*Regular Paper*

# An Acceleration Process for Iterated Vectors Generated by a Real Symmetric Matrix

Ken Iguchi †

We define an acceleration process for iterated vectors generated by a matrix iterative process. We consider here an algorithm in which the acceleration process is applied once after every $m+2$ iterations of a matrix iterative process. Then our aim is to determine the number $m$ (precisely $m+2$) required for each application of the acceleration process to be effective. Two examples are given to demonstrate the analytical results and to compare then with Jennings' acceleration process.

## 1. Introduction

We consider accelerating the convergence of the vector sequence $\{\boldsymbol{x}^{(k)}\}$ generated by a stationary iterative process of the form:

$$\boldsymbol{x}^{(k+1)}=\boldsymbol{H}\boldsymbol{x}^{(k)}+\boldsymbol{d} \qquad \text{for } k=0,\ 1,\ \cdots \tag{1.1}$$

with a given starting vector $\boldsymbol{x}^{(0)}$, where $H$ is an $n\times n$ real symmetric matrix and $\boldsymbol{d}$ is an $n$-column vector.

We consider an algorithm in which the following acceleration process[2] is applied to (1. 1) once after every $m+2$ iterations with a fixed $m$:

$$\left.\begin{aligned}\tilde{\boldsymbol{x}}&=\boldsymbol{x}^{(m+2)}+\{\lambda^2/(1-\lambda^2)\}(\boldsymbol{x}^{(m+2)}-\boldsymbol{x}^{(m)}),\\ \lambda^2&=\|\boldsymbol{x}^{(m+2)}-\boldsymbol{x}^{(m+1)}\|^2/\|\boldsymbol{x}^{(m+1)}-\boldsymbol{x}^{(m)}\|^2,\end{aligned}\right\} \tag{1.2}$$

where $\|\ \ \|$ denotes the Euclidean norm, that is, $\|\boldsymbol{y}\|^2=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}$ for vector $\boldsymbol{y}$. Then we determine the least integer $m_0$ (precisely $m_0+2$) required for each application of the acceleration process to be effective. The idea of repeating Aitken $\delta^2$ process has been used and analysed in a generalized vector version of the process by Brezinski[1] for the computation of matrix eigenvalues. Other version of this extension can be found in Ref. 4). Jennings' acceleration process for vector sequence is found in Ref. 3).

In this paper, it is only shown that the vector sequence generated by (1.1) get the rapid convergence when the acceleration process is applied after so many iterations of (1.1). If, by repeating use of the acceleration process, the rapid convergence were to be achieved, it should be considered how many of the iterations of (1.1) are needed before the acceleration process is used effectively. The acceleration process (1.2) is sophisticated. Therefore, the use of (1.2) may be restricted within some particular cases, but it seems to be worthwhile to know the way which the process (1.2) is used effectively.

Now we introduce terms and notations used in this paper. The procedure which the acceleration process (1.2) is applied once, is called a stage in our algorithm. Our algorithm can be carried out through some procedures. Therefore, at the $r$-th stage, (1.2) has been applied $r$ times. We assume that, in the $r$-th stage of our algorithm, the sequence $\{\boldsymbol{x}^{(k)}\}$ generated by (1.1) and the predicted vector $\tilde{\boldsymbol{x}}$ given by (1.2) are denoted respectively by $\{\boldsymbol{x}^{(k,r)}\}$ and $\tilde{\boldsymbol{x}}^{(r)}$. We further assume that a starting vector for (1. 1) at the $r$-th stage is denoted by $\boldsymbol{x}^{(0,r)}$, and is given by $\boldsymbol{x}^{(0,r)}=\tilde{\boldsymbol{x}}^{(r-1)}$ with $\boldsymbol{x}^{(0,1)}=\boldsymbol{x}^{(0)}$.

Then, (1. 1) and (1. 2) are written as follows:

$$\boldsymbol{x}^{(k+1,r)}=H\boldsymbol{x}^{(k,r)}+\boldsymbol{d}$$
$$\text{for } k=0,\ 1,\ \cdots,\ m+1;\ r=1,\ 2,\ \cdots \tag{1.1$'$}$$

$$\left.\begin{aligned}\tilde{\boldsymbol{x}}^{(r)}&=\boldsymbol{x}^{(m+2,r)}+\{\lambda^2/(1-\lambda^2)\}(\boldsymbol{x}^{(m+2,r)}-\boldsymbol{x}^{(m,r)})\\ \lambda^2&=\|\boldsymbol{x}^{(m+2,r)}-\boldsymbol{x}^{(m+1,r)}\|^2/\|\boldsymbol{x}^{(m+1,r)}-\boldsymbol{x}^{(m,r)}\|^2.\end{aligned}\right\} \tag{1.2$'$}$$

