<table>
<thead>
<tr>
<th>項目</th>
<th>内容</th>
</tr>
</thead>
<tbody>
<tr>
<td>タイトル</td>
<td>An Adaptive Acceleration of General Linear Iterative Processes for Solving Systems of Linear Equations (大型の数値計算に関する諸問題)</td>
</tr>
<tr>
<td>著者</td>
<td>TANABE, KUNIO</td>
</tr>
<tr>
<td>引用</td>
<td>数理解析研究所講究録 (1971), 129: 100-118</td>
</tr>
<tr>
<td>発行日</td>
<td>1971-11</td>
</tr>
<tr>
<td>URL</td>
<td><a href="http://hdl.handle.net/2433/106557">http://hdl.handle.net/2433/106557</a></td>
</tr>
<tr>
<td>タイプ</td>
<td>Departmental Bulletin Paper</td>
</tr>
<tr>
<td>テキストバージョン</td>
<td>publisher</td>
</tr>
</tbody>
</table>

京都大学学術情報リポジトリ

京都大学
An Adaptive Acceleration of General Linear Iterative Processes for Solving Systems of Linear Equations

Kunio TANABE
The Institute of Statistical Mathematics
(Submitted to Numerische Mathematik)

Summary
An adaptive acceleration of general linear iterative processes is devised for solving singular and nonsingular systems of linear equations. At each step of the procedure, an acceleration parameter is controlled from information available in the current two iterations of the original process to effect nonlinear feedback to the process which causes 'attraction' of virtual sequences, similar to the phenomenon suggested by Wiener, and produces extremely rapid convergence of the accelerated process. The method can be efficiently applied to almost all of linear iterative processes. The relations to generalized inverses are discussed.
1. Introduction

The purpose of this paper is to give an adaptive acceleration technique of general iterative processes for solving a system of \( m \) linear equations in \( n \) unknowns, written

\[ Ax = b, \]

where \( A \) is a \( m \times n \) complex matrix, \( x \) and \( b \) are \( n \) and \( m \) dimensional complex column vectors respectively.

In many linear iterative processes, matrix equation (1) is reduced to the analogous equation

\[ x = Qx + Rb, \]

where the \( n \times n \) convergent matrix \( Q \) and \( n \times m \) matrix \( R \) depend only on \( A \) and satisfy the relation

\[ Q + RA = I \]

\[ \text{Im}A \cap \text{Ker}R = \{ 0 \}, \]

then an stationary linear iterative formula

\[ x^{i+1} = \varphi(x^i, b) \equiv Qx^i + Rb, \quad i=0,1,\ldots. \]

is obtained. We have many iterative processes of form (4); named Gauss-Seidel, Jacobi, Extrapolated Seidel, Extrapolated Jacobi, SOR, ADI, Block SOR, etc. for nonsingular matrix \( A \) [2][6], and Weak Steepest Descent [5], Kaczmarz [4], etc. for possibly singular \( A \). Different processes are distinguished by the choice of \( Q \) and \( R \). Nonstationary (cyclic) processes such as Richardson method can be considered as a kind of stationary process (4)
if their one cycle of acceleration process is considered as a single step.

The iterative process (4) generates a convergent sequence \( \{x^i\} \), whose limit point is a solution of (1) in some sense, more precisely, we have

\[
\lim_{i \to \infty} x^i = P_{\text{Ker}A} x^0 + A^- b
\]

\[
A^- = (I - \bar{Q})^{-1} R,
\]

where \( x^0 \) is an initial vector, \( P_{\text{Ker}A} \) is the orthogonal projection onto \( \text{Ker}A \) and \( \bar{Q} \) is the restriction of \( Q \) onto \( \text{Im}A^* \). Note that the matrix \( A^- \) is a generalized inverse of minimum norm reflexive type [3], i.e. \( AA^- A = A, A^- AA^- = A^- \) and \( (A^- A)^* = A^- A \).

