統計数理 第36巻 第1号（1988）

大規模非線形方程式系における丸め誤差の振舞いについて

統計数理研究所　土谷　隆
東京大学工学部　伊理正夫

（1988年3月　受付）

1. はじめに

数値計算の結果に含まれる丸め誤差の評価のために従来から用いられている区間演算(Alefeld and Herzberger (1985)を参照)は，〝丸め誤差の打消し〟などの現象を捉えられないため，実用規模の問題に適用すると，しばしば非現実的に過大な誤差評価を与えてしまうことが問題とされてきた．より精密な誤差評価を行うためには〝計算過程で基本演算を実行するごとに発生する誤差の計算結果（関数値）への影響〟を計算する必要があるが，実際にそれを行うことは困難であると考えられていた．しかし，最近著者の一人によって提案された高速自動微分法を利用すると，精密な誤差評価を行うために必要な諸量を効率良く計算することができる（Iri (1984)；伊理 他 (1985)；土谷 (1986)；伊理，久保田 (1986)；Iri et al. (1988)）．それによると，

（i） 各基本演算（四則演算，初等関数など）において発生する丸め誤差の計算結果（関数値）への影響は線形に加算される［微小誤差の仮定］；
（ii） 各基本演算において発生する丸め誤差はある幅の一様分布に従う独立な確率変数と見なしうる［独立誤差の仮定］；

という2つの仮定に基づいて，計算結果に含まれる丸め誤差の精密な評価を効率の良い算法で得ることが可能になる．本論文では，化学プラントの水・メタノール蒸溜塔の平衡状態を定める108元非線形連立方程式の108個の関数を例として，それらの計算値に含まれる丸め誤差の振舞いを解析する．計算された関数値に含まれる丸め誤差の仮定 (i), (ii) に基づいた確率モデルが実用的な規模の計算過程における丸め誤差の振舞いを十分に良く捉えていることを，これら108個の関数の丸め誤差に関する統計的諸量の理論値と実測値を様々な側面から比較することによって検証する．さらに，様々な計算過程によって計算される関数の丸め誤差の(〝確率変数〟としての）分布形を類型化して捉えるために，上述の仮定から自然に導かれる1パラメータの確率分布族を導入し，パラメータを適当に定めることによって，108個の各関数の丸め誤差の分布の特徴が良く捉えられることを検証する．

2. 関数の計算過程と丸め誤差

いくつかの基本演算 (+, −, ×, ／, sin, cos, exp, log, ...) があらかじめ定められていて，最

初に入力変数の値，計算に必要な定数の値が与えられると，関数の値は「それまでに得られている中間変数（入力変数，定数も中間変数に含む）に対し，基本演算のうちの一つを施して新たな中間変数を計算する」ことを繰り返すことによって求められる．すなわち，n 個の入力変数 $\{x_1, \ldots, x_n\}$, n_c 個の定数 $\{c_1, \ldots, c_{n_c}\}$ から，n_v 個の中間変数を計算して，関数fの値を求める過程は以下のように記述できる：

$$
\begin{aligned}
v_1 &= \phi_1(u_{11}, \ldots, u_{1m_1}), \\
&\vdots \\
v_i &= \phi_i(u_{i1}, \ldots, u_{im_i}), \\
&\vdots \\
f = v_{n_v} &= \phi_{n_v}(u_{n_v1}, \ldots, u_{n_vm_{n_v}}).
\end{aligned} \tag{2.1}
$$

ここで，

$$
\begin{aligned}
&u_{ij} \in \{x_1, \ldots, x_n\} \cup \{c_1, \ldots, c_{n_c}\} \cup \{v_1, \ldots, v_{i-1}\} \\
&\qquad\qquad (i=1, \ldots, n_v;\ j=1, \ldots, m_i), \\
&\phi_i \ (i=1, \ldots, n_v) \text{は基本演算}
\end{aligned}
$$

