数理解析研究所講究録
第 129 巻 1971 年 100-118

March 1971

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

(1) $$Ax = b,$$

where A is a m×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

(2) $$x = Qx + Rb,$$

where the n×n convergent matrix Q and n×m matrix R depend only on A and satisfy the relation

(3) $$Q + RA = I \qquad \mathrm{Im}A \cap \mathrm{Ker}R = \{0\},$$

then an stationary linear iterative formula

(4) $$x^{i+1} = \phi(x^i, b) \equiv Qx^i + Rb, \quad i=0,1,\cdots.$$

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. Nonstatioary (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

$$(5)\qquad \begin{aligned} \lim_{i\to\infty} x^i &= P_{KerA}\, x^0 + A^- b \\ A^- &= (I - \tilde{Q})^{-1} R, \end{aligned}$$

where x^0 is an initial vector, P_{KerA} is the orthogonal projection onto KerA and $\tilde{Q}$ is the restriction of Q onto ImA^*. 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 ϕ_α from C^n to C^n for any real α defined as

(6) $$\begin{aligned}\phi_\alpha(x,\ b) &\equiv x + \alpha\,(\phi(x,\ b) - x)\\ &= Q_\alpha x + R_\alpha b,\end{aligned}$$

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

ϕ_α 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 $\phi(x)$, $\phi^o(x)$, $\phi_\alpha(x)$, $\phi_\alpha^o(x)$ respectively.

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

$$x_{i+1,j} = \phi(x_{i,j}) = Qx_{i,j} + Rb$$

$$x_{i,j+1} = \phi_{\alpha_j}(x_{i,j}) = Q_{\alpha_j}x_{i,j} + R_{\alpha_j}b$$

where $i,j \in N = \{0,1,2,\cdots \infty\}$.

The sequence is well defined since ϕ_α and ϕ are commutative. Note that the first column sequence $\{x_{i.0}\}_{i\in N}$ is generated by the original iteration (4) and that

$$\lim_{i\to\infty} x_{i.k} = \lim_{i\to\infty} x_{i.0} \qquad k \in N.$$

DEFINITION: Let sequences $\{d_{i.j}\}_{i.j\in N}$, $\{e_{i.j}\}_{i.j\in N}$ of column vectors defined as

$$(8) \qquad \begin{aligned} d_{i.j} &= x_{i.j+1} - x_{i.j} \\ e_{i.j} &= x_{i+1.j} - x_{i.j} \quad , \quad i,j \in N. \end{aligned}$$

Then we have

$$(9) \qquad \begin{aligned} e_{i+1.j} &= \phi^{o}(e_{i.j}) = Qe_{i.j} \\ e_{i.j+1} &= \phi^{o}_{\alpha_j}(e_{i.j}) = Q_{\alpha_j} e_{i.j} \quad , \quad i,j \in N. \end{aligned}$$

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

LEMMA 1: $e_{i.j} \in (ES(0) + ES(1))^{\perp}$ for $i \in N^{+}$, $j \in N$, where $N^{+} = N - \{0\}$, $ES(\lambda) = \{ x \; ; \; Qx = \lambda x \}$ and $(*)^{\perp}$ denotes the orthogonal complement of $*$.

Now we introduce an acceleration algorithm.

ACCELERATION ALGORITHM: Given an initial vector $x_{0.0}$, determine the sequence $\{x_{i.i}\}_{i\in N}$ of vectors from the reccurrence formula

$$(10) \qquad \begin{aligned} x_{i+1.i} &= \phi(x_{i.i}) \\ e_{i.i} &= x_{i+1.i} - x_{i.i} \\ e_{i+1.i} &= \phi^{o}(e_{i.i}) \qquad\qquad \text{for } i \in N, \\ \alpha_i &= \frac{\| e_{i.i} \|^2 - \langle e_{i.i} , e_{i+1.i} \rangle}{\| e_{i+1.i} - e_{i.i} \|^2} \\ x_{i+1.i+1} &= x_{i+1.i} + \alpha_i e_{i+1.i} \end{aligned}$$

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 $e_{i_0 i_0} = 0$ 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 α_i on the i-th stage of the algorithm (10) is made so as to minimize the function

$$(11) \qquad f_i(\alpha) \equiv \| e_{i,i+1} \|^2 = \| Q_\alpha e_{i,i} \|^2,$$

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

Let ρ_i and σ_i be defined as

$$(12) \qquad \rho_i \equiv \frac{\langle e_{ii} , e_{i+1,i} \rangle}{\| e_{ii} \|^2}, \qquad \sigma_i \equiv \frac{\| e_{i+1,i} \|^2}{\| e_{i,i} \|^2}$$

where $e_{ii} \neq 0$ is assumed, then we have

$$(13) \qquad f_i(\alpha) = \| e_{ii} \|^2 (1 - 2(1-\rho_i)\alpha + (1 - 2\rho_i + \sigma_i)\alpha^2).$$