2 Acceleration procedure

In this section we develop an acceleration procedure to improve the convergence of the process (4) with nonnegative definite Hermitian matrix \( Q \).

If we know all the eigenvalues of \( Q \), acceleration can be executed perfectly. But they are ordinarily unknown. Then in ordinary acceleration process such as Richardson method, enclosure bound of the eigenvalues of \( Q \) is estimated and acceleration parameters are chosen by the knowledge of Chebyshev polynomial on this interval [2][6]. These acceleration methods adopt minimax criterion in the choice of acceleration parameters, that is, this choice is made, regardless of current value \( x \),
so as to minimize the spectral radius of the basic iteration matrix of the accelerated process.

Contrarily, the method proposed below uses current information of \( x \)'s, that is, at each step of the procedure, an acceleration parameter is controlled from information available in the current two iterations of the original process to effect nonlinear feedback which causes 'attraction' of virtual sequences, similar to the phenomenon suggested by Wiener[7], and produces extremely rapid convergence of the accelerated process.

Given a linear iterative process (4) with nonnegative definite Hermitian matrix \( Q \), we consider the following mapping \( \phi_\alpha \) from \( C^n \) to \( C^n \) for any real \( \alpha \) defined as

\[
(6) \quad \phi_\alpha(x, b) = x + \alpha (\phi(x, b) - x) = Q_\alpha x + R_\alpha b,
\]

where \( Q_\alpha = (1-\alpha)I + \alpha Q \) and \( R_\alpha = \alpha R \).

\( \phi_\alpha \) is again of the form (4) since \( Q_\alpha + R_\alpha A = I \). In the following, \( \phi(x, b) \), \( \phi(x, 0) \), \( \phi_\alpha(x, b) \), \( \phi_\alpha(x, 0) \) will be abbreviated as \( \phi(x) \), \( \phi^0(x) \), \( \phi_\alpha(x) \), \( \phi^0_\alpha(x) \) respectively.

**DEFINITION:** Given a sequence of real numbers \( \{\alpha_i\}_{i \in \mathbb{N}} \) and an initial vector \( x_{80} \), we consider the following virtual sequence \( \{x_{ij}\}_{i, j \in \mathbb{N}} \), of which only diagonal and lower subdiagonal elements are really computed in the acceleration, defined as

\[
\begin{align*}
x_{i+1, j} &= \phi(x_{i, j}) = Qx_{i, j} + Rb \\
x_{i, j+1} &= \phi_\alpha(x_{i, j}) = Q_\alpha x_{i, j} + R_\alpha b
\end{align*}
\]
where \( i, j \in \mathbb{N} = \{0, 1, 2, \ldots, \infty\} \).

The sequence is well defined since \( \phi_k \) and \( \phi \) are commutative.

Note that the first column sequence \( \{x_{i,0}\}_{i \in \mathbb{N}} \) is generated by the original iteration (4) and that

\[
\lim_{k \to \infty} x_{i,k} = \lim_{k \to \infty} x_{i,0} \quad k \in \mathbb{N}.
\]

**DEFINITION:** Let sequences \( \{d_{i,j}\}_{i,j \in \mathbb{N}} \), \( \{e_{i,j}\}_{i,j \in \mathbb{N}} \) of column vectors defined as

\[
d_{i,j} = x_{i,j+1} - x_{i,j}, \\
e_{i,j} = x_{i+1,j} - x_{i,j}, \quad i, j \in \mathbb{N}.
\]

Then we have

\[
e_{i+1,j} = \phi^*(e_{i,j}) = Qe_{i,j} \\
e_{i,j+1} = \phi_{\alpha_j}(e_{i,j}) = Q_{\alpha_j} e_{i,j}, \quad i, j \in \mathbb{N}.
\]

The following lemma is easily deduced from (7) and (9).

