数理解析研究所講究録
第382巻 1980年 1-14

連立一次方程式の数値解の誤差の事後評価

中央大 理工　西見二昭

§1 序

Gauss-Seidel 反復法:

(0) $x^{(\nu+1)} = B_1 x^{(\nu+1)} + B_2 x^{(\nu)} + c \quad , \nu = 0, 1, 2, \ldots$.

ただし; $Ax = b \Leftrightarrow x = (B_1 + B_2)x + c$, $A\alpha = b$

の数値解 $x^{(\nu+1)}$ の誤差の事後評価についての既知の結果;

(1) $\|x^{(\nu+1)} - \alpha\| \leqq \dfrac{\|B_2\|}{1 - \|B_1 + B_2\|} \|x^{(\nu+1)} - x^{(\nu)}\|$ (Dück, 1959 [1]),

(2) $\|x^{(\nu+1)} - \alpha\| \leqq \dfrac{\lambda}{1-\lambda} \|x^{(\nu+1)} - x^{(\nu)}\|$.

ただし $\lambda = \max_i \lambda_i$, $\lambda_i \triangleq \sum_{k=1}^{i-1} |B_{ik}| \lambda_k + \sum_{k=i+1}^{n} |B_{ik}|$

$B = B_1 + B_2$, B_1 は下半三角行列, B_2 は上半三角行列

(Sassenfeld, 1951 [2])

しかし，これらの結果は実際には使いものにならない。それは反復法の計算における，丸め誤差，ケタ落ち，情報落ちなどの影響をまったく考慮していないからである。たとえば，ある ν について $x^{(\nu+1)} = x^{(\nu)}$ となるような簡単な例をつくることができるが，この場合これらの評価式によると $x^{(\nu+1)}$ の誤差はゼロとなるが，これはあきらかに矛盾である。

本稿では計算誤差の影響を正しく考慮に入れた、Sück型、Sassenfeld型の事後評価式を導く。さらにこれらの評価式を非線型方程式の場合に拡張し、「一般のGauss-Seidel法」の収束条件をあきらかにする（§2, §3）。

最後に、上記結果を利用して、連立一次方程式の近似解が与えられたとき、「sharpな誤差評価をともなった数値解」をもとめる実用的な処方箋を与える（§4, §5, §6）。

§2　線型方程式に対するGauss-Seidel法

反復公式：

$$\tilde{x}^{(\nu+1)} = [[B_1\tilde{x}^{(\nu+1)}]_{FL} + [B_2\tilde{x}^{(\nu)}]_{FL} + c]_{FL}, \quad \nu=0,1,2,\cdots \qquad (3)$$

（ただし $[\cdot]_{FL}$ は浮動小数点演算の結果）

による数値解の誤差は次の式によって正しく評価される。

$$(4) \quad \|\tilde{x}^{(\nu+1)} - \alpha\| \leqq \frac{1}{1-\|B\|}\left\{\|\tilde{x}^{(\nu+1)} - \hat{x}^{(\nu+1)}\| + \|B_2\|\,\|\tilde{x}^{(\nu+1)} - \tilde{x}^{(\nu)}\|\right\},$$

（ただし、$\|B\|<1$）

$$(5) \quad \|\tilde{x}^{(\nu+1)} - \alpha\|_\infty \leqq \frac{1}{1-\lambda}\left\{\delta + \lambda\|\tilde{x}^{(\nu+1)} - \tilde{x}^{(\nu)}\|_\infty\right\}.$$

ただし、$\hat{x}^{(\nu+1)} = B_1\tilde{x}^{(\nu+1)} + B_2\tilde{x}^{(\nu)} + c$,

$$\delta \triangleq \max_i \delta_i, \quad \delta_i \triangleq \sum_{k=1}^{i-1}|B_{ik}|\delta_k + |\tilde{x}_i^{(\nu+1)} - \hat{x}_i^{(\nu+1)}|.$$

(4), (5) はそれぞれ (1), (2) の改良である。これらは次節の結果の特別の場合である（カッコ内第1項が新しく追加された項）。

§3. 非線型方程式に対するGauss-Seidel型反復法

［定理1］　X：ノルム空間，D：完備集合，

$\tilde{D} \subsetneqq D \subset X$，　$f: D\times D \to D$，$F: D\to D$

$x \in D \Rightarrow F(x) \triangleq f(x,x)$, かつ $\alpha = F(\alpha)$ となる $\alpha \in D$ 存在。