である．各中間変数を計算するための計算ステップを〝基本計算ステップ〟と呼ぶ．現実には，実数が浮動小数点表現で近似され，基本演算も有限桁の精度で行われるために，各ステップの計算は正確には行われず，丸め誤差が発生する．中間変数v_iを計算するための基本計算ステップ

$$
v_i = \phi_i(u_{i1}, \ldots, u_{im_i}) \tag{2.2}
$$

において，基本演算 ϕ_i に対応して実際に行われる演算を $\hat{\phi}_i$ とし，v_i の計算値を $\hat{v}_i$ と記すことにすると，(2.2)に対応して行われる実際の基本演算は以下のようになる：

$$
\hat{v}_i = \hat{\phi}_i(\hat{u}_{i1}, \ldots, \hat{u}_{im_i}). \tag{2.3}
$$

中間変数 v_i の計算値 $\hat{v}_i$ の正確な値からのずれ $\Delta v_i = \hat{v}_i - v_i$ を v_i の〝集積丸め誤差〟と呼ぶことにする．関数 f の丸め誤差は，関数値 f に対応する中間変数 v_{n_v} の集積丸め誤差である．式(2.2), (2.3)より，Δv_i は次のように表される：

$$
\begin{aligned}
\Delta v_i &= \hat{v}_i - v_i \\
&= \hat{\phi}_i(\hat{u}_{i1}, \ldots, \hat{u}_{im_i}) - \phi_i(u_{i1}, \ldots, u_{im_i}) \\
&= \phi_i(\hat{u}_{i1}, \ldots, \hat{u}_{im_i}) - \phi_i(u_{i1}, \ldots, u_{im_i}) + \delta v_i,
\end{aligned} \tag{2.4}
$$

$$
\delta v_i = \hat{\phi}_i(\hat{u}_{i1}, \ldots, \hat{u}_{im_i}) - \phi_i(\hat{u}_{i1}, \ldots, \hat{u}_{im_i}). \tag{2.5}
$$

ここで，δv_i は，v_i を計算する際に新たに発生する丸め誤差で〝発生誤差〟と呼ばれる．δv_i は，v_i より後に計算される中間変数の計算値に影響を及ぼしながら，最終的に計算された関数値に影響を与える．微小誤差の仮定はより具体的には次のように表現できる．

仮定 1.

$$
\phi_i(\hat{u}_{i1}, \ldots, \hat{u}_{im_i}) - \phi_i(u_{i1}, \ldots, u_{im_i}) = \sum_{k=1}^{m_i} \frac{\partial \phi_i}{\partial u_{ik}} \Delta u_{ik}. \tag{2.6}
$$

これを(2.4)に代入するとΔv_iは，v_i を計算するための引数 u_{ik} $(k=1, \ldots, m_i)$ に含まれる〝伝幡誤差〟Δu_{ik} と発生誤差 δv_i とを用いて，以下のように表現される：

$$\Delta v_i = \sum_{k=1}^{m_i} \frac{\partial \phi_i}{\partial u_{ik}} \Delta u_{ik} + \delta v_i \,. \tag{2.7}$$

これを関数 f を計算するための全過程 (2.1) に適用すると，f の丸め誤差 Δf は，各中間変数 v_i を計算する際の発生誤差 δv_i を用いて，

$$\Delta f = \sum_i \frac{\partial f}{\partial v_i} \delta v_i \tag{2.8}$$

と書ける（和は人力変数，定数を除いた全中間変数についてとる）．すなわち，仮定1のもとでは，関数 f の計算値に含まれる丸め誤差は "計算過程で各中間変数 v_i を計算する際の発生誤差 δv_i に，v_i の値の変動が f の値に与える影響を表す偏微分係数 $\partial f/\partial v_i$ を掛けたものの（入力変数，定数を除いた全中間変数に関する）総和" で表現できる．本論文では (2.8) を丸め誤差の基本的な表現式として採用する．高速自動微分法を使うと，計算過程の全中間変数値を保存しておくことによって，関数 f だけを計算する手間の高々数倍の手間で全ての影響係数 $\partial f/\partial v_i$ を評価することができる（Iri (1984)；伊理，久保田 (1986)）．

式 (2.8) に基づいて丸め誤差の解析を行うには，さらに，実際の計算機の数値表現を考慮した，発生誤差 δv_i に関するいくつかの仮定が必要である．ここでは，数値表現，発生誤差について以下のことを仮定する（Wilkinson (1963)）．

仮定 2. 数値は底 β, 仮数部 t 桁，指数部 e（整数）の浮動小数点数として

$$x = u \times \beta^e \quad (u \text{ は } \beta \text{ 進 } t \text{ 桁の小数} : \beta^{-1} \leqq u < 1) \tag{2.9}$$

の形で表現される．基本演算の結果は，正確な値を上述の浮動小数点数に"丸める"ことによって得られる．丸め方式としては，四捨五入（仮数部 $t+1$ 桁目が $\beta/2$ 以上ならば仮数部 t 桁目に1を加えた後，$t+1$ 桁目以下を切捨てる）と 切捨て（仮数部 $t+1$ 桁目以下を切捨てる）とを考える．

このとき，変数 v_i を計算する際の発生誤差 δv_i の分布幅は

$$|\delta v_i| \leqq m(v_i)\varepsilon \quad (\leqq |v_i|\varepsilon) \tag{2.10}$$

で見積ることができる．ここで，$m(v_i)$ は，v_i と同じ指数部を持つ最小の正の浮動小数点数，ε

表 1. 本論文で使用した計算機の数値表現および丸め方式.
Table 1. The floating-point representations and rounding methods employed in the computers for the experiments.