**LEMMA 1:** \( e_{i,j} \in (ES(0) + ES(1))^L \) for \( i \in \mathbb{N}^+, \ j \in \mathbb{N} \),

where \( \mathbb{N}^+ = \mathbb{N} - \{0\} \), \( ES(\lambda) = \{ x ; Qx = \lambda x \} \) and \((*)^T\) denotes the orthogonal complement of \(*\).

Now we introduce an acceleration algorithm.

**ACCELERATION ALGORITHM:** Given an initial vector \( x_\infty \), determine the sequence \( \{x_{i,i}\}_{i \in \mathbb{N}} \) of vectors from the recurrence formula

\[
x_{i+1,i} = \phi(x_{i,i}) \\
e_{i,i} = x_{i+1,i} - x_{i,i} \\
e_{i+1,i} = \phi^*(e_{i,i}) \quad \text{for} \ i \in \mathbb{N},
\]

\[
a_i = \frac{\|e_{i,i}\|^2 - \langle e_{i,i}, e_{i+1,i} \rangle}{\|e_{i+1,i} - e_{i,i}\|^2} \\
x_{i+1,i+1} = x_{i+1,i} + a_i e_{i+1,i}
\]
where only diagonal and lower subdiagonal elements of the sequences \( \{ x_{ij} \}_{i,j} \in N \) and \( \{ e_{ij} \}_{i,j} \in N \) appear and if \( e_{i_0} = 0 \) for some \( i_0 \in N \), then the algorithm is terminated yielding a solution but generally the algorithm does not stop in finite steps.

The choice of the acceleration parameter \( a_i \) on the \( i \)-th stage of the algorithm (10) is made so as to minimize the function

\[
 f_i(a) \equiv \| e_{i,i+1} \|^2 = \| Q_a e_{i,i} \|^2,
\]

that is, \( a_i \) is so controlled that the two neighbouring row sequences come as close as possible on the \((i, i+1)\) stage.

Let \( \rho_i \) and \( \sigma_i \) be defined as

\[
\begin{align*}
\rho_i & \equiv \frac{\langle e_{i,i}, e_{i,i+1} \rangle}{\| e_{i,i} \|^2} \\
\sigma_i & \equiv \frac{\| e_{i,i+1} \|^2}{\| e_{i,i} \|^2}
\end{align*}
\]

where \( e_{i,i} \neq 0 \) is assumed, then we have

\[
 f_i(a) = \| e_{i,i} \|^2 (1 - 2(1 - \rho_i)a + (1 - 2\rho_i + \sigma_i)a^2).
\]

Thus the minimum

\[
 f_i(a_i) = \| e_{i,i} \|^2 \frac{\sigma_i - \rho_i^2}{1 - 2\rho_i + \sigma_i}
\]

is attained at

\[
 a_i = \frac{1 - \rho_i}{1 - 2\rho_i + \sigma_i}.
\]

Let \( \lambda_i \)'s are the eigenvalues of \( Q \) in the open interval \((0, 1)\) ordered as
\[ 0 < \lambda_1 < \lambda_2 < \cdots < \lambda_r < 1, \]

where \( \lambda_i \) and \( \lambda_r \) will also be denoted by \( \lambda \) and \( \Lambda \) respectively. Corresponding to the direct sum decomposition of \((ES(0) + ES(1))^\perp\)

\[
(ES(0) \oplus ES(1))^\perp = ES(\lambda_1) \oplus \cdots \oplus ES(\lambda_r),
\]

we have the unique decomposition of \( e_{ij} \)

\[
e_{ij} = e_i^1 \oplus e_i^2 \oplus \cdots \oplus e_i^r,
\]

\[ e_i^j \in ES(\lambda_j), \quad i \in \mathbb{N}^+. \]

Then \( \rho_i \) and \( \sigma_i \) are convex combination of \( \lambda_j \) and \( \lambda_i^2 \) respectively, i.e.