## 2. Convergence Property

We first investigate the convergence properties of (1.1)′ and (1.2)′. We suppose that the matrix $H$ has $n$ orthogonal eigenvectors $\boldsymbol{u}_j$ $(j=1,\ 2,\ \cdots,\ n)$ corresponding to eigenvalues $\lambda_j$ $(j=1,\ 2,\ \cdots,\ n)$ ordered such that

$$1>|\lambda_1|>|\lambda_2|\geqq|\lambda_3|\geqq\cdots\geqq|\lambda_n|. \tag{2.1}$$

Now putting $\alpha_j=\lambda_j/\lambda_1$ $(j=1,\ 2,\ \cdots,\ n)$, we have

$$\alpha_1=1 \quad\text{and}\quad 1>|\alpha_2|\geqq|\alpha_3|\geqq\cdots\geqq|\alpha_n|. \tag{2.2}$$

Suppose further that the vector sequence $\{\tilde{\boldsymbol{x}}^{(k)}\}$ converges to the vector $\boldsymbol{x}$. We define the

† Toyota College of Technology, 2-1 Eisei-Cho, Toyota 471, Japan

error vector $\boldsymbol{\varepsilon}^{(k,r)}$ of $\boldsymbol{x}^{(k,r)}$ and the error vector $\tilde{\boldsymbol{\varepsilon}}^{(r)}$ of $\tilde{\boldsymbol{x}}^{(r)}$ as follows:

$$\boldsymbol{\varepsilon}^{(k,r)}=\boldsymbol{x}-\boldsymbol{x}^{(k,r)} \quad (2.3)$$

$$\tilde{\boldsymbol{\varepsilon}}^{(r)}=\boldsymbol{x}-\tilde{\boldsymbol{x}}^{(r)}. \quad (2.4)$$

Here we derive the formula showing the convergence property of the sequence $\{\boldsymbol{x}^{(k,r)}\}$.

Since $\boldsymbol{x}=H\boldsymbol{x}+\boldsymbol{d}$, using (1.1)′ and (2.3)

$$\boldsymbol{\varepsilon}^{(k,r)}=H\boldsymbol{\varepsilon}^{(k-1,r)}=H^k\boldsymbol{\varepsilon}^{(0,r)}. \quad (2.5)$$

The error vector $\boldsymbol{\varepsilon}^{(0,r)}$ can be expressed by

$$\boldsymbol{\varepsilon}^{(0,r)}=\sum_{j=1}^{n}c_j^{(r)}\boldsymbol{u}_j \quad (2.6)$$

where the $c_j^{(r)}$ are scalar constants.

Since $H\boldsymbol{u}_j=\lambda_j\boldsymbol{u}_j$, for (2.5) we obtain after a little manipulation

$$\boldsymbol{\varepsilon}^{(k,r)}=\sum_{j=1}^{n}c_j^{(r)}\lambda_j^k\boldsymbol{u}_j. \quad (2.7)$$

This can be written as

$$\boldsymbol{\varepsilon}^{(k,r)}=\lambda_1^k\sum_{j=1}^{n}\alpha_j^k\boldsymbol{u}_j^{(r)}$$
$$\text{for } k=0, 1, \cdots, m+2;\ r=1, 2, \cdots, \quad (2.8)$$

where

$$\boldsymbol{u}_j^{(r)}=c_j^{(r)}\boldsymbol{u}_j \quad \text{for } j=1, 2, \cdots, n. \quad (2.9)$$

The (2.7) or (2.8) shows the behaviour of the error vector $\boldsymbol{\varepsilon}^{(k,r)}$. When (1.2)′ is applied to the three iterates: $\boldsymbol{x}^{(m,r)}$, $\boldsymbol{x}^{(m+1,r)}$ and $\boldsymbol{x}^{(m+2,r)}$, it is given in the following theorem how $\tilde{\boldsymbol{\varepsilon}}^{(r)}$ can be expressed.

