j

$$\text{either}\quad S_i\to\to S_j\quad\text{or}\quad S_j\to\to S_i.$$

Next we have

Theorem 4.5. Let A_N be of full column rank and $A_N=Y_NG_N$ be its GS factorization. Suppose $D^{-N}A_N$ converges, say, to $\tilde{B}$ which satisfies (β). Then the following statement are true

1) If $Y_ND_r^{-N}$ converges to $\tilde{\boldsymbol{Z}}=diag(\tilde{Z}_1,...,\tilde{Z}_s)$ in which each block $\tilde{Z}_i$ $(i=1,\cdots,s)$ is indecomposable, then $D^{(s)}=\delta_sI_{p_s}$, $s=1,\cdots,t$

2) If $D^{(s)}=\delta_sI_{p_s}$, $s=1,\cdots,t$, then $Y_ND_r^{-N}$ converges.

Proof. From Theorem 4.2., the convergence of $B_N=D^{-N}A_N$ implies the convergence of G_N and $D^{-N}Y_N$. Suppose $D^{-N}Y_N\to\tilde{Z}$. Then it follows, on account of Lemma 4.4, that $Y_ND_r^{-N}$ converges to $\tilde{\boldsymbol{Z}}=diag(\tilde{Z}_1,\cdots,\tilde{Z}_s)$ if

a) $(\tilde{Z}_j)_{u,v}=\tilde{Z}_{q_{j-1}+u,\,q_{j-1}+v}\ne 0\Rightarrow\theta_u^{(j)}=\theta_v^{(j)}$, and

b) if $i<j$, then

$$(Y_ND_r^{-N})_{q_{i-1}+u,\,q_{j-1}+v}=\left(\frac{\delta_i}{\delta_j}\right)^N(D^{-N}Y_N)_{q_{i-1}+u,\,q_{j-1}+v}e^{iN(\theta_u^{(i)}-\theta_v^{(j)})}$$

converges to zero. Now Corollary 4 says that for $i<j$

$$(D^{-N}Y_N)_{q_{i-1}+u,\,q_{j-1}+v}=\frac{Y_{q_{i-1}+u,\,q_{j-1}+v}^{(N)}}{d_{q_{i-1}+u}^N}=O\left(\left|\frac{\delta_j}{\delta_i}\right|^{2N}\right)$$

and so b) is automatically statisfied in this case. Therefore $Y_ND_r^{-N}$ converges iff a) holds. Since each $\tilde{Z}_i$ is indecomposable, for arbitrary (u,v) there exists a path either from $S_{q_{i-1}+u}$ to $S_{q_{i-1}+v}$ or vice verse. In either case this implies that $\theta_u^{(i)}=\theta_v^{(i)}$ for any u and v. We complete the proof of 1).

2) This time D is locally primitive, so we have $\theta_u^{(j)}=\theta_v^{(j)}$ $(j=1,\cdots,t)$ and hence

$$(Y_N D_r^{-N})_{q_{i-1}+u,\, q_{j-1}+v}$$

$$=\begin{cases}\left(\dfrac{\delta_i}{\delta_j}\right)^N (D^{-N}Y_N)_{q_{i-1}+u,\, q_{j-1}+v} e^{iN(\theta_u^{(i)}-\theta_v^{(j)})} & \text{if } i \neq j \\ (D^{-N}Y_N)_{q_{j-1}+u,\, q_{j-1}+v} & \text{if } j = i\end{cases}$$

By hypothesis, the above converges for $j = i$. The convergence for $i > j$ is obvious; while the convergence for $i < j$ can be easily achieved by noticing that

$$(D^{-N}Y_N)_{q_{i-1}+u,\, q_{j-1}+v} = \frac{Y^{(N)}_{q_{i-1}+u,\, q_{j-1}+v}}{d^N_{q_{i-1}+u}} = O\left(\left|\frac{\delta_j}{\delta_i}\right|^{2N}\right).$$

Remark. From Theorem 4.5. we know that in the case of multiple eigenvalues, if $k = q_{i-1} + u$, then

$$\frac{y_k^{(N)}}{d_i^N} \to [\overbrace{0,\cdots,0}^{1\cdots q_{i-1}},\ \overbrace{\tilde{p}_{q_{i-1}+1},\cdots,\tilde{p}_{q_i}}^{(q_{i-1}+1)\cdots q_i},\ 0,\cdots,0]^T.$$

Let us now turn to the applications of this theorem. Our first application is the following result gives the general convergence result of power scaled triangular matrix.