\[
\rho_i = \frac{\sum \lambda_j \| e_i^j \|^2}{\sum \lambda_j \| e_i^j \|^2}, \quad \sigma_i = \frac{\sum \lambda_i^2 \| e_i^j \|^2}{\sum \lambda_i^2 \| e_i^j \|^2}.
\]

Let the points \((\lambda_j, \lambda_i^2)\) in the \((\rho, \sigma)\)-plane be denoted by \( P_j \), and let the convex hull of the points \( P_1, P_2, \ldots, P_r \) be denoted by \( CH(P_1, \ldots, P_r) \) (or abbr. CH). The equation (17) means

\[
(\rho_i, \sigma_i) \in CH(P_1, \ldots, P_r).
\]

Define functions \( g \) and \( h \) as

\[
g(\rho, \sigma) \equiv \frac{1 - \rho}{1 - 2\rho + \sigma}, \quad h(\rho, \sigma) \equiv \frac{\sigma - \rho^2}{1 - 2\rho + \sigma}.
\]

Elaborate calculation shows

\[
\begin{align*}
\min_{(\rho, \sigma) \in CH} g(\rho, \sigma) &= g(\lambda, \lambda^2) = \frac{1}{1 - \lambda}, \\
\max_{(\rho, \sigma) \in CH} g(\rho, \sigma) &= g(\Lambda, \Lambda^2) = \frac{1}{1 - \Lambda},
\end{align*}
\]

7
Min \( h(\rho, \sigma) = h(\lambda_i, \lambda_i^2) = 0 \), \( i \in \{1, 2, \ldots, r\} \),

\[
\begin{align*}
\text{Max} \ h(\rho, \sigma) &= h(u, v) = \frac{(\Lambda - \lambda)^2}{(2 - \Lambda - \lambda)^2}, \quad \text{where} \\
u &= \frac{\Lambda + \lambda - 2 \Lambda \lambda}{2 - \Lambda - \lambda}, \quad v = \frac{\Lambda^2 (1 - \lambda) + \lambda^2 (1 - \Lambda)}{2 - \Lambda - \lambda}.
\end{align*}
\]

Thus we have the following lemma.

**Lemma 2:** \( 1 < \frac{1}{1 - \lambda} \leq \alpha_i \leq \frac{1}{1 - \lambda} \),

\[
0 \leq \delta_i \leq \frac{\Lambda - \lambda}{2 - \Lambda - \lambda} < \Lambda,
\]

where \( \delta_i = \sqrt{h(\rho, \sigma)} \), and \( \frac{\Lambda - \lambda}{2 - \Lambda - \lambda} \) decreases monotonically from \( \frac{\Lambda}{2 - \Lambda} \) to 0 as \( \lambda \) varies from 0 to \( \Lambda \).

Now we prove the convergence of the algorithm (10).

**Theorem 1:** The acceleration algorithm generates a convergent sequence \( \{x_{i+1}\}_{i \in N} \) whose limit point coincides with that of the original process (4), i.e.

\[
\lim_{i \to \infty} x_{i+1} = \lim_{i \to \infty} x_{i+1} = P_{\text{Ker} A^*} x_{00} + (I - Q) R b.
\]

**Proof** Since \( \text{Im} A^* = \text{ES}(1)^+ \), we have

\[
e_{i+1, i+1} = Q e_{i+1, i+1} = Q e_{i+1, i+1} = Q e_{i+1, i+1}
\]

then \( \|e_{i+1, i+1}\| = \|Q\| Q_{i+1, i+1} \leq \Lambda \delta_i \|e_{i+1}\| \leq \Lambda^i \left( \prod_{j=0}^{i} \delta_j \right) \|e_{0, i}\| \).