**Theorem 1.** If (1.1)′ is iterated $m+2$ times with the starting vector $\boldsymbol{x}^{(0,r)}$, and if (1.2)′ is applied to the last three iterates, $\boldsymbol{x}^{(m,r)}$, $\boldsymbol{x}^{(m+1,r)}$ and $\boldsymbol{x}^{(m+2,r)}$, then $\tilde{\boldsymbol{\varepsilon}}^{(r)}$ and $\tilde{\boldsymbol{x}}^{(r)}$ can be expressed respectively, using

$$\theta_j^{(r)}=\|\boldsymbol{u}_j^{(r)}\|/\|\boldsymbol{u}_1^{(r)}\| \quad (j=2, 3, \cdots, n) \quad (2.10)$$

as follows:

$$\tilde{\boldsymbol{\varepsilon}}^{(r)}=\lambda_1^{m+2}\left(\frac{\sum_{i=2}^{n}P_1(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2}{1+\sum_{i=2}^{n}Q(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2}\right)\boldsymbol{u}_1^{(r)}$$
$$+\lambda_1^{m+2}\sum_{j=2}^{n}((-P_2(\lambda_1, \alpha_j, m)$$
$$+\sum_{i=2}^{n}P_3(\alpha_j;\ \lambda_1, \alpha_i, m)(\theta_i^{(r)})^2)$$
$$/(1+\sum_{i=2}^{n}Q(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2)\boldsymbol{u}_j^{(r)} \quad (2.11)$$

where

$$P_1(\lambda_1, \alpha_i, m)=\alpha_i^{2m}\frac{(1-\alpha_i^2)(1-\alpha_i\lambda_1)^2}{(1-\lambda_1^2)(1-\lambda_1)^2}$$

$$P_2(\lambda_1, \alpha_j, m)=\alpha_j^{m}\frac{(1-\alpha_j^2)}{(1-\lambda_1^2)}$$

$$P_3(\alpha_j:\ \lambda_1, \alpha_i, m)=\alpha_j^m\alpha_i^{2m}\frac{(\alpha_j^2-\alpha_i^2)(1-\alpha_i\lambda_1)^2}{(1-\lambda_1^2)(1-\lambda_1)^2}$$

$$Q(\lambda_1, \alpha_i, m)=\alpha_i^{2m}\frac{(1-\alpha_i^2\lambda_1^2)(1-\alpha_i\lambda_1)^2}{(1-\lambda_1^2)(1-\lambda_1)^2} \quad (2.12)$$

and

$$\tilde{\boldsymbol{x}}^{(r)}$$
$$=\boldsymbol{x}-\lambda_1^{m+2}\left(\frac{\sum_{i=2}^{n}P_1(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2}{1+\sum_{i=2}^{n}Q(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2}\right)\boldsymbol{u}_1^{(r)}$$
$$-\lambda_1^{m+2}\sum_{j=2}^{n}((-P_2(\lambda_1, \alpha_j, m)$$
$$+\sum_{i=2}^{n}P_3(\alpha_j:\ \lambda_1, \alpha_i, m)(\theta_i^{(r)})^2)$$
$$/(1+\sum_{i=2}^{n}Q(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2)\boldsymbol{u}_j^{(r)}. \quad (2.13)$$

**Proof.** we obtain from (1.2)' using (2.3) and (2.4)

$$\left.\begin{array}{l}\tilde{\boldsymbol{\varepsilon}}^{(r)}=\boldsymbol{\varepsilon}^{(m+2,r)}+\{\lambda^2/(1-\lambda^2)\}(\boldsymbol{\varepsilon}^{(m+2,r)}-\boldsymbol{\varepsilon}^{(m,r)})\\ \lambda_2/(1-\lambda^2)\\ =\dfrac{-\|\boldsymbol{\varepsilon}^{(m+2,r)}-\boldsymbol{\varepsilon}^{(m+1,r)}\|^2}{(\boldsymbol{\varepsilon}^{(m+2,r)}-2\boldsymbol{\varepsilon}^{(m+1,r)}+\boldsymbol{\varepsilon}^{(m,r)},\ \boldsymbol{\varepsilon}^{(m+2,r)}-\boldsymbol{\varepsilon}^{(m,r)})}\end{array}\right\} \quad (2.14)$$

denoting by ( , ) the inner product, so that $(\boldsymbol{y}, \boldsymbol{y})=\boldsymbol{y}^{\mathrm{T}}\boldsymbol{y}$ for vector $\boldsymbol{y}$.

By repeating use of (2.8), from (2.14) we obtain (2.11) with (2.12). Using (2.4), we have (2.13) from (2.11). This completes the proof.

In our algorithm, we take $\tilde{\boldsymbol{x}}^{(r)}$ as the starting vector for (1.1)′ at the $r+1$-th stage. Then we have the following results.