また任意の $x, y, z \in D$ に対して,

(6) $\| f(x,y) - f(x,z) \| \leqq \ell(x) \| y - z \|$,

(7) $\| F(x) - F(y) \| \leqq L \| x - y \|$, $0 \leqq \ell < 1$

となる $\ell(x)$, L が存在。

$\tilde{f} : \tilde{D} \times \tilde{D} \to \tilde{D}$ による反復法:

(8) $\hat{y}^{(\nu+1)} = \tilde{f}(\tilde{y}^{(\nu+1)}, \tilde{y}^{(\nu)})$, $\nu = 0, 1, 2, \cdots$

が定義できるとき(即ち, $\hat{y}^{(0)} \in \tilde{D} \Rightarrow \forall \nu, \hat{y}^{(\nu)} \in \tilde{D}$ のとき)

(9) $\| \tilde{y}^{(\nu+1)} - \alpha \| \leqq \dfrac{1}{1-L} \left\{ \| \tilde{y}^{(\nu+1)} - \hat{y}^{(\nu+1)} \| + L_2 \| \tilde{y}^{(\nu+1)} - \hat{y}^{(\nu)} \| \right\}$.

(10) ただし $\hat{y}^{(\nu+1)} = f(\tilde{y}^{(\nu+1)}, \tilde{y}^{(\nu)})$,

(11) $\ell(\tilde{y}^{(\nu+1)}) \leqq L_2$.

(証明)

$$\tilde{y}^{(\nu+1)} - \alpha = \tilde{f}(\hat{y}^{(\nu+1)}, \tilde{y}^{(\nu)}) - f(\alpha, \alpha)$$

$$= \tilde{f}(\tilde{y}^{(\nu+1)}, \tilde{y}^{(\nu)}) - f(\tilde{y}^{(\nu+1)}, \tilde{y}^{(\nu)}) + f(\tilde{y}^{(\nu+1)}, \tilde{y}^{(\nu)}) - f(\alpha, \alpha)$$

$$= \tilde{y}^{(\nu+1)} - \hat{y}^{(\nu+1)} + f(\tilde{y}^{(\nu+1)}, \hat{y}^{(\nu+1)}) - f(\alpha, \alpha) + f(\tilde{y}^{(\nu+1)}, \tilde{y}^{(\nu)}) - f(\tilde{y}^{(\nu+1)}, \tilde{y}^{(\nu+1)}).$$

$$\therefore \| \tilde{y}^{(\nu+1)} - \alpha \| \leqq \| \tilde{y}^{(\nu+1)} - \hat{y}^{(\nu+1)} \| + \| F(\tilde{y}^{(\nu+1)}) - F(\alpha) \| + \ell(y^{(\nu+1)}) \| \tilde{y}^{(\nu+1)} - \tilde{y}^{(\nu)} \|.$$

(7), (11) により

$$\| \tilde{y}^{(\nu+1)} - \alpha \| \leqq \| \tilde{y}^{(\nu+1)} - \hat{y}^{(\nu+1)} \| + L \| \tilde{y}^{(\nu+1)} - \alpha \| + L_2 \| \tilde{y}^{(\nu+1)} - \tilde{y}^{(\nu)} \|,$$

これから(9)がえられる。 QED

[定理1]は (8)を計算誤差をふくむ近似反復法とみなすとき，Collatz [3] による「一般化された Dück の定理」の改良となっている。(9)のカッコ内第2項が Dück-Collatz が与えた項，第1項が新しく追加された項で計算誤差の影響を正しく反映している。第1項の追加により§1で述べた矛盾が解決された。

次の系1-1は前定理から容易に導かれる。説明のない記号は[定理1]と同じ意味である。

[系1-1] $X = R^n$, $\tilde{D} \subsetneq D \subset X$, D：凸体。
$\tilde{D} = \{ {}^t(\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n) \mid \forall i,\ x_i \in \mathcal{F} \}$, $\mathcal{F}$：浮動小数点数全体からなる有限集合。