Since \( d_{i+1, i} = \alpha_i e_{i+1, i} = \alpha_i \bar{Q} e_{i+1, i} \), we have

\[
\|d_{i+1, i}\| \leq \alpha_i \Lambda \|e_{i+1}\|.
\]

Thus

\[
\|x_{i+1, i+1} - x_{i+1}| \leq \|d_{i+1, i}\| + \|e_{i+1, i+1}\| \leq \Lambda (\alpha_i + \delta_i) \|e_{i+1}\|
\]

\[
\leq \Lambda (\alpha_i + \delta_i) \|e_{i+1}\|.
\]
where $\delta_i < \Lambda$ and $K \equiv \frac{1}{1 - \frac{1}{\Lambda}} + \Lambda > \alpha_i + \delta_i$.

Since the right hand side of the above inequality is summable, the sequence $\{x_{ii}\}$ is convergent. Since

$$\lim_{i \to \infty} e_{ii} = \lim_{i \to \infty} (\phi(x_{ii}) - x_{ii}) = 0,$$

then $\phi(\bar{x}, b) = \bar{x}$, where $\bar{x} = \lim_{i \to \infty} x_{ii}$. Noting that the component $P_{\text{Ker}A}x_{\infty}$ is invariant under the transformations $\phi$ and $\phi_{\alpha}$'s we have the desired result.

**Theorem 2:**

$$\|x_{ii} - \bar{x}\| \leq \frac{\Lambda^i}{1 - \frac{1}{\Lambda}} \|x_{\infty}\|,$$

where $\bar{x} = P_{\text{Ker}A}x_{\infty} + (I - \bar{Q})^tRb$.

**Proof:** It is easily seen from (3) that

$$\text{Ker}A = \text{ES}(1) \text{ and } \text{Im}A^* \subset \text{Im}A^*,$$

then we have $\text{RA}A^* = RA(I - \bar{Q})^tR = (I - Q)(I - \bar{Q})^tR = R$.

We have

$$\text{Im}A^* \ni x_{ii} - \bar{x} = \bar{x}_{ii} - A^*b$$

$$= (I - \bar{Q})^t(I - Q)(x_{ii} - A^*b)$$

$$= (I - \bar{Q})^tRA(\bar{x}_{ii} - A^*b)$$

$$= (I - \bar{Q})^t(RAx_{ii} - Rb)$$

$$= (I - \bar{Q})^t((I - Q)x_{ii} - Rb)$$

$$= -(I - \bar{Q})^t e_{ii},$$
where $\bar{x}_{ii} = x_{ii} - P_{\text{Ker}A} x_{ii}$. Then we have
\[
\| x_{ii} - \bar{x} \| \leq \frac{1}{1 - \Lambda} \| e_{ii} \| \leq \frac{\Lambda^i}{1 - \Lambda} \left( \sum_{j=0}^{\infty} \delta_j \right) \| e_{oo} \| .
\]
This completes the proof.

Although the behaviour of $(\rho_i, \sigma_i)$ is very complicated in general, we have the following theorem.

**THEOREM 3:** If $\lambda_1 + \lambda_2 > \lambda_\rho$, then
\[
\lim_{i \to \infty} (\rho_{2i}, \sigma_{2i}) = (\rho_{\text{odd}}, \sigma_{\text{odd}}) = P_{\text{odd}}
\]
\[
\lim_{i \to \infty} (\rho_{2i-1}, \sigma_{2i-1}) = (\rho_{\text{even}}, \sigma_{\text{even}}) = P_{\text{even}}
\]
\[
\frac{P_i P_{\text{odd}}}{P_{\text{odd}} P} \frac{P_i P_{\text{even}}}{P_{\text{even}} P} = \xi
\]
where
\[
\cdot \xi \text{ or } 1/\xi = \frac{\lambda^2 (1 - \Lambda)^2}{\Lambda^2 (1 - \lambda)^2} \neq 1.
\]
Either $P_{\text{odd}} \in \text{LS}^+$, $P_{\text{even}} \in \text{LS}^-$ or $P_{\text{even}} \in \text{LS}^+$, $P_{\text{odd}} \in \text{LS}^-$ occurs,
where
\[
\text{LS}^+ = \left\{ (\rho, \sigma) \in P_i P_r ; \Delta(\rho, \sigma) > 0 \right\},
\]
\[
\text{LS}^- = \left\{ (\rho, \sigma) \in P_i P_r ; \Delta(\rho, \sigma) < 0 \right\},
\]
\[
\Delta(\rho, \sigma) = H(\rho, \sigma, \lambda) + H(\rho, \sigma, \Lambda)
\]
\[
H(\rho, \sigma, \lambda) = \frac{1}{1 - 2\rho + \sigma} \left\{ (\sigma - \rho) \lambda^i + (1 - \rho) \lambda^2 \right\}.
\]
and $ST$ denotes the length of the line segment $ST$ between two points $S$ and $T$.