Thus the minimum

$$(14) \qquad f_i(\alpha_i) = \| e_{ii} \|^2 \frac{\sigma_i - \rho_i^2}{1 - 2\rho_i + \sigma_i}$$

is attained at

$$(15) \qquad \alpha_i = \frac{1 - \rho_i}{1 - 2\rho_i + \sigma_i} .$$

Let λ_i's are the eigenvalues of Q in the open interval (0, 1) ordered as

$$0 < \lambda_1 < \lambda_2 < \cdots\cdots < \lambda_r < 1,$$

where λ_1 and λ_r will also be denoted by λ and Λ respctively. Corresponding to the direct sum decomposition of $(ES(0)+ES(1))^{\perp}$

$$(ES(0) \oplus ES(1))^{\perp} = ES(\lambda_1) \oplus \cdots\cdots \oplus ES(\lambda_r),$$

we have the unique decomposition of e_{ii}

$$(16) \qquad \begin{aligned} e_{i.i} &= \varepsilon_i^1 \oplus \varepsilon_i^2 \oplus \cdots\cdots \oplus \varepsilon_i^r \\ \varepsilon_i^j &\in ES(\lambda_j), \qquad i \in N^+. \end{aligned}$$

Then ρ_i and σ_i are convex combination of λ_j and λ_j^2 respectively, i.e.

$$(17) \qquad \rho_i = \frac{\sum_{j=1}^{r} \lambda_j \|\varepsilon_i^j\|^2}{\sum_{j=1}^{r} \|\varepsilon_i^j\|^2}, \qquad \sigma_i = \frac{\sum_{j=1}^{r} \lambda_j^2 \|\varepsilon_i^j\|^2}{\sum_{j=1}^{r} \|\varepsilon_i^j\|^2}.$$

Let the points (λ_j, λ_j^2) in the (ρ,σ)- plane be denoted by P_j, and let the convex hull of the points $P_1, P_2, \cdots, P_r$ be denoted by $CH(P_1, \cdots, P_r)$ (or abbr. CH). The equation (17) means

$$(\rho_i, \sigma_i) \in CH(P_1, \cdots, P_r).$$

Define functions **g** and **h** as

$$(18) \qquad \begin{aligned} g(\rho,\sigma) &\equiv \frac{1 - \rho}{1 - 2\rho + \sigma} \\ h(\rho,\sigma) &\equiv \frac{\sigma - \rho^2}{1 - 2\rho + \sigma}. \end{aligned}$$

Elaborate calculation shows

$$(19) \qquad \begin{aligned} \min_{(\rho\sigma)\in CH} g(\rho,\sigma) &= g(\lambda,\lambda^2) = \frac{1}{1 - \lambda}, \\ \max_{(\rho\sigma)\in CH} g(\rho,\sigma) &= g(\Lambda,\Lambda^2) = \frac{1}{1 - \Lambda}, \end{aligned}$$

$$(20)\quad \begin{aligned} &\underset{(\rho\sigma)\in CH}{\text{Min}}\ h(\rho,\sigma) = h(\lambda_i,\lambda_i^2) = 0, \quad i \in \{1,2,\cdots,r\}, \\ &\underset{(\rho\sigma)\in CH}{\text{Max}}\ h(\rho,\sigma) = h(u,v) = \frac{(\Lambda-\lambda)^2}{(2-\Lambda-\lambda)^2}, \text{ where} \\ &u = \frac{\Lambda+\lambda-2\Lambda\lambda}{2-\Lambda-\lambda}, \quad v = \frac{\Lambda^2(1-\lambda)+\lambda^2(1-\Lambda)}{2-\Lambda-\lambda}. \end{aligned}$$

Thus we have the following lemma.

LEMMA 2:

$$(21)\quad \begin{aligned} &1 < \frac{1}{1-\lambda} \leqq \alpha_i \leqq \frac{1}{1-\Lambda}, \\ &0 \leqq \delta_i \leqq \frac{\Lambda-\lambda}{2-\Lambda-\lambda} < \Lambda, \end{aligned}$$

where $\delta_i = \sqrt{h(\rho_i,\sigma_i)}$, and $\frac{\Lambda-\lambda}{2-\Lambda-\lambda}$ decreases monotonically from $\frac{\Lambda}{2-\Lambda}$ to 0 as λ varies from 0 to Λ.