**Theorem 2.** If, in our algorithm, we put

$$\boldsymbol{x}^{(0,r+1)}=\tilde{\boldsymbol{x}}^{(r)} \quad \text{for } r=1, 2, \cdots \quad (2.15)$$

with $\boldsymbol{x}^{(0,1)}=\boldsymbol{x}^{(0)}$, then we have

$$\boldsymbol{u}_1^{(r+1)}=\lambda_1^{m+2}\left(\frac{\sum_{i=2}^{n}P_1(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2}{1+\sum_{i=2}^{n}Q(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2}\right)\boldsymbol{u}_1^{(r)} \quad (2.16)$$

and, for $j=2, 3, \cdots, n$

$$\boldsymbol{u}_j^{(r+1)}$$
$$=\lambda_1^{m+2}(-P_2(\lambda_1, \alpha_j, m)$$
$$+\sum_{i=2}^{n}P_3(\alpha_j;\ \lambda_1, \alpha_i, m)(\theta_i^{(r)})^2)$$
$$/(1+\sum_{i=2}^{n}Q(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2)\boldsymbol{u}_j^{(r)}. \quad (2.17)$$

Furthermore we have, for $j=2, 3, \cdots, n$, considering that all of $\theta_i^{(r)}$ is not zero:

$$\theta_j^{(r+1)}$$
$$=\frac{|-P_2(\lambda_1, \alpha_j, m)+\sum_{i=2}^{n}P_3(\alpha_j;\ \lambda_1, \alpha_i, m)(\theta_i^{(r)})^2|}{\sum_{i=2}^{n}P_1(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2}$$
$$\times\theta_j^{(r)} \quad (2.18)$$

**Proof.** Setting $k=0$ in (2.8) at the $r+1$-th stage, we have

$$\boldsymbol{\varepsilon}^{(0,r+1)}=\sum_{j=1}^{n}\boldsymbol{u}_j^{(r+1)}. \tag{2.19}$$

Subtracting each side of (2.15) from $\boldsymbol{x}$ respectively and using (2.3) and (2.4), we have

$$\boldsymbol{\varepsilon}^{(0,r+1)}=\tilde{\boldsymbol{\varepsilon}}^{(r)} \tag{2.20}$$

with $\boldsymbol{\varepsilon}^{(0,1)}=\boldsymbol{x}-\boldsymbol{x}^{(0)}$.

The right-hand side of (2.20) can be written as (2.11). The left-hand side of (2.20) is expressed as (2.8) at the $r+1$-th stage. Then considering (2.9) and matching term of $\boldsymbol{u}_j^{(r+1)}$ to term of $\boldsymbol{u}_j^{(r)}(j=1, 2, \cdots, n)$, we obtain (2.16) and (2.17).

Dividing each side of (2.17) by corresponding side of (2.16), we obtain

$$\frac{\|\boldsymbol{u}_j^{(r+1)}\|}{\|\boldsymbol{u}_1^{(r+1)}\|}=\frac{\left|-P_2(\lambda_1, \alpha_j, m)+\sum_{i=2}^{n}P_3(\alpha_j;\ \lambda_1, \alpha_i, m)(\theta_i^{(r)})^2\right|}{\sum_{i=2}^{n}P_1(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2}\times\frac{\|\boldsymbol{u}_j^{(r)}\|}{\|\boldsymbol{u}_1^{(r)}\|}, \tag{2.21}$$

from which (2.18) is obtained using (2.10). These complete the proof.

We shall consider how (1.2)′ is applied effectively. We begin by evaluating the magnitudes between the functions of (2.12).

**Theorem 3.** In our algorithm, there exists the least integer $m_0$ which satisfies

$$\frac{2}{m+2}\left(\frac{m}{m+2}\right)^{\frac{m}{2}}\frac{|\alpha_2|^{m+2}}{1-\alpha_2^2}<1. \tag{2.22}$$

If $m\geqq m_0$, it holds that, for $i=2, 3, \cdots, n$; $j=2, 3, \cdots, n$ $(i\neq j)$

$$|P_3(\alpha_j;\ \lambda_1, \alpha_i, m)|<P_1(\lambda_1, \alpha_i, m). \tag{2.23}$$

**Proof.** We define a function of real variable $\alpha$ with $|\alpha|\leqq|\alpha_2|$ by

$$f(\alpha)=|\alpha|^m\frac{|\alpha^2-\beta^2|}{1-\beta^2} \tag{2.24}$$

where $\beta$ is a real value with $|\beta|\leqq|\alpha_2|$ and $m$ is a positive number. Since $f(\alpha)$ is continuous in closed interval, then it has the maximum value. By the differential calculus, it is easily seen that $f(\alpha)$ takes on the extreme value,

$$\frac{2}{m+2}\left(\frac{m}{m+2}\right)^{\frac{m}{2}}\frac{|\beta|^{m+2}}{1-\beta^2} \tag{2.25}$$

at

$$\alpha^2=\frac{m}{m+2}\beta^2. \tag{2.26}$$

Therefore, the maximum of $f(\alpha)$ is given by

$$\max\left\{\max_{\beta}f\left(\left(\frac{m}{m+2}\right)^{\frac{1}{2}}|\beta|\right),\ \max_{\beta}f(\alpha_2)\right\}.$$