Computer		DEC VAX-11/780	HITAC M-280H
Operating System		UNIX 4.2BSD	VOS3
FORTRAN Compiler		F77	FORTRAN77
仮数部長	単精度	2進24桁	16進 6桁
	倍精度	2進56桁	16進14桁
丸め方式		四捨五入（0捨1入）	切捨て
Machine epsilon	単精度	$2^{-24} \fallingdotseq 6.0\times10^{-8}$	$16^{-5} \fallingdotseq 1.0\times10^{-7}$
	倍精度	$2^{-56} \fallingdotseq 1.4\times10^{-17}$	$16^{-13} \fallingdotseq 2.2\times10^{-16}$

はいわゆる machine epsilon（計算機上で $1+\varepsilon \neq 1$ なる最小の数）で，$\varepsilon=\beta^{1-t}/2$（四捨五入の場合）あるいは $\varepsilon=\beta^{1-t}$（切捨ての場合）である．本論文で取り扱う計算機 HITAC M-280H, DEC VAX-11/780 の数値表現，丸め方式を表1に示す．

仮定 1, 2 と式 (2.8), (2.10) より，関数 f の計算値に含まれる丸め誤差 Δf の大きさの限界として，次の評価式が得られる．

〈1〉絶対評価

$$|\Delta f| \leqq \varepsilon \sum_i \left|\frac{\partial f}{\partial v_i}\right| m(v_i) = A[f]\varepsilon, \tag{2.11}$$

ここで，

$$A[f] \equiv \sum_i \left|\frac{\partial f}{\partial v_i}\right| m(v_i). \tag{2.12}$$

式 (2.11) は，$|\Delta f|$ に対する絶対的な限界であるので，これを"絶対評価"と呼ぶ．しかし，絶対評価の限界が実際に達成されることはまれである．より現実的な評価のためには，式 (2.8) の中の δv_i が実際にどのように振舞うかの確率的なモデルを考える必要がある．そのために，次の仮定を置く．

仮定 3. 中間変数 v_i の発生誤差 δv_i は，

$$\text{平均 } \mathrm{E}[\delta v_i] = \begin{cases} 0 & \text{(四捨五入の場合)} \\ -\dfrac{\varepsilon}{2}\,\mathrm{sign}(v_i)m(v_i) & \text{(切捨ての場合)}, \end{cases} \tag{2.13}$$

$$\text{分布幅} = \begin{cases} 2m(v_i)\varepsilon & \text{(四捨五入の場合)} \\ m(v_i)\varepsilon & \text{(切捨ての場合)} \end{cases} \tag{2.14}$$

の互いに独立な一様分布に従う確率変数である．

仮定 1～3 の下では，関数 f の丸め誤差は，"多数の独立な同一の一様分布に従う確率変数 U_i の加重和" として，

$$\Delta f = \sum_i \frac{\partial f}{\partial v_i}\delta v_i = \sum_i w_i U_i \tag{2.15}$$

と表現される．式 (2.15) を"丸め誤差の確率モデル"と呼ぶことにする．この確率変数 Δf の分布形は，加重 $|w_i| = |\partial f/\partial v_i| \cdot (\delta v_i$ の分布幅$)$ $(i=1, \ldots)$ の性質で決まる．もし加重の中に極端に大きいものが一つあれば (2.15) の分布形は，一様分布に近く，同程度の大きさのものが多数あれば正規分布に近くなる．このとき，Δf の平均 $\mathrm{E}[\Delta f]$，分散 $\mathrm{V}[\Delta f]$，および，2つの関数 f, g の丸め誤差 $\Delta f, \Delta g$ の共分散 $\mathrm{Cov}[\Delta f, \Delta g]$ は以下のように書ける（Iri et al. (1988)）．