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

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

$$\lim_{i\to\infty} x_{ii} = \lim_{i\to\infty} x_{i0} = P_{\text{Ker}A}x_{00} + (I-\tilde{Q})^{-1}Rb.$$

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

$$e_{i+1.i+1} = Qe_{i.i+1} = \tilde{Q}e_{i.i+1} = \tilde{Q}Q_{\alpha_i}e_{ii}$$

then $\|e_{i+1.i+1}\| \leqq \|\tilde{Q}\|\,\|Q_{\alpha_i}e_{ii}\| = \Lambda\delta_i\|e_{ii}\| \leqq \Lambda^{i+1}(\prod_{j=0}^{i}\delta_j)\|e_{00}\|$.

Since $d_{i+1.i} = \alpha_i e_{i+1.i} = \alpha_i\tilde{Q}e_{ii}$, we have

$$\|d_{i+1.i}\| \leqq \alpha_i\Lambda\|e_{ii}\|.$$

Thus
$$\begin{aligned}\|x_{i+1.i+1} - x_{i.i}\| &\leqq \|d_{i+1.i}\| + \|e_{i+1.i+1}\| \\ &\leqq \Lambda(\alpha_i+\delta_i)\|e_{i.i}\|\end{aligned}$$

$$\leq \Lambda^{i+1}\left(\prod_{j=0}^{i}\delta_j\right)(\alpha_i + \delta_i)\|e_{oo}\|$$
$$< \Lambda^{2i+2} K \|e_{oo}\|,$$

since $\delta_i < \Lambda$ and $K \equiv \frac{1}{1-\Lambda} + \Lambda > \alpha_i + \delta_i$.

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 $\phi(\tilde{x}, b) = \tilde{x}$, where $\tilde{x} = \lim_{i\to\infty} x_{ii}$. Noting that the component $P_{KerA}x_{oo}$ is invariant under the transformations ϕ and ϕ_α's we have the desiered result.

THEOREM 2: $\|x_{ii} - \tilde{x}\| \leq \frac{\Lambda^i}{1-\Lambda}\left(\prod_{j=0}^{i-1}\delta_j\right)\|e_{oo}\|$,

where $\tilde{x} = P_{KerA}x_{oo} + (I - \tilde{Q})^{-1}Rb$.

Proof) It is easily seen from (3) that

$$KerA = ES(1) \text{ and } ImA^- \subset ImA^*,$$

then we have $RAA^- = RA(I - \tilde{Q})^{-1}R = (I - Q)(I - \tilde{Q})^{-1}R = R$.

We have

$$\begin{aligned}
ImA^* \ni x_{ii} - \tilde{x} &= \bar{x}_{ii} - A^- b \\
&= (I - \tilde{Q})^{-1}(I - Q)(x_{ii} - A^- b) \\
&= (I - \tilde{Q})^{-1}RA(\bar{x}_{ii} - A^- b) \\
&= (I - \tilde{Q})^{-1}(RAx_{ii} - Rb) \\
&= (I - \tilde{Q})^{-1}((I - Q)x_{ii} - Rb) \\
&= -(I - \tilde{Q})^{-1}e_{ii},
\end{aligned}$$

where $\bar{x}_{ii} = x_{ii} - P_{\mathrm{KerA}} x_{ii}$. Thenwe have

$$\| x_{ii} - \tilde{x} \| \leqq \frac{1}{1-\Lambda} \| e_{ii} \| \leqq \frac{\Lambda^{i}}{1-\Lambda} \left(\prod_{j=0}^{i-1} \delta_j \right) \| e_{oo} \| .$$

This completes the proof.

Although the behaviour of (ρ_i, σ_i) is very complicated in general, we have the following theorem.

THEOREM 3: If $\lambda_1 + \lambda_2 > \lambda_r$, then

$$\lim_{i\to\infty} (\rho_{2i-1}, \sigma_{2i-1}) = (\rho_{odd}, \sigma_{odd}) = P_{odd}$$

(22) $$\lim_{i\to\infty} (\rho_{2i}, \sigma_{2i}) = (\rho_{even}, \sigma_{even}) = P_{even}$$

$$\frac{\overline{P_l P_{odd}}}{\overline{P_{odd} P_r}} \frac{\overline{P_l P_{even}}}{\overline{P_{even} P_r}} = \xi$$

where

$$\xi \text{ or } 1/\xi = \frac{\lambda^2 (1-\Lambda)^2}{\Lambda^2 (1-\lambda)^2} \neq 1.$$

Either $P_{odd} \in LS^+$, $P_{even} \in LS^-$ or $P_{even} \in LS^+$, $P_{odd} \in LS^-$ occurs, where

$$LS^+ = \{ (\rho, \sigma) \in P_l P_r \ ; \ \Delta(\rho, \sigma) > 0 \},$$

$$LS^- = \{ (\rho, \sigma) \in P_l P_r \ ; \ \Delta(\rho, \sigma) < 0 \},$$

$$\Delta(\rho,\sigma) = H(\rho,\sigma,\lambda) + H(\rho,\sigma,\Lambda)$$

$$H(\rho,\sigma,\bar{\lambda}) = \frac{1}{1-2\rho+\sigma} \{ (\sigma-\rho)\bar{\lambda}' + (1-\rho)\bar{\lambda}^2 \}.$$

and $\overline{ST}$ denotes the length of the line segment ST between two points S and T.