$f(x, y)$, $F(x)$ がそれぞれ $D \times D$, D をふくむ領域で連続的微分可能で，$\left|\frac{\partial F_i}{\partial x_k}\right| \leqq M_{ik}$, $\left|\frac{\partial f_i(x,y)}{\partial y_k}\right| \leqq m_{ik}$,
$\| \mathrm{matrix}(M_{ik}) \| \leqq L < 1$, $\| \mathrm{matrix}(m_{ik}) \| \leqq L_2$
のとき

$$(12)\quad \| \tilde{y}^{(\nu+1)} - \alpha \| \leqq \frac{1}{1-L} \left\{ \| \tilde{y}^{(\nu+1)} - \hat{y}^{(\nu+1)} \| + L_2 \| \tilde{y}^{(\nu+1)} - \tilde{y}^{(\nu)} \| \right. \blacksquare$$

不等式(4)はこの系の特別の場合(f, F が linear mapping のとき)である。

[定理2]　$X = R^n$; $f, \tilde{f}, \mathfrak{F}, \tilde{\mathfrak{F}}$ などの意味は [系1-1] におけるものと同様。

$f(x, y)$ が $x = \alpha$, $y = \alpha$ をふくむ適当な近傍で連続的微分可能で, $\left|\frac{\partial f_i(x,y)}{\partial x_k}\right| \leqq M_{ik}$ $(k < i)$, $\left|\frac{\partial f_i(x,y)}{\partial y_k}\right| \leqq M'_{ik}$ $(k \geqq i)$,

$l \triangleq \max_i l_i$,　$l_i \triangleq \sum_{k=1}^{i-1} M_{ik} l_k + \sum_{k=i}^{n} M'_{ik}$,

$\delta \triangleq \max_i \delta_i$,　$\delta_i \triangleq \sum_{k=1}^{i-1} M_{ik}\delta_k + |\tilde{y}_i^{(\nu+1)} - \hat{y}_i^{(\nu+1)}|$, で

$l < 1$ のとき,

(13)　$\|\tilde{y}^{(\nu+1)} - \alpha\|_\infty \leqq \dfrac{1}{1-l}\left\{\delta + l\|\tilde{y}^{(\nu+1)} - \hat{y}^{(\nu)}\|_\infty\right\}$.

($\tilde{y}^{(\nu)}$ の条件つき収束定理).

(証明)

$\varepsilon_i^{(\nu)} \triangleq \tilde{y}_i^{(\nu)} - \alpha_i$ とすると,

$$\begin{aligned}
\varepsilon_i^{(\nu+1)} &= \tilde{f}_i(\tilde{y}^{(\nu+1)}, \tilde{y}^{(\nu)}) - f_i(\alpha, \alpha) \\
&= \tilde{f}_i(\tilde{y}^{(\nu+1)}, \tilde{y}^{(\nu)}) - f_i(\tilde{y}^{(\nu+1)}, \tilde{y}^{(\nu)}) \\
&\quad + f_i(\alpha + \varepsilon^{(\nu+1)}, \alpha + \varepsilon^{(\nu)}) - f_i(\alpha, \alpha).
\end{aligned}$$

$$\therefore |\varepsilon_i^{(\nu+1)}| \leqq |\tilde{y}_i^{(\nu+1)} - \hat{y}_i^{(\nu+1)}| + \sum_{k=1}^{i-1} M_{ik}|\varepsilon_k^{(\nu+1)}| + \sum_{k=i}^{n} M'_{ik}|\varepsilon_k^{(\nu)}|.$$

$\varepsilon \triangleq \max_i |\varepsilon_i^{(\nu)}|$,　$\delta'_i \triangleq |\varepsilon_i^{(\nu+1)}|/\varepsilon$ とすると

$$\delta'_i \leqq \frac{1}{\varepsilon}|\tilde{y}_i^{(\nu+1)} - \hat{y}_i^{(\nu+1)}| + \sum_{k=1}^{i-1} M_{ik}\delta'_k + \sum_{k=i}^{n} M'_{ik}.$$