$$=\max\left\{\max_{\beta}\frac{2}{m+2}\left(\frac{m}{m+2}\right)^{\frac{m}{2}}\frac{|\beta|^{m+2}}{1-\beta^2},\right.$$

$$\left.\max_{\beta}|\alpha_2|^m\frac{\alpha_2^2-\beta^2}{1-\beta^2}\right\}$$

$$\leqq\max\left\{\frac{2}{m+2}\left(\frac{m}{m+2}\right)^{\frac{m}{2}}\frac{|\alpha_2|^{m+2}}{1-\alpha_2^2},\ |\alpha_2|^{m+2}\right\}. \tag{2.27}$$

Since $|\alpha_2|^{m+2}<1$ for all $m$, if (2.22) holds, it follows that, for all $i$ and $j$,

$$|\alpha_j^m|\frac{|\alpha_j^2-\alpha_i^2|}{1-\alpha_i^2}<1. \tag{2.28}$$

Since $\left(\frac{m}{m+2}\right)^{\frac{m}{2}}=1/\left(1+\frac{2}{m}\right)^{\frac{m}{2}}$, it is found that $\left(\frac{m}{m+2}\right)^{\frac{m}{2}}$ decreases as $m$ increases. Therefore, there exists the least integer $m_0$ which satisfy (2.22). If $m\geqq m_0$, (2.28) holds for all $i$ and $j$.

We have from (2.12)

$$P_3(\alpha_j;\ \lambda_1, \alpha_i, m)=\alpha_j^m\frac{(\alpha_j^2-\alpha_i^2)}{(1-\alpha_i^2)}P_1(\lambda_1, \alpha_i, m). \tag{2.29}$$

Taking the absolute value for the both sides of (2.29) and using (2.28), we obtain (2.23). This completes the proof.

To verify that $\|\boldsymbol{\varepsilon}^{(m,r)}\|$ vanishes as $r$ increases, it is required that the value in parentheses of the first term in (2.11) is less than unity. Hence, we have the following theorems.

**Theorem 4.** In our algorithm, it holds that, for all $\alpha_i$ and $m$,

$$P_1(\lambda_1, \alpha_i, m)<Q(\lambda_1, \alpha_i, m). \tag{2.30}$$

**Proof.** We have from (2.12)

$$P_1(\lambda_1, \alpha_i, m)=\frac{1-\alpha_i^2}{1-\alpha_i^2\lambda_1^2}Q(\lambda_1, \alpha_i, m). \tag{2.31}$$

*Since* $(1-\alpha_i^2)/(1-\alpha_i^2\lambda_1^2)<1$, we have (2.31).

**Theorem 5.** In our algorithm, it holds that, for any value of $\theta_i^{(r)}$ and all $m$,

$$\frac{\sum_{i=2}^{n}P_1(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2}{1+\sum_{i=2}^{n}Q(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2}<1. \tag{2.32}$$

**Proof.** Using Theorem 4, we have, for any $\theta_j^{(r)}$

$$\sum_{i=2}^{n}P_1(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2<\sum_{i=2}^{n}Q(\lambda_1, \alpha_i, m)(\theta_i^{(r)})^2$$

from which we obtain (2.32).

In the next theorem, we show that, as $r$ increases, all of $|\alpha_j^m|(\theta_j^{(r)})^2(j=2, 3, \cdots, n)$ get bounded in the magnitude within 0 and some constant value.

**Theorem 6.** In our algorithm, when $m\geqq m_0$, if $\theta_j^{(r)}$ satisfies

$$|P_2(\lambda_1, \alpha_j, m)|\leqq P_1(\lambda_1, \alpha_j, m)(\theta_j^{(r)})^2 \tag{2.33}$$

or

$$|\alpha_j^m|(\theta_j^{(r)})^2 \geqq (1-\lambda_1)^2/(1-\alpha_j\lambda_1)^2, \qquad (2.34)$$

then it holds that

$$\theta_j^{(r+1)} < \theta_j^{(r)} \qquad (2.35)$$

**Proof.** When $m \geqq m_0$, from Theorem 3 we obtain, for all $\alpha_j$ $(j \neq 1)$,

$$\sum_{i=2}^{n}|P_3(\alpha_j;\ \lambda_1,\ \alpha_i,\ m)|(\theta_i^{(r)})^2 < \sum_{\substack{i=2\\(i\neq j)}}^{n} P_1(\lambda_1,\ \alpha_i,\ m)(\theta_i^{(r)})^2. \qquad (2.36)$$