〈四捨五入の場合〉

$$\mathrm{E}[\Delta f] = \sum_i \frac{\partial f}{\partial v_i}\mathrm{E}[\delta v_i] = 0, \tag{2.16}$$

$$(2.17)\qquad \mathrm{V}[\Delta f]=\sum_i\left|\frac{\partial f}{\partial v_i}\right|^2\mathrm{V}[\delta v_i]=\frac{\varepsilon^2}{3}\sum_i\left|\frac{\partial f}{\partial v_i}\right|^2 m(v_i)^2,$$

$$(2.18)\qquad \mathrm{Cov}[\Delta f,\Delta g]=\sum_i\frac{\partial f}{\partial v_i}\frac{\partial g}{\partial v_i}\mathrm{V}[\delta v_i]=\frac{\varepsilon^2}{3}\sum_i\frac{\partial f}{\partial v_i}\frac{\partial g}{\partial v_i}m(v_i)^2.$$

〈切捨ての場合〉

$$(2.19)\qquad \mathrm{E}[\Delta f]=\sum_i\frac{\partial f}{\partial v_i}\mathrm{E}[\delta v_i]=-\frac{\varepsilon}{2}\sum_i\frac{\partial f}{\partial v_i}\mathrm{sign}(v_i)m(v_i),$$

$$(2.20)\qquad \mathrm{V}[\Delta f]=\sum_i\left|\frac{\partial f}{\partial v_i}\right|^2\mathrm{V}[\delta v_i]=\frac{\varepsilon^2}{12}\sum_i\left|\frac{\partial f}{\partial v_i}\right|^2 m(v_i)^2.$$

特に，$\mathrm{E}[(\Delta f)^2]^{1/2}$ を丸め誤差の確率評価と呼ぶことにする．具体的には，

〈2〉確率評価

$$(2.21)\qquad \mathrm{E}[(\Delta f)^2]=\mathrm{E}[\Delta f]^2+\mathrm{V}[\Delta f]=P[f]^2\varepsilon^2.$$

ここで，

$$(2.22)\qquad P[f]\equiv\left(\frac{\varepsilon^2}{3}\sum_i\left|\frac{\partial f}{\partial v_i}\right|^2 m(v_i)^2\right)^{1/2}\qquad\text{(四捨五入の場合)},$$

$$(2.23)\qquad P[f]\equiv\left(\frac{\varepsilon^2}{3}\sum_i\left|\frac{\partial f}{\partial v_i}\right|^2 m(v_i)^2+\frac{\varepsilon^2}{4}\sum_{i\neq j}\frac{\partial f}{\partial v_i}\frac{\partial f}{\partial v_j}\cdot\mathrm{sign}(v_i)\,\mathrm{sign}(v_j)m(v_i)m(v_j)\right)^{1/2}\qquad\text{(切捨ての場合)}.$$

四捨五入の場合確率評価は丸め誤差の標準偏差である．以上で導入した丸め誤差の確率モデルが現実の計算過程における丸め誤差の振舞いを十分に良く説明しうることを，以下の節で具体例を通じて実証する．以下では，これらの確率モデルに基づいて計算された平均，分散などの統計的諸量を "理論平均"，"理論分散" などと呼んで引用する．

3. 対象とする連立方程式

化学プロセス・シミュレーション・システム DPS（日本科学技術研修所(1980)）によって水・メタノール蒸溜塔（15 段）の平衡状態を求める際に現れる 108 元の非線形連立方程式を例題として取り上げる．