**Proof:** Since $e_{ii,ii} = QQ_{\alpha_i} e_{ii}$, we have
\[ e_{i+1}^\prime = \eta_i^\prime e_i^\prime \]
\[ \eta_i^\prime = \lambda_j (1 - \alpha_i (1 - \lambda_j)) = H(\rho_i^\prime, \sigma_i^\prime, \lambda_j). \]

Since \( H(\rho, \sigma, \lambda) \) is a quadratic function in \( \lambda \), it attains its minimum at
\[ \lambda = \frac{\rho - \sigma}{2(1 - \rho)} = \Theta(\rho, \sigma). \]

Since
\[ \Theta(\rho_i^\prime, \sigma_i^\prime) = \frac{\frac{\lambda_j}{2} \lambda_j (1 - \lambda_j) \|e_i^\prime\|^2}{\frac{1}{1 - \lambda_j} \|e_i^\prime\|^2}, \; i \in N^+ \]
we have \( \lambda/2 = \Theta(\rho_i^\prime, \sigma_i^\prime) \leq \lambda/2 \). Since
\[ H(\rho_i^\prime, \sigma_i^\prime, \lambda_j) = \frac{\frac{\lambda_j}{2} \lambda_j (\lambda_j - \lambda_j) (1 - \lambda_j) \|e_i^\prime\|^2}{\frac{1}{1 - \lambda_j} \|e_i^\prime\|^2}, \; i \in N^+ \]
we have for \( i \in N^+ \)
\[ (23) \quad \eta_i^1 \leq 0 \leq \eta_i^r. \]

We have
\[ \lambda + 2(\Theta(\rho, \sigma) - \lambda) = \frac{\frac{\lambda_j}{2} (\lambda_j - \lambda) (1 - \lambda_j) \|e_i^\prime\|^2}{\frac{1}{1 - \lambda_j} \|e_i^\prime\|^2} \leq \lambda - \lambda, \; i \in N^+ \]
then if \( \lambda - \lambda < \lambda_2 \),
\[ (24) \quad H(\rho_i^\prime, \sigma_i^\prime, \lambda_i) < H(\rho_i^\prime, \sigma_i^\prime, \lambda_i) < \cdots < H(\rho_i^\prime, \sigma_i^\prime, \lambda_i), \]
for \( i \in N^+ \) (See Fig. 5-(a)). i.e.
\[ (25) \quad \eta_i^1 < \eta_i^2 < \cdots < \eta_i^r, \; \text{for } i \in N^+. \]

We have also (25) if \( 2\lambda \geq \lambda \) (See Fig. 5-(b).) which is implied in the assumption of (24).
Noting (23) and (25), we have, as $i \to \infty$

$$e_i \sim e_i^1 + e_i^\prime,$$

(26) $$p_i \sim \frac{\lambda \|e_i^1\|^2 + \Lambda \|e_i^\prime\|^2}{\|e_i^1\|^2 + \|e_i^\prime\|^2},$$

$$\sigma_i \sim \frac{\lambda^2 \|e_i^1\|^2 + \Lambda^2 \|e_i^\prime\|^2}{\|e_i^1\|^2 + \|e_i^\prime\|^2}.$$