Hence it follows that

$$\begin{aligned}
&\left|-P_2(\lambda_1,\ \alpha_j,\ m)+\sum_{i=2}^{n}P_3(\alpha_j;\ \lambda_1,\ \alpha_i,\ m)(\theta_i^{(r)})^2\right| \\
&< |P_2(\lambda_1,\ \alpha_j,\ m)| + \sum_{i=2}^{n}|P_3(\alpha_j;\ \lambda_1,\ \alpha_i,\ m)|(\theta_i^{(r)})^2 \\
&< P_1(\lambda_1,\ \alpha_j,\ m)(\theta_j^{(r)})^2 + \sum_{\substack{i=2\\(i\neq j)}}^{n} P_1(\lambda_1,\ \alpha_i,\ m)(\theta_i^{(r)})^2 \\
&< \sum_{i=2}^{n}P_1(\lambda_1,\ \alpha_i,\ m)(\theta_i^{(r)})^2. \qquad (2.37)
\end{aligned}$$

Substituting this result into (2.20), we have (2.35). This establishes Theorem 6.

We now have the theorem concerning the convergence of the sequence $\{\boldsymbol{x}^{(m,r)}\}_{r=1}^{\infty}$.

**Theorem 7.** In our algorithm, if $m \geqq m_0$, then it holds that

$$\|\boldsymbol{\varepsilon}^{(k,r)}\| = |\lambda_1|^k \|\boldsymbol{u}_1^{(r)}\| \left(1+\sum_{j=2}^{n}\alpha_j^{2k}(\theta_j^{(r)})^2\right)^{\frac{1}{2}}. \qquad (2.38)$$

where

$$\|\boldsymbol{u}_1^{(r)}\| = |\lambda_1|^{(m+2)(r-1)} \prod_{k=1}^{r-1}\left(\frac{\sum_{i=2}^{n}P_1(\lambda_1,\ \alpha_i,\ m)(\theta_i^{(k)})^2}{1+\sum_{i=2}^{n}Q(\lambda_1,\ \alpha_i,\ m)(\theta_i^{(k)})^2}\right)|c_1^{(1)}|\|\boldsymbol{u}_1\|. \qquad (2.39)$$

Moreover, $\|\boldsymbol{\varepsilon}^{(m,r)}\|$ vanishes as $r$ increases.

**Proof.** Taking the inner product for (2.8), we obtain

$$\|\boldsymbol{\varepsilon}^{(k,r)}\|^2 = \lambda_1^{2k}\sum_{j=1}^{n}\alpha_j^{2k}\|\boldsymbol{u}_j^{(r)}\|^2. \qquad (2.40)$$

from which we obtain (2.38) using (2.10).

Using (2.16) iteratively, and using (2.9), we obtain (2.39). Next we show that the quantity $\|\boldsymbol{\varepsilon}^{(m,r)}\|$ vanishes as $r$ increases.

Substituting $k=m$ into (2.38), we have

$$\|\boldsymbol{\varepsilon}^{(m,r)}\| = |\lambda_1^m| \|\boldsymbol{u}_1^{(r)}\| \left(1+\sum_{j=2}^{n}\alpha_j^{2m}(\theta_j^{(r)})^2\right)^{\frac{1}{2}} \qquad (2.41)$$

Since, for all $\alpha_i$, and $\theta_i^{(r)}$ which satisfy (2.33), Theorem 6 holds, the magnitude of $\theta_j^{(r)}$ is bounded with increasing $r$, so that the magnitude of $\alpha_j^{2m}(\theta_j^{(r)})^2$ get bounded within some constant value. When $\theta_j^{(r)}$ does not satisfy (2.33), we have for the $\theta_j^{(r)}$ the following inequalities,

$$|P_2(\lambda_1,\ \alpha_j,\ m)| > P_1(\lambda_1,\ \alpha_j,\ m)(\theta_j^{(r)})^2, \qquad (2.42)$$

that is,

$$|\alpha_j^m|(\theta_j^{(r)})^2 < (1-\lambda_1)^2/(1-\alpha_j\lambda_1)^2. \qquad (2.43)$$

Since we have the following inequalities,

$$(1-\lambda_1)^2/(1-\alpha_j\lambda_1)^2 < \begin{cases}1 & \text{for } \lambda_1>0,\\ 4/(1-|\lambda_1|)^2 & \text{for } \lambda_1<0,\end{cases} \qquad (2.44)$$

the magnitude of $|\alpha_j^m|(\theta_j^{(r)})^2$ is bounded with increasing $r$.

Since $\|\boldsymbol{u}_1^{(r)}\|$ vanishes due to Theorem 5 as $r$ increases, $\|\varepsilon^{(m,r)}\|$ vanishes from (2.41). These complete the proof.