Corollary 4.6. Let D be diagonal and T be upper triangular. Then $D^{-N}T^N$ converges if and only if

1) Either $|\lambda_i/d_i| < 1$ or $\lambda_i = d_i$ for each i

2) If $S_i \to\to S_j \Rightarrow |\lambda_i| > |\lambda_j|$.

Proof. Let $A_N = T^N$. This time the GS factorization for $A_N = T^N$ becomes $D_T^N(D_T^{-N}T^N)$ and from Theorem 4.2., $D^{-N}T^N$ converges if and only if both $G_N = D_T^{-N}T^N$ and $D^{-N}D_T^N$ converge.

The convergence of $D^{-N}D_T^N$ is equivalent to 1); while the convergence of $G_N = D_T^{-N}T^N$, on account of Theorem 3.4., is exactly the same as the path condition 2).

A relevant application of Theorem 3.4. is to the question of subspace iteration. Armed with Theorem 3.4. we can get a sharper theoretical result than was previously given.

5. Application to the Subspace Iteration

Next, suppose T is an block upper diagonal matrix of the form

$$\boldsymbol{T} = \begin{bmatrix} \lambda_1 I_{p_1} & T_{12} & \cdots & T_{1t} \\ 0 & \lambda_2 I_{p_2} & \cdots & T_{2t} \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_t I_{p_t} \end{bmatrix}, \qquad (33)$$

where $|\lambda_1| > |\lambda_2| > \cdots > |\lambda_t|$. Let $\boldsymbol{D_T} = diag(T) = diag(\lambda_1 I_{p_1},\cdots,\lambda_t I_{p_t})$ and denote $D_{T_r} = (D_T)^{\langle r\rangle}_{\langle r\rangle}$. Then from Theorem 3.4. it follows that $D_T^{-N}T^N$ converges.

Assume B is $n\times r$ matrix of full column rank. Therefore $r \le n = \sum_{i=1}^t p_i$ and without loss of generality we can write $r = \sum_{i=1}^s p_i + w$ for some $w \le p_{s+1}$. Thus we can write $T_r = diag(\lambda_1 I_{p_1},\cdots,\lambda_s I_{p_s},\ \lambda_{s+1} I_w)$. We now have

Corollary 5.1. Let T be $n\times n$ upper triangular matrix defined as in (33), and let B be $n\times r$ matrix whose columns are linearly independent. If

$$T^N B = Y_N G_N$$

is its GS factorization, then the followings hold

1) $D_T^{-N}T^N B$ converges, say, to a limit $\tilde{A}$.

2) $Y_N D_{T_r}^{-N}$ converges to $\begin{bmatrix} P \\ 0 \end{bmatrix}$, where $P = diag(P_1,\cdots,P_s,\ \tilde{P})$ and each P_i ($i = 1,\cdots,s$) is a $p_i\times p_i$ matrix and $\tilde{P}$ is a $p_{s+1}\times w$ matrix.

Proof. The result follows by simply choosing $A_N = T^N B$ in Theorem 4.2.

Let us now turn to the question of subspace iteration for a restricted class of matrices. Suppose that

$$A = VTV^* \qquad (34)$$

is $n\times n$ matrix, where V is unitary and T is as in (33). Then using the same P_i as above we have

Corollary 5.2. Suppose that A is an $n\times n$ matrix which satisfies (34). let Y_0 be an $n\times r$ matrix whose columns are linearly independent and $\{Y_N\}$ be sequence of matrices defined by the following factorization

$$A^N Y_0 = Y_N G_N.$$

Then

$$Y_N D_{T_r}^{-N} \to [V_1P_1,\cdots,V_sP_s,\ V_{s+1}\tilde{P}]. \qquad (35)$$

Proof. Since

$$A^N Y_0 = Y_N G_N,$$

it follows that

$$VT^N(V^*Y_0) = Y_N G_N.$$

Partition $V = [V_1,\cdots,V_t]$ conformally to that of T in (33) and set $B = V^*Y_0$, then