Proof) Since $e_{i+1,i+1} = QQ_{\alpha_i} e_{i,i}$, we have

$$\varepsilon_{i+1}^{j} = \eta_i^j \varepsilon_i^j$$

$$\eta_i^j = \lambda_j (1 - \alpha_i (1 - \lambda_j))$$

$$= H(\rho_i, \sigma_i, \lambda_j).$$

Since $H(\rho, \sigma, \bar{\lambda})$ is a quadratic function in $\bar{\lambda}$, it attains its minimum at

$$\bar{\lambda} = \frac{\rho - \sigma}{2(1 - \rho)} \equiv \Theta(\rho,\sigma).$$

Since

$$\Theta(\rho_i, \sigma_i) = \frac{\sum_{j=1}^{r} \lambda_j (1 - \lambda_j) \|\varepsilon_i^j\|^2}{2 \sum_{j=1}^{r} (1 - \lambda_j) \|\varepsilon_i^j\|^2}, \quad i \in N^+$$

we have $\lambda/2 \leqq \Theta(\rho_i, \sigma_i) \leqq \Lambda/2$. Since

$$H(\rho_i, \sigma_i, \lambda_\ell) = \frac{\sum_{j=1}^{r} \lambda_\ell (\lambda_\ell - \lambda_j)(1 - \lambda_j) \|\varepsilon_i^j\|^2}{\sum_{j=1}^{r} (1 - \lambda_j)^2 \|\varepsilon_i^j\|^2}, \quad i \in N^+$$

we have for $i \in N^+$

(23) $$\eta_i^1 \leqq 0 \leqq \eta_i^r .$$

We have $$\lambda + 2(\Theta(\rho_i, \sigma_i) - \lambda) = \frac{\sum_{j=1}^{r} (\lambda_j - \lambda)(1 - \lambda_j) \|\varepsilon_i^j\|^2}{\sum_{j=1}^{r} (1 - \lambda_j) \|\varepsilon_i^j\|^2}$$

$$\leqq \Lambda - \lambda, \quad i \in N^+$$

then if $\Lambda - \lambda < \lambda_2$,

(24) $$H(\rho_i, \sigma_i, \lambda_1) < H(\rho_i, \sigma_i, \lambda_2) < \cdots < H(\rho_i, \sigma_i, \lambda_r),$$

for $i \in N^+$, (See Fig. 5-(a)) i.e.

(25) $$\eta_i^1 < \eta_i^2 < \cdots\cdots < \eta_i^r, \quad \text{for } i \in N^+.$$

We have also (25) if $2\lambda \geqq \Lambda$ (See Fig. 5-(b).) which is implied in the assumption of (24).

Noting (23) and (25), we have, as $i \to \infty$

$$
(26)\qquad
\begin{aligned}
e_{ii} &\sim \varepsilon_i^1 + \varepsilon_i^r ,\\
\rho_i &\sim \frac{\lambda \|\varepsilon_i^1\|^2 + \Lambda \|\varepsilon_i^r\|^2}{\|\varepsilon_i^1\|^2 + \|\varepsilon_i^r\|^2} ,\\
\sigma_i &\sim \frac{\lambda^2 \|\varepsilon_i^1\|^2 + \Lambda^2 \|\varepsilon_i^r\|^2}{\|\varepsilon_i^1\|^2 + \|\varepsilon_i^r\|^2} .
\end{aligned}
$$