Then we have

$$\eta_i^1 \sim \lambda(\lambda - \Lambda)(1 - \Lambda) \|e_i^\prime\|^2 / L$$

$$\eta_i^\prime \sim \Lambda(\Lambda - \lambda)(1 - \lambda) \|e_i^1\|^2 / L$$

where $L = (1 - \lambda)^2 \|e_i^1\|^2 + (1 - \Lambda)^2 \|e_i^\prime\|^2$, then

$$\frac{\|e_i^{1+}\|^2}{\|e_i^1\|^2} \sim \left(\frac{\eta_i^1}{\eta_i^\prime}\right)^2 \frac{\|e_i^1\|^2}{\|e_i^\prime\|^2}$$

(27) $$= \frac{\lambda^2(1 - \Lambda)^2 \|e_i^\prime\|^2}{\Lambda^2(1 - \lambda)^2 \|e_i^1\|^2},$$

and then

$$\frac{\|e_i^{1+}\|^2}{\|e_i^1\|^2} \sim \frac{\|e_i^{1-}\|^2}{\|e_i^\prime\|^2}.$$

Thus we have the first half of the theorem.

If there exists an integer $i^\star$ such that $P_i P_{i^\star}$ and the line defined by $\Delta(\rho, \sigma)$ intersect at $(\rho_{i^\star}, \sigma_{i^\star})$, i.e.

$$\rho_{i^\star} = \frac{\Lambda^2(1 - \Lambda) + \lambda^2(1 - \Lambda)}{\Lambda + \lambda - 2\Lambda \lambda}$$

(28) $$\sigma_{i^\star} = \frac{\Lambda^3(1 - \lambda) + \lambda^3(1 - \Lambda)}{\Lambda + \lambda - 2\Lambda \lambda}$$

then it is easily seen that
(29) \(( \rho_i, \sigma_i ) = ( \rho_{x_i}, \sigma_{x_i} )\), for any \( i \in \mathbb{N}^* \),

that is, The point \(( \rho_{x_i}, \sigma_{x_i} )\) is a stationary point of the discrete dynamical system defined by (10) and (12). Generally we have

\(( \rho_i, \sigma_i ) \neq ( \rho_{x_i}, \sigma_{x_i} )\).

Assume that both \( P_{\text{odd}} \) and \( P_{\text{even}} \) are in the same line segment \( LS^+ \) or \( LS^- \), the we have either

\[ 0 \leq -H(\rho_i, \sigma_i, \lambda_1) < H(\rho_i, \sigma_i, \lambda_0) \]

or

\[ 0 \leq H(\rho_i, \sigma_i, \lambda_0) < -H(\rho_i, \sigma_i, \lambda_1), \]

which means \( \lim_{{i \to \infty}} (\rho_i, \sigma_i) = P_r \) or \( P_1 \), contradicting (27).

Thus \( P_{\text{odd}} \) and \( P_{\text{even}} \) are separated by the line \( \Delta(\rho, \sigma) = 0 \), and \(( \rho_{x_i}, \sigma_{x_i} )\) is an unstable stationary point of (10) and (12).

This completes the proof.

For solving (1), 'Back and Forth' technique often yields iterative processes with symmetric iteration matrices \( Q \) [2][6]. If two iterations of the process with nondefinite symmetric matrix \( Q \) are taken as a single step, the resulting iteration has a nonnegative definite iteration matrix. So our acceleration algorithm is applicable to most of linear iterative processes. Moreover it is applicable to already accelerated processes such as Extrapolated Jacobi, SOR and Richardson methods.
3. Practical Consideration

In practical computations, it often happens that some leading digits of $x_i$ and $x_{i+1}$ coincide with one another, then cancelation occurs in the computation of $e_{ii}$. But this is remedied, without destroying the acceleration schema (10), by transferring $x$ to another memory locations and by continuing the original iterations (4) in the schema (10) with new right hand side $b' = b - Ax$, just as in the iterative improvement process.