## 3. Practical Considerations

For the practical purpose, the total number of iterations of (1.1)′ are limited under a certain degree of accuracy. In order to get the rapid convergence, it can be seen from (2.38) and (2.39) that we should take the value $r$ as large as possible, so that the integer $m$ in (1.1)′ should be taken as small as possible.

## 4. Numerical Results

To demonstrate the analytical results, let us use (1.2)′ to accelerate the convergence of stationary iterative process with specified $H$ and $\boldsymbol{d}$. Then acceleration process (1.2)′ is applied with $m=m_0$ and the other $m$'s, and the numerical results obtained for $m=m_0$ are compared with those for the other $m$'s.

Here we mention Jennings' acceleration process to be compared with the present process (1.2)′:

$$\tilde{\boldsymbol{x}}^{(r)} = \boldsymbol{x}^{(m+2,r)} - s(\boldsymbol{x}^{(m+2,r)} - \boldsymbol{x}^{(m+1,r)}), \qquad (4.1)$$

$$\begin{aligned}s = &((\boldsymbol{x}^{(m+1,r)}-\boldsymbol{x}^{(m,r)})^{\mathrm{T}}(\boldsymbol{x}^{(m+2,r)}-\boldsymbol{x}^{(m+1,r)}))\\ &/((\boldsymbol{x}^{(m+1,r)}-\boldsymbol{x}^{(m,r)})^{\mathrm{T}}(\boldsymbol{x}^{(m+2,r)}\\ &-2\boldsymbol{x}^{(m+1,r)}+\boldsymbol{x}^{(m,r)})). \qquad (4.2)\end{aligned}$$

The whole computations were carried out in the single-precision arithmetics.

**Example 1.** The $H$ is 15×15 real symmetric matrix and the $\boldsymbol{d}$ is 15-column vector as fol-

**Table 1** Sample data of the least iteration number $m_0$ (precisely $m_0+2$) of stationary iterative process (1.1)′ required before applying the acceleration process (1.2)′.

| $\lvert\alpha_2\rvert$ | 0.840 | 0.910 | 0.932 | 0.945 | 0.954 | 0.960 | 0.970 | 0.985 | 0.990 | 0.995 |
|---|---|---|---|---|---|---|---|---|---|---|
| $m_0$ | 1 | 2 | 3 | 4 | 5 | 6 | 9 | 18 | 27 | 55 |

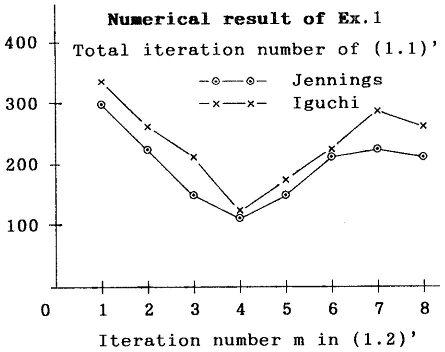



**Fig. 1** Total iteration number of stationary iterative process (1.1)′ to get the five correct digits to the correct value when acceleration process (1.2)′ or (4.1) is applied once every after (1.1)′ is iterated $m+2$ times with specified value as $m$.

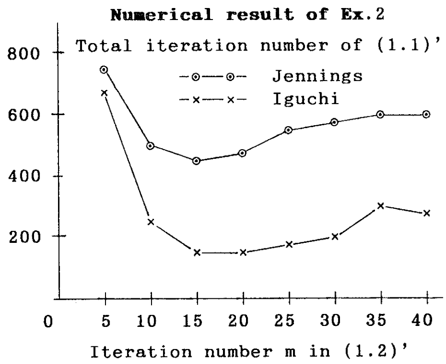



**Fig. 2** Total iteration number of stationary iterative process (1.1)′ to get the five correct digits to the correct value when acceleration process (1.2)′ or (4.1) is applied once every after (1.1)′ is iterated $m+2$ times with specified value as $m$.

lows:

$$H=\begin{Bmatrix} 0.315 & -0.252 & 0.063 & & & & \\ & 0.378 & -0.252 & 0.063 & & \mathbf{0} & \\ & & 0.378 & \ddots & \ddots & & \\ & & & \ddots & \ddots & 0.063 & \\ \text{symmetric} & & & & 0.378 & -0.252 & 0.063 \\ & & & & & 0.378 & -0.252 \\ & & & & & & 0.315 \end{Bmatrix},$$

$$\boldsymbol{d}=(0.2,\ 0.2,\ \cdots,\ 0.2)^{\mathrm{T}}.$$

The dominant eigenvalue of $H$ is $\lambda_1=0.989$ and the dominant ratio of the eigenvalues to $\lambda_1$ is $\alpha_2=0.943$, so that we get $m_0=4$ from (2.22). The numerical results are shown in **Fig. 1**.