Then we have

$$
\begin{aligned}
\eta_i^1 &\sim \lambda(\lambda - \Lambda)(1 - \Lambda)\|\varepsilon_i^r\|^2 \quad /L\\
\eta_i^r &\sim \Lambda(\Lambda - \lambda)(1 - \lambda)\|\varepsilon_i^1\|^2 \quad /L
\end{aligned}
$$

where $L := (1 - \lambda)^2\|\varepsilon_i^1\|^2 + (1 - \Lambda)^2\|\varepsilon_i^r\|^2$, then

$$
(27)\qquad
\begin{aligned}
\frac{\|\varepsilon_{i+1}^1\|^2}{\|\varepsilon_{i+1}^r\|^2} &\sim \left(\frac{\eta_i^1}{\eta_i^r}\right)^2 \frac{\|\varepsilon_i^1\|^2}{\|\varepsilon_i^r\|^2}\\
&= \frac{\lambda^2(1 - \Lambda)^2}{\Lambda^2(1 - \lambda)^2}\,\frac{\|\varepsilon_i^r\|^2}{\|\varepsilon_i^1\|^2} ,
\end{aligned}
$$

and then

$$
\frac{\|\varepsilon_{i+1}^1\|^2}{\|\varepsilon_{i+1}^r\|^2} \sim \frac{\|\varepsilon_{i-1}^1\|^2}{\|\varepsilon_{i-1}^r\|^2} .
$$

Thus we have the first half of the theorem.

If there exists an integer i^* such that $P_1 P_r$ and the line defined by $\Delta(\rho,\sigma)$ intersect at $(\rho_{i^*},\sigma_{i^*})$, i.e.

$$
(28)\qquad
\begin{aligned}
\rho_{i^*} &= \frac{\Lambda^2(1 - \lambda) + \lambda^2(1 - \Lambda)}{\Lambda + \lambda - 2\Lambda\lambda}\\
\sigma_{i^*} &= \frac{\Lambda^3(1 - \lambda) + \lambda^3(1 - \Lambda)}{\Lambda + \lambda - 2\Lambda\lambda}
\end{aligned}
$$

then it is easily seen that

(29) $(\rho_i, \sigma_i) = (\rho_{i*}, \sigma_{i*})$, for any $i \in N^+$,

that is, The point (ρ_{i*}, σ_{i*}) is a stationary point of the discrete dynamical system defined by (10) and (12). Generally we have

$$(\rho_i, \sigma_i) \neq (\rho_{i*}, \sigma_{i*}).$$

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

$$0 \leqq -H(\rho_i, \sigma_i, \lambda_1) < H(\rho_i, \sigma_i, \lambda_r)$$

or $$0 \leqq H(\rho_i, \sigma_i, \lambda_r) < -H(\rho_i, \sigma_i, \lambda_1),$$

which means $\lim_{i \to \infty} (\rho_i, \sigma_i) = P_r$ or P_1, contradicting (27). Thus P_{odd} and P_{even} are separated by the line $\Delta(\rho, \sigma) = 0$, and (ρ_{i*}, σ_{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 accelerat-ion algorithm is applicable to most of linear iterative process-es. 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_{ii} and $x_{i+1,i}$ coinside 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

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

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

$$(30) \qquad \alpha_i = \frac{\| e_{ii} \|^2}{\| e_{ii} \|^2 - \langle e_{ii} , e_{i+1,i} \rangle},$$

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

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

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.

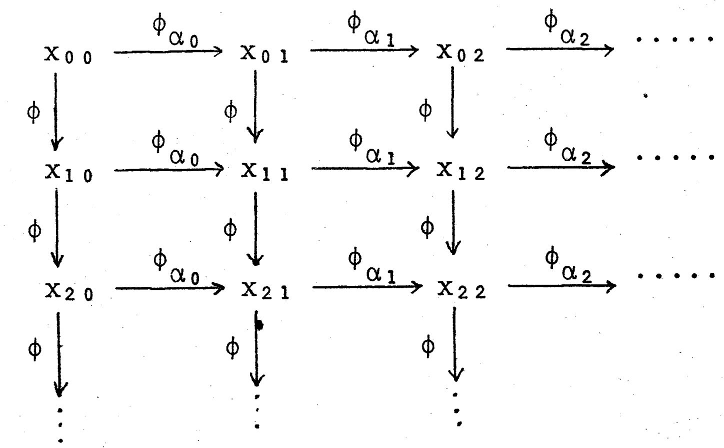
x00 x01 x02
x10 x11 x12
x20 x21 x22
φα0 φα1 φα2
φ

Fig. 1

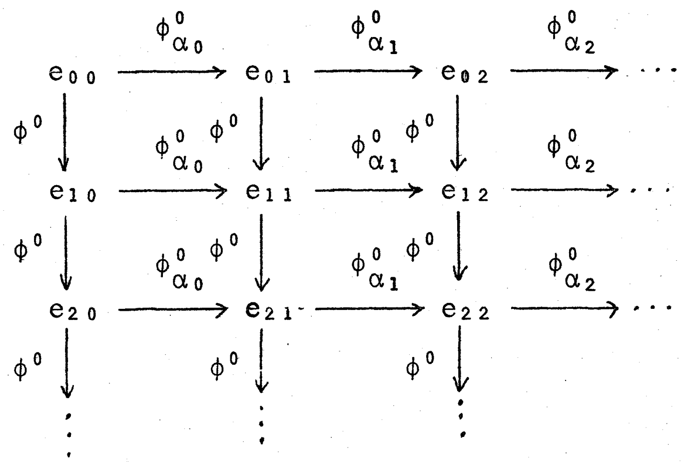
e00 e01 e02
e10 e11 e12
e20 e21 e22
φ0α0 φ0α1 φ0α2
φ0