とくに

$$\delta'_1 \leqq \frac{1}{\varepsilon}|\tilde{y}_1^{(\nu+1)} - \hat{y}_1^{(\nu+1)}| + \sum_{k=1}^{n} M'_{1k} = \frac{\delta_1}{\varepsilon} + l_1.$$

いま $\delta_k' \leqq \dfrac{\delta_k}{\varepsilon} + l_k$,$k=1,\cdots,i-1$ が成立していれば

$$\delta_i' \leqq \frac{1}{\varepsilon}|\tilde{y}_i^{(\nu+1)} - \hat{y}^{(\nu+1)}| + \sum_{k=1}^{i-1} M_{ik}\frac{\delta_k}{\varepsilon} + \sum_{k=1}^{i-1} M_{ik}\, l_k + \sum_{k=i}^{n} M_{ik}'$$

$$= \frac{1}{\varepsilon}\delta_i + l_i .$$

従って $i=1,2,\cdots,n$ に対して $\delta_i' \leqq \dfrac{\delta_i}{\varepsilon} + l_i$ が成立。

即ち $|\varepsilon_i^{(\nu+1)}| \leqq \delta_i + l_i \cdot \max_j |\varepsilon_j^{(\nu)}|$.

$$\therefore \quad \|\tilde{y}^{(\nu+1)} - \alpha\|_\infty \leqq \delta + l\|\tilde{y}^{(\nu)} - \alpha\|_\infty$$

$$\leqq \delta + l\|\hat{y}^{(\nu)} - \tilde{y}^{(\nu+1)}\|_\infty + l\|\tilde{y}^{(\nu+1)} - \alpha\|_\infty .$$

これから不等式 (13) が得られる。 QED

前節の (5) は [定理2] の特別な場合である。

連立非線型方程式 $\varphi(x)=0$ に対する一次反復法

(14) $\qquad \varphi(x)=0 \Leftrightarrow x = F(x) \;;\; x^{(\nu+1)} = F(x^{(\nu)})$

は、F が縮小写像であるときに収束することはよく知られている。(計算誤差がある場合の(14)の収束のようすも知られている[4])。

いくつかの例で試みてみるとすでに得られている($x_1^{(\nu+1)},\cdots,x_{i-1}^{(\nu+1)}$)を反復公式の右辺に feed back して $x_i^{(\nu+1)}$ を求める方法を採用する (Gauss-Seidel型) と収束が加速されることがわかる。

しかし連立一次方程式で知られているような Jacobi型

反復法(14)が収束しても Gauss-Seidel 型反復法

(15) $y^{(\nu+1)} = f(y^{(\nu+1)}, y^{(\nu)})$

が収束するとはかぎらない。

次の[定理3]は(15)の収束条件を与えるものである。

[定理3] M_{ik}, M'_{ik}, l_i, l などの意味は定理2と同様(ただし,計算誤差をふくまない反復法(15)では前定理における δ_i, δ はゼロとなる)。

$0 \leqq l < 1$ のとき反復法(15)は収束する。

(証明) $\delta_i' \leqq \dfrac{\delta_i}{\varepsilon} + l_i \quad (i=1,2,\cdots,n)$ を前定理証明の中で示した。今の場合は $\delta_i = 0$ であるから

$$\varepsilon\delta_i' \leqq l_i\varepsilon \quad (i=1,2,\cdots,n)$$ が成立。

$\varepsilon = \|y^{(\nu)} - \alpha\|_\infty$, $\varepsilon\delta_i' = |\varepsilon_i^{(\nu+1)}| = |y_i^{(\nu+1)} - \alpha_i|$ であったから

$$\|y^{(\nu+1)} - \alpha\|_\infty \leqq l\,\|y^{(\nu)} - \alpha\|_\infty .$$

(16) $\therefore \quad \|y^{(\nu+1)} - \alpha\|_\infty \leqq l^{\nu+1}\|y^{(0)} - \alpha\|_\infty .$

$$\therefore\ 0 \leq l < 1 \Rightarrow \lim_{\nu\to\infty} y^{(\nu+1)} = \alpha .$$ QED

[付記]

計算誤差がある場合,各 ν に対する δ の上界を $\bar{\delta}$ とするとき(16)のかわりに

(17) $\|\hat{y}^{(\nu+1)} - \alpha\|_\infty \leqq \dfrac{\bar{\delta}}{1-l} + l^{\nu+1}\|\hat{y}^{(0)} - \alpha\|$