$$T^N B = (V^*Y_N)G_N. \qquad (36)$$

It is easily seen that the columns of V^*Y_N are orthogonal. Therefore (36) can be regarded as the GS factorization of $T^N B$. From Corollary 5.1., we have that for $V = [V_1,\cdots,V_t]$

$$V^*Y_N D_{T_r}^{-N} \to \begin{bmatrix} P \\ 0 \end{bmatrix},$$

which is equivalent to

$$Y_N D_{T_r}^{-N} \to V\begin{bmatrix} P \\ 0 \end{bmatrix} = V_r P = [V_1 P_1, \cdots, V_s P_s,\ V_{s+1}\tilde{P}] .$$

6. References

[1] F. L. Bauer, "Das Verfahren der Treppeniteration und verwandte Verfahren zur Lösung Algebraischer Eigenwertprobleme," *Zeitschrift für Angewandte Mathe- matik und Physik*, Vol. 8, No. 3, 1957, pp. 214-235. doi:10.1007/BF01600502

[2] A. Ben-Israel, "A Volume Associated with $m \times n$ Matrices," *Linear Algebra and Its Applications*, Vol. 167, No. 1, pp. 87-111, 1992. doi:10.1016/0024-3795(92)90340-G

[3] X. Chen and R. E. Hartwig, "On Simultaneous Iteration for Computing the Schur Vectors of Matrices," *Proceedings of the 5th SIAM Conference on Applied Linear Algebra*, Snowbird, June 13-19, 1994, pp. 290-294.

[4] X. Chen and R. E. Hartwig, "On the Convergence of Power Scaled Cesaro Sums," *Linear Algebra and Its Applications*, Vol. 267, pp. 335-358, 1997. doi:10.1016/S0024-3795(97)80056-0

[5] X. Chen and R. E. Hartwig, "The Semi-iterative Method Applied to the Hyperpower Iteration," *Numerical Linear Algebra with Applications*, Vol. 12, No. 9, pp. 895-910, 2005. doi:10.1002/nla.429

[6] X. Chen and R. E. Hartwig, "The Picard Iteration and Its Application," *Linear and Multi-linear Algebra*, Vol. 54, No. 5, 2006, pp. 329-341. doi:10.1080/03081080500209703

[7] R. A. Horn and C. R. Johnson, "Matrix Analysis," Cambridge University Press, Cambridge, 1985.

[8] R. A. Horn and C. R. Johnson, "Topics in Matrix Analysis," Cambridge University Press, Cambridge, 1991.

[9] A. S. Householder, "The Theory of Matrices in Numerical Analysis," Dover, New York, 1964.

[10] B. N. Parlett and W. G. Poole, Jr., "A Geometric Theory for the QR, LU, and Power Iterations," *SIAM Journal on Numerical Analysis*, Vol. 10, No. 2, 1973, pp. 389-412. doi:10.1137/0710035

[11] H. Rutishauser, "Simultaneous Iteration Method for Symmetric Matrices," *Numerische Mathematik*, Vol. 16, No. 3, 1970, pp. 205-223. doi:10.1007/BF02219773

[12] Y. Saad, "Numerical Methods for Large Eigenvalue Problems," Manchester University Press, Manchester, 1992.

[13] G. W. Stewart, "Methods of Simultaneous Iteration for Calculating Eigenvectors of Matrices" In: John J. H. Miller, Eds., *Topics in Numerical Analysis II*, Academic Press, New York, 1975, pp. 185-169.

[14] G. R. Wang, Y. Wei and S. Qiao, "Generalized Inverses: Theory and Computations," Science Press, Beijing/New York, 2004.

[15] D. S. Watkins, "Understanding the QR Algorithm," *SIAM Review*, Vol. 24, No. 4, 1982, pp. 427-440. doi:10.1137/1024100

[16] D. S. Watkins, "Some Perspectives on the Eigenvalue Problem," *SIAM Review*, Vol. 35, No. 3, 1993, pp. 430-470, doi:10.1137/1035090

[17] J. H. Wilkinson, "The Algebraic Eigenvalue Problem," Oxford University Press (Clarendon), London and New York, 1964.