Fig.2

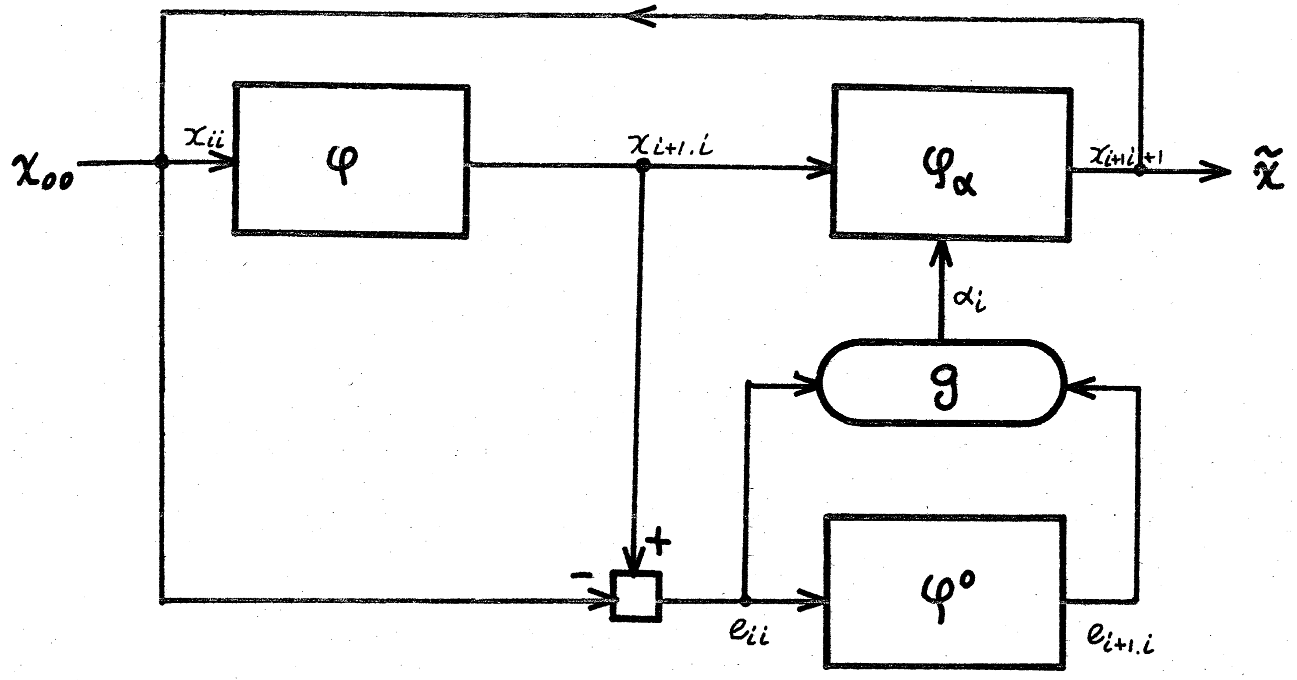

Fig. 3. Block Diagram of Algorithm (10).