が成立 (Isaacson [4] に相当)。

§4 連立一次方程式における誤差限界つきの数値解の求めかた

いま，近似解 $\tilde{x}^{(0)}$ が与えられたとする。これから $\tilde{x}^{(1)}$ を次式で求める:

$$\hat{x}^{(1)} = [[B\hat{x}^{(1)}]_{Fl} + c]_{Fl}, \quad \text{ただし } Ax = b \Leftrightarrow x = Bx + c.$$

$\|B\| < 1$ ならば (4)

$$\|\tilde{x}^{(1)} - \alpha\| \leqq \frac{1}{1 - \|B\|}\left\{\|\hat{x}^{(1)} - \hat{x}^{(1)}\| + \|B\|\,\|\tilde{x}^{(1)} - \tilde{x}^{(0)}\|\right\}$$

により $\hat{x}^{(1)}$ の誤差限界が与えられる。

$\|B\|$ が 1 にくらべて十分小さければカッコ内の数値は第1項により支配される。そして $\|\hat{x}'^{(1)} - \hat{x}^{(1)}\|$ の値はほんの2〜3ケタもあればよいのだから $\hat{x}^{(1)}$ のかわりに

$$\hat{x}' \doteq B\hat{x}^{(0)} + c \quad \text{([多]倍長計算)}$$

を用いてよい。

Bは次のようにしてきめれば上記の目的にかなう。すなわち

$$Ax = b \Leftrightarrow x = x + \Lambda(Ax - b), \quad \Lambda \text{は未定行列}$$

とすると，$x = x + \Lambda(Ax - b) \Leftrightarrow x = Bx + c$

であるためには $I + \Lambda A = B$, $-\Lambda b = c$ であればよい。

$\|B\|$ を小さくするためには，Λを−Aの「近似」逆行列にえらべばよい。結局，

(18) $\tilde{x}^{(1)} = (I + \Lambda A)\tilde{x}^{(0)} - \Lambda b$ （単精度計算），

(19) $\hat{x}'^{(1)} = (I + \Lambda A)\tilde{x}^{(0)} - \Lambda b$ （[多]倍長計算），

により $\tilde{x}^{(1)}$, $\hat{x}'^{(1)}$ を定めれば

$$\Delta x = \frac{1}{1-\|B\|}\left\{\|\tilde{x}^{(1)}-\hat{x}^{(1)}\| + B\|\tilde{x}^{(1)}-\tilde{x}^{(0)}\|\right\} \tag{20}$$

により $x^{(1)}$ の誤差を sharp に事後評価できる。ただし

$B = I + \Lambda A$, Λ は $-A$ の「近似」逆行列。

§5 数値例

［例 1］

$$A = \begin{bmatrix} 16 & -120 & 240 & -140 \\ -120 & 1220 & -2700 & 1680 \\ 240 & -2700 & 6500 & -4200 \\ -140 & 240 & -4200 & 2800 \end{bmatrix}, \quad b = \begin{bmatrix} 516 \\ -5720 \\ 13640 \\ -7380 \end{bmatrix}.$$

この例の右辺は，解 α が ${}^t(1,-1,1,-1)$ となるように定めてある。$Ax=b$ を FACOM 230-48 により単精度計算（仮数部 16 進 6 ケタ）により解いた結果を $\tilde{x}^{(0)}$ とすると

$\tilde{x}^{(0)} = {}^t(0.999723, -0.9999965, 1.000089, -0.9998822)$

であった。(18) 式により $\tilde{x}^{(1)}$ を計算し，(19), (20) 式により Δx をもとめた ($\Delta x = \|\tilde{x}^{(1)}-\alpha\|_\infty$)。結果は次の通りである。

$$\tilde{x}^{(1)} = \begin{bmatrix} 0.9999784 \\ -0.9999990 \\ 1.0000330 \\ -1.0000350 \end{bmatrix}, \quad \Delta x = 0.0000354.$$

成分ごとの誤差は次の通り。

$$\xi = \begin{bmatrix} -0.0000216 \\ 0.0000010 \\ 0.0000330 \\ \underline{-0.0000350} \end{bmatrix}.$$