We have generally

$$\langle e_{ii}, e_{i+1i} \rangle^2 \leq \| e_{ii} \|^2 \| e_{i+1i} \|^2.$$  

But the computation with finite precision violates occasionally this inequality resulting $\alpha_i < 0$, when $\Lambda$ is close to 1. In this case we put

$$\alpha_i = \frac{\| e_{ii} \|^2}{\| e_{ii} \|^2 - \langle e_{ii}, e_{i+1i} \rangle},$$

instead of the fourth equality of the algorithm (10), noting that

$$\alpha_i = \frac{1}{1 - \rho_i}, \text{ when } \alpha_i = \rho_i^2.$$  

4. Acknowledgement

The author would like to thank H. Akaike, Inst. Statist. Math. and M. Sibuya, Scientific Center, IBM Japan for their encouragement during the preparation of the paper.
Fig. 1

Fig. 2
Fig. 3. Block Diagram of Algorithm (10).

Fig. 4. Polygon CH
REFERENCES


CORRECTIONS FOR
"An adaptive acceleration of general linear iterative processes for solving systems of linear equations"
Kunio TANABE

(b: From the bottom)

<table>
<thead>
<tr>
<th>Page</th>
<th>Line</th>
<th>Old</th>
<th>New</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>9</td>
<td>$\bar{Q}$ is the restriction of $Q$ onto $\text{Im}A^*$.</td>
<td>$\bar{Q} = QP_{\text{Im}A^<em>}$, $P_{\text{Im}A^</em>} = I - P_{\text{Ker}A^*}$.</td>
</tr>
<tr>
<td>5</td>
<td>b2</td>
<td>$\frac{|e_{i\cdot i}|^2 - \langle e_{i\cdot i}, e_{i\cdot n} \rangle}{|e_{i\cdot n}|-e_{i\cdot i}|^2}$</td>
<td>$\langle e_{i\cdot i}, e_{i\cdot i} - e_{i\cdot n} \rangle$</td>
</tr>
<tr>
<td>7</td>
<td>b7</td>
<td>$(p_i, \sigma_i) \in \text{CH}(p, \ldots, pr)$.</td>
<td>$(p_i, \sigma_i) \in \text{CH}(p, \ldots, pr)$ for $i \in \mathbb{N}^+$.</td>
</tr>
<tr>
<td>8</td>
<td>b2</td>
<td>$|d_{i\cdot n\cdot i\cdot n}| + |e_{i\cdot n\cdot i\cdot i}|$.</td>
<td>$|d_{i\cdot n\cdot i\cdot i}| + |e_{i\cdot i}|$.</td>
</tr>
<tr>
<td>8</td>
<td>b1</td>
<td>$\Lambda (\alpha_i + \delta_i) |e_{i\cdot i}|$.</td>
<td>$(\alpha_i \Lambda + 1) |e_{i\cdot i}|$.</td>
</tr>
<tr>
<td>9</td>
<td>1</td>
<td>$\Lambda_i \left( \sum_{j=0}^{i-1} \delta_j \right) (\alpha_i + \delta_i) |e_{0\cdot 0}|$.</td>
<td>$\Lambda_i \left( \sum_{j=0}^{i-1} \delta_j \right) (\alpha_i \Lambda + 1) |e_{0\cdot 0}|$.</td>
</tr>
<tr>
<td>9</td>
<td>2</td>
<td>$\Lambda_i^{i+2} \leq |e_{0\cdot 0}|$.</td>
<td>$\Lambda_i \leq |e_{0\cdot 0}|$.</td>
</tr>
<tr>
<td>9</td>
<td>3</td>
<td>$K \equiv \frac{1}{1 - \Lambda} + \Lambda \cdot \alpha_i + \delta_i$.</td>
<td>$K \equiv \frac{1}{1 - \Lambda} \cdot \alpha_i \Lambda + 1$.</td>
</tr>
</tbody>
</table>