**Example 2.** The $H$ is $4\times 4$ symmetric matrix and the $\boldsymbol{d}$ is 4-column vector as follows:

$$H=\begin{bmatrix} 0.248 & 0.124 & 0.372 & 0.496 \\ & -0.372 & 0.124 & 0.620 \\ \text{symmetric} & & 0.744 & -0.248 \\ & & & -0.124 \end{bmatrix},$$

$$\boldsymbol{d}=\begin{bmatrix} 0.01 \\ 0.01 \\ 0.01 \\ 0.01 \end{bmatrix}.$$

The dominant eigenvalue of $H$ is $\lambda_1=-0.996$ and the ratios of the other eigenvalues to $\lambda_1$ are $\alpha_2=-0.988$, $\alpha_3=-0.583$ and $\alpha_4=0.196$, so that we get $m_0=24$ from (2.22). The numerical results are shown in **Fig. 2**.

For the both examples, a starting vector $\boldsymbol{x}_0$ was taken as

$$\boldsymbol{x}_0=(1.0,\ 1.0,\ \cdots,\ 1.0)^{\mathrm{T}}.$$

The value $\alpha_i$ 's in the example 1 are all positive, and some of the value $\alpha_i$ 's in the example 2 are negative.

The numbers of iterations of (1.1)′ were counted until the significant digits become correct to five decimals. When the process (1.1)′ for the examples 1 and 2 was iterated without acceleration until the same accuracy was achieved, the numbers of the iterations of (1.1)′ were respectively 1328 and 2292. It can be seen from Fig. 1 and Fig. 2 that the minimum total numbers of the iterations of (1.1)′ are taken in the vicinity of $m=m_0$. The justification of our algorithm depends on the way of the distribution of the eigenvalues of $H$. Our numerical results for two examples show the identification to our analytical results.

## 5. Conclusions

We consider an algorithm in which the acceleration process (1.2)′ is applied once after every $m+2$ times iterations of (1.1)′. Then if we take the number $m+2$ larger than or equal to $m_0+2$, it is warranted that the vector sequence $\{\boldsymbol{x}^{(m,r)}\}_{r=1}^{\infty}$ generated by (1.1)′ converges. If we take $m$ equal or near to $m_0$, by the application of our algorithm, the total number of iterations of (1.1)′ will be reduced to about one-tenth of that required without acceleration.

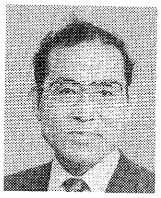
The value $m_0$'s obtained by our analysis will give the good estimations of their exact values when some of the eigenvalues of $H$ in (1.1)′ are distributed in the neighborhood of the value $\left(\frac{m_0}{m_0+2}\right)^{\frac{1}{2}}|\alpha_2|$. The acceleration process (1.2)′ is applied effectively in our algorithm when $\alpha_i<0$ as well as when $\alpha_i \geqq 0$. However, when Jennings' acceleration process (4.1) instead of the process (1.2)′ was applied in our algorithm, it was found by similar analysis as (1.2)′ that the process (4.1) had good convergence property only when $\alpha_i \geqq 0$, $i=2, 3, \cdots, n$.

**Acknowledgement** I would like to thank the anonymous refrees for their valuable comments and suggestions.

## References

1) Brezinski, C.: Computation of the Eigenelements of a Matrix by the ε-Algorithm, *Linear Algebra Applics*, Vol. 11, pp. 7-20 (1975).
2) Iguchi, K.: Convergence Property of Aitken $\delta^2$ Process and the Applicable Acceleration Process, *J. Inf. Process.*, Vol. 7, pp. 22-30 (1984).
3) Jennings, A.: Accelerating the Convergence of Matrix Iterative Processes, *J. Inst. Maths. Applics*, Vol. 8, pp. 99-110 (1971).
4) Tan, R. C.: Implementation of the Topological ε-Algorithm, *SIAM. J. Sci. Stat. Comp.*, Vol. 9, pp. 839-848 (1988).

(Received August 19, 1994)
(Accepted November 17, 1994)



**Ken Iguchi** was born in 1937. He received the B. E. degree from Nagoya Institute of Technology, the M. E. and Dr. Eng. degrees from Nagoya University in 1966 and 1985, respectively. He is now a professor in the department of information and computer engineering, Toyota National College of Technology, Toyota, Japan. His research interests include numerical analysis, operations research, statistics. He is a member of the IPSJ and JSIAM.