この例に示されるように本法は極めてsharpな誤差評価を与えている。

以下に示す[例2]では係数行列Aが4次のHilbert行列(条件数$=1.55\times10^4$)，右辺は解が前例と同一となるように定めた。[例3]は4次のHilbert行列の逆行列を係数Aとし(Hilbert行列の逆行列の要素はすべて整数)、bは[例1]と同一方針で定めてある。このようにしたときに限り途中計算の丸め，ケタ落ち，情報落ち等による，見かけ上のbの攪乱δbが，確実に$\delta x=(\text{条件数})\times\delta b$となって数値解$\tilde{x}$の誤差となる。[例3]は最もタチの悪い例であるが誤差評価のsharpnessは失われていない。以下係数行列Aと$\tilde{x}^{(1)}$，成分ごとの誤差ξと本法による誤差評価Δxのみを示す。

[例2]

$$A=\begin{bmatrix}1 & 1/2 & 1/3 & 1/4\\ 1/2 & 1/3 & 1/4 & 1/5\\ 1/3 & 1/4 & 1/5 & 1/6\\ 1/4 & 1/5 & 1/6 & 1/7\end{bmatrix}$$

$$\tilde{x}^{(1)}=\begin{bmatrix}1.0000410\\ -0.9994811\\ 1.0009760\\ -1.0006100\end{bmatrix},\ \xi=\begin{bmatrix}0.0000410\\ 0.0005189\\ \underline{0.0009760}\\ -0.0006100\end{bmatrix},\ \Delta x=0.0009849$$

例3

$$A = \begin{bmatrix} 16 & -120 & 240 & -140 \\ -120 & 1200 & -2700 & 1680 \\ 240 & -2700 & 6480 & -4200 \\ -140 & 1680 & -4200 & 2800 \end{bmatrix}$$

$$\tilde{x}^{(1)} = \begin{bmatrix} 0.9956512 \\ -1.0007310 \\ 0.9995270 \\ -1.0004730 \end{bmatrix}, \quad \xi = \begin{bmatrix} -0.0043488 \\ 0.0007310 \\ -0.0004730 \\ 0.0004730 \end{bmatrix}, \quad \Delta x = 0.0043586$$

§6. 数値例についての考察

考察に先だって公式 (4) を，他の定理とは独立に導く。

$$\alpha = B\alpha + c, \qquad (\alpha = F(\alpha)).$$

$$\therefore \quad \tilde{x}^{(1)} - \alpha = \tilde{x}^{(1)} - B\alpha - c.$$

$$\hat{x}'^{(1)} = B\tilde{x}^{(0)} + c \qquad \text{([多]倍長計算)}.$$

$$\tilde{x}^{(1)} - \alpha \doteqdot \tilde{x}^{(1)} - \hat{x}'^{(1)} \underline{\underline{- B\alpha}} \; \underline{- c} + B\tilde{x}^{(0)} \underline{+ c} - B\tilde{x}^{(1)} \underline{\underline{+ B\tilde{x}^{(1)}}}.$$

(21) $$\therefore \quad \tilde{x}^{(1)} - \alpha \doteqdot \tilde{x}^{(1)} - \hat{x}'^{(1)} + \underline{\underline{B(\tilde{x}^{(1)} - \alpha)}} + B(\tilde{x}^{(0)} - \tilde{x}^{(1)}).$$

(22) $$\therefore \quad \| \tilde{x}^{(1)} - \alpha \| \lessapprox \| \tilde{x}^{(1)} - \hat{x}^{(1)} \| + \|B\| \, \| \tilde{x}^{(1)} - \alpha \| + \|B\| \, \| \tilde{x}^{(0)} - \tilde{x}^{(1)} \|.$$

matrix norm が vector norm から誘導されていれば，上記不等式は，norm の各公理の不等式がすべて「きつい」ので（つまり各不等式を等号で成立させる vector and/or matrix が必ず存在するから）式 (4) のもととなっている不等式 (22) は「きつい」不等式である。いいかえれば，matrix B に対して，不等式 (4) を等号で成立させる

vector $\tilde{x}^{(1)}-\alpha$, $\tilde{x}^{(1)}-\hat{x}^{(1)}$, $\tilde{x}^{(0)}-\tilde{x}^{(1)}$ が必ず存在する。