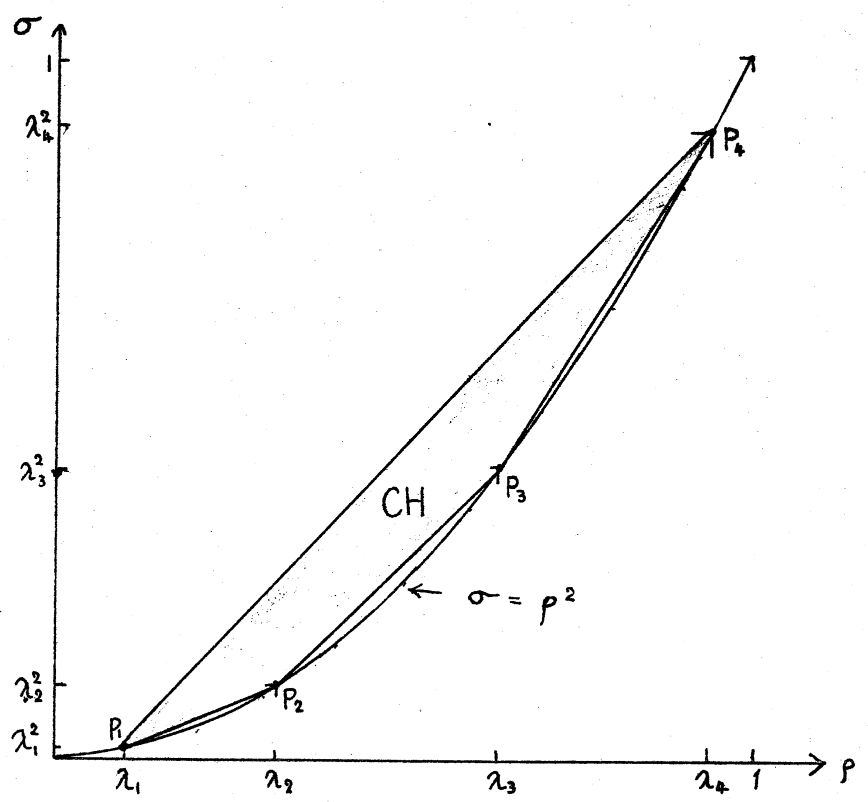

Fig.4.Polygon CH

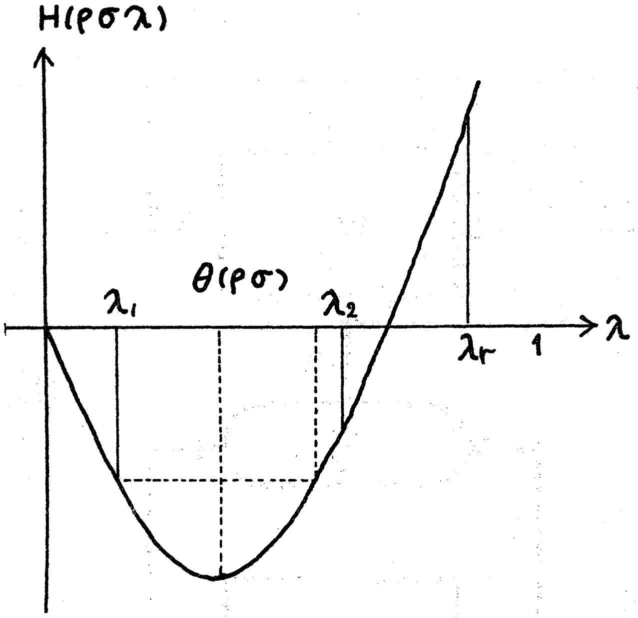

Fig. 5-(a)

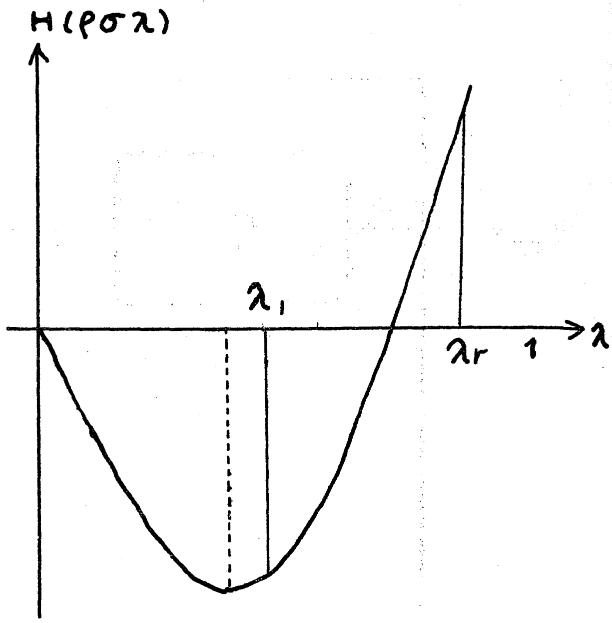

Fig. 5-(b)

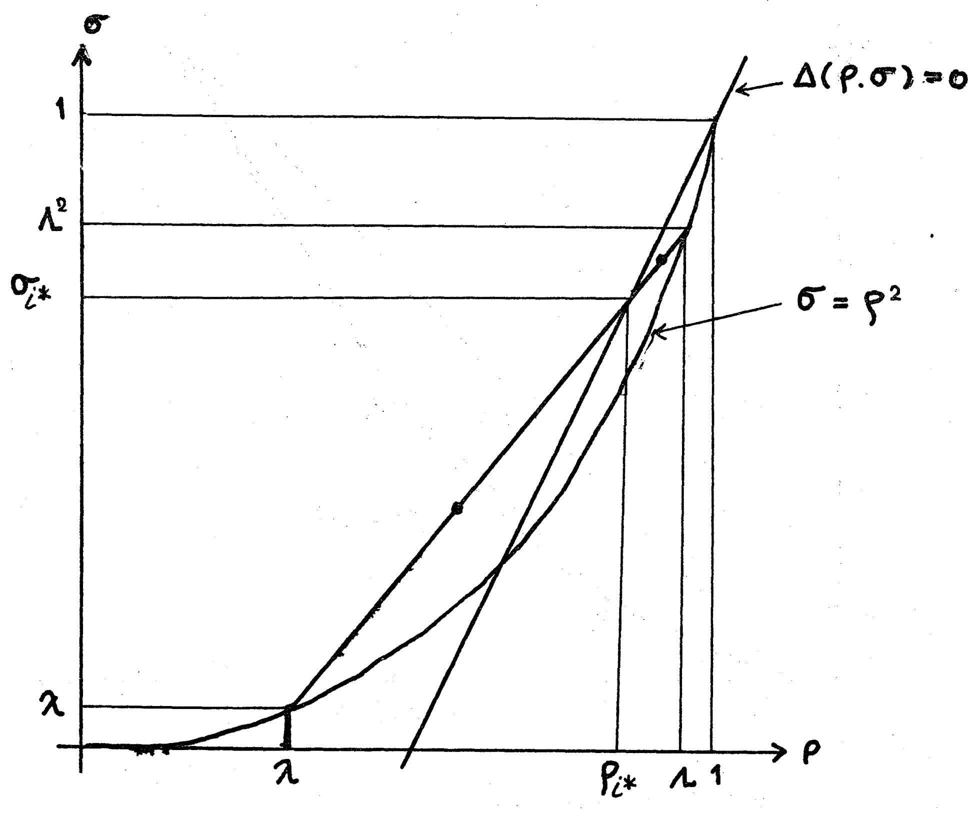

Fig. 6.

REFERENCES

1. Ben-Israel, A., Charnes, A., Contributions to the theory of generalized inverses, J. Soc. Indust. Appl. Math., 11, 667-699 (1963).
2. Forsythe, G.E., Solving linear algebraic equations can be interesting, Bull. Amer. Math. Soc., 59, 299-329 (1953).
3. Sibuya, M., Subclasses of generalized inverses of matrices, Annals of the Institute of Statistical Mathematics, 22-3, 543-556 (1970).
4. Tanabe, K., Projection method for solving a singular system of linear equations, to appear in Numerische Mathematik.
5. Tewarson, R.P., On some representations of generalized inverses, SIAM Rev., 11, 272-276 (1969).
6. Westlake, J.R., A handbook of numerical matrix inversion and solution of linear equations. New York: John Wiley 1968.
7. Wiener, N., Cybernetics. New York: John Wiley 1961.

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

Kunio TANABE
Inst. Statist. Math.

∘∘∘∘∘∘∘∘∘∘∘∘∘∘∘∘∘

(b: From the bottom)

Page	Line	Old	New
3	9	$\tilde{Q}$ is the restriction of Q onto $\mathrm{Im}A^*$.	$\tilde{Q} = QP_{\mathrm{Im}A^*}$, $P_{\mathrm{Im}A^*} = I - P_{\mathrm{Ker}A}$.
5	b2	$\dfrac{\\|e_{ii}\\|^2 - \langle e_{ii}, e_{i+1.i}\rangle}{\\|e_{i+1.i} - e_{ii}\\|^2}$	$\dfrac{\langle e_{ii}, e_{ii} - e_{i+1,i}\rangle}{\\|e_{ii} - e_{i+1.i}\\|^2}$
7	b7	$(\rho_i, \sigma_i) \in CH(P_1, \cdots, P_r)$.	$(P_i, \sigma_i) \in CH(P_1, \cdots P_r)$ for $i \in N^+$.
8	b2	$\\|d_{i+1.i}\\| + \\|e_{i+1.i+1}\\|$	$\\|d_{i+1.i}\\| + \\|e_{i.i}\\|$
8	b1	$\Lambda(\alpha_i + \delta_i)\\|e_{ii}\\|$	$(\alpha_i\Lambda + 1)\\|e_{ii}\\|$
9	1	$\Lambda^{i+1}(\prod_{j=0}^{i}\delta_j)(\alpha_i + \delta_i)\\|e_{00}\\|$	$\Lambda^i(\prod_{j=0}^{i-1}\delta_j)(\alpha_i\Lambda + 1)\\|e_{00}\\|$
9	2	$\Lambda^{2i+2}K\\|e_{00}\\|$	$\Lambda^{2i}K\\|e_{00}\\|$
9	3	$K \equiv \frac{1}{1-\Lambda} + \Lambda > \alpha_i + \delta_i$.	$K \equiv \frac{1}{1-\Lambda} > \alpha_i\Lambda + 1$