それでは この不等式(4)が最も「ゆるく」なるのは どのような場合であろうか。この不等式の「ゆるさ」が，とりもなおさず数値解 $\tilde{x}^{(1)}$ の誤差の事後評価 Δx の over-estimation の程度を示すこととなる。

次の仮定は容易に許容できるものであろう。

1° $\|B\|$ は1にくらべてかなり小さい（$\varepsilon\times 10^{-2}$ 程度とする）。

2° $\|\tilde{x}^{(1)}-\alpha\|$, $\|\tilde{x}^{(1)}-\hat{x}^{(1)}\|$, $\|\tilde{x}^{(0)}-\tilde{x}^{(1)}\|$ のオーダーは大体同じくらいの大きさ（大きさ a の程度）。（$\|\tilde{x}^{(0)}-\tilde{x}^{(1)}\| \ll a$ なら以下の 2ε を ε とすればよい。）

まず(21)式から(22)式にうつるとき，2回

$$\|Bx\| \leqq \|B\|\,\|x\|$$

の不等式が用いられている。$\|x\| \fallingdotseq a$ ならば

$$0 \leqq \|Bx\| \leqq \varepsilon a \times 10^{-2}$$

であるから，$\|Bx\|$ を $\|B\|\,\|x\|$ におきかえるさい，第2，3項において第1項にくらべて高々 2ε% の「ガタ」が生ずる。

次に（式の変形においてはこちらの方が先なのだが）「三角不等式」を用いて「和のノルム」を「ノルムの和」でおきかえるときにも同様に 2ε% 程度以下の「ガタ」を生じる。

しかしながらこれら二つの「ガタ」は両立しえないのである。即ち前者の理由で 2ε% の「ガタ」を生じたとすれば，$\|Bx\|=0$ となり後者の「ガタ」はゼロとなる。逆に後者の理由で 2ε% の

「ガタ」を生じたときは $\|Bx\| = \|B\|\|x\|$ となり前者の理由による「ガタ」はゼロとなるからである。

従って $\|B\| \fallingdotseq 2 \times 10^{-2}$ のとき (4) による誤差の事後評価の over-estimation の程度はせいぜい 2% 程度であることがわかる。

よく「残差を用いた誤差ノルムの評価は、かなり over-estimation となる」といわれているが、上記の考察により、本法では形式上残差を利用しているにもかかわらず、そのようなことはありえないことがわかる。

また $\tilde{A}^{-1}$ を A^{-1} の「近似」行列とするとき、$B \triangleq I - \tilde{A}^{-1}A$ であるが $\|B\|$ の大きさを極端に小さくする必要はないこともこの考察からわかる。ごく特別の場合を除いて $\|B\|$ は 0.2〜0.1 程度、特別の場合でも 0.05 程度であれば、実用上さしつかえないであろう。このような rough な「近似逆行列」を用いても、この方法は従来のどの方法よりも sharp な誤差評価を与える。

従って、大次元の Band Matrix を係数とする Sparse Problem の解の誤差評価にあたっては、解法のプログラムの趣旨を生かして $\|I - \tilde{A}^{-1}A\|$ を前記の程度に小さくする「近似」逆 Band Matrix $\tilde{A}^{-1}$ を用いれば十分であろう。

また $\|B\|$ を計算することにより、上記の論証から、事後評価

Δx の over-estimation の程度を estimate することができることもこの考察により明らかであろう。

参考文献

[1] Dück, W.: Eine Fehlerabschätzung zum Einzelschrittverfahren bei linearen Gleichungssytemen, Num. Math. 1 (1959), 73-77.

[2] Sassenfeld, H.: Ein Hinreichendes Konvergenzkriterium und eine Fehleraschätzung für die Iteration in Einzelschritten bei Linearen Gleichungen, Z. Angew. Math. Mech. 31 (1951), 92-94.

[3] Collatz, L.: "Funktionalanalysis und Numerische Mathematik" (1964) Springer.

[4] Isaacson: "Introduction to Numerical Analysis."