未知変数の個数＝108,
定数の個数＝214,
計算される中間変数の個数＝関数計算に要する基本計算ステップ数＝2851

で，中間変数のうち 108 個が関数値として返される．この蒸溜塔の概念図を図 1 に掲げる．

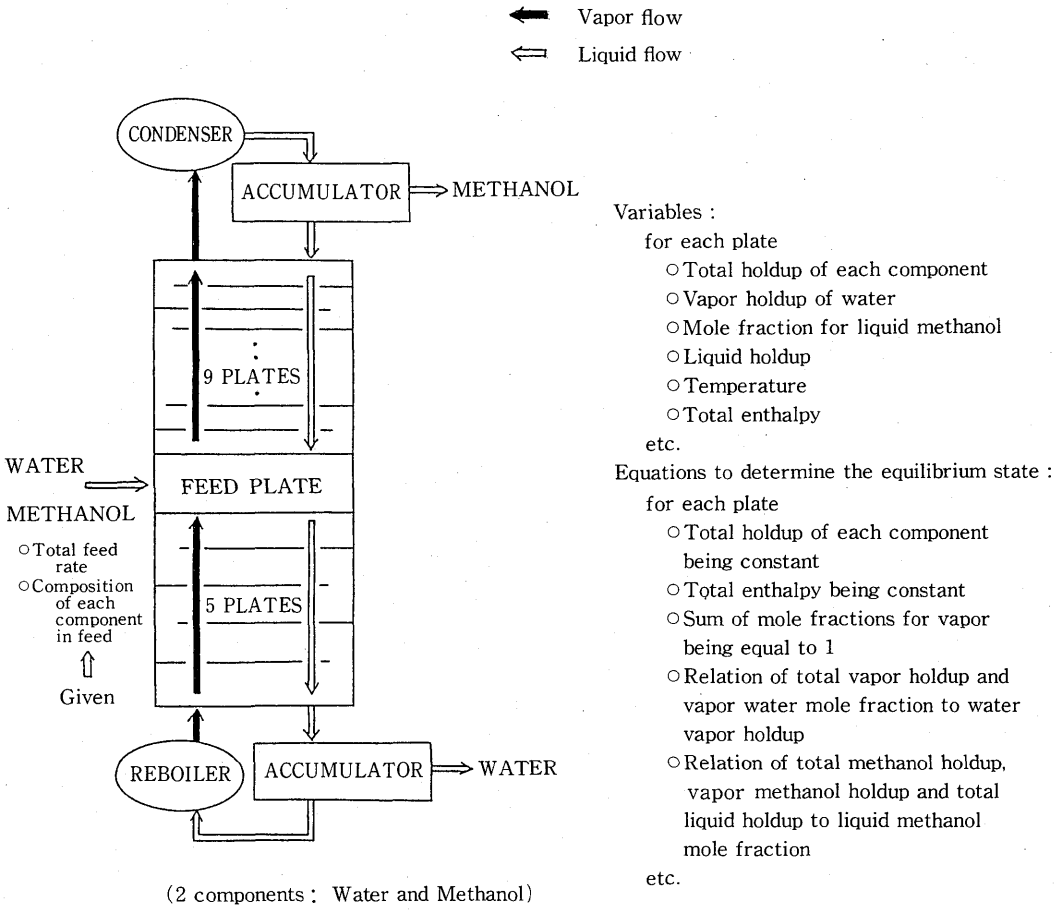


図1. 水・メタノール蒸溜塔の概念図.
Fig. 1. Schematic diagram of the water-methanol distillation tower.

4. 標本作成と丸め誤差の統計的諸量の実測

連立方程式を構成する108個の関数を，未知変数 $\boldsymbol{x}=(x_1, \ldots, x_{108})$ と定数 $\boldsymbol{c}=(c_1, \ldots, c_{214})$ から計算されるベクトル関数 $\boldsymbol{f}=(f_1, \ldots, f_{108})$ と見なして $\boldsymbol{f}(\boldsymbol{x}, \boldsymbol{c})$ と記し，$\boldsymbol{f}$ を単精度計算した結果を $\hat{\boldsymbol{f}}_S$，倍精度計算した結果を $\hat{\boldsymbol{f}}_D$ と記すことにする．連立方程式の解に近い2点 $(\boldsymbol{x}_1, \boldsymbol{c})$, $(\boldsymbol{x}_2, \boldsymbol{c})$ を選び，各点で（単精度計算時の）の丸め誤差の"標本"を作るために，以下のようなことを行った．まず，$(\boldsymbol{x}_l, \boldsymbol{c})$ $(l=1, 2)$ の値に乱数を用いて摂動を加えて

$$(4.1)\qquad (\boldsymbol{x}_l', \boldsymbol{c}') = (x_{l1}\times(1+\varepsilon_1), \ldots, x_{l108}\times(1+\varepsilon_{108}), c_1\times(1+\varepsilon_{109}), \ldots, c_{214}\times(1+\varepsilon_{322}))$$

（ε_i は区間 $[0, 1]\times 10^{-4}$ 上の一様乱数）

を作り，108個の関数値に含まれる丸め誤差 $\Delta\boldsymbol{f}$ の実測値を

$$(4.2)\qquad \Delta\hat{\boldsymbol{f}} = \hat{\boldsymbol{f}}_S(\boldsymbol{x}_l', \boldsymbol{c}') - \hat{\boldsymbol{f}}_D(\boldsymbol{x}_l', \boldsymbol{c}')$$

で定める；322個の乱数 ε_i を100回繰り返し選んで100個の丸め誤差標本 $\Delta\hat{\boldsymbol{f}}^{(1)}, \ldots, \Delta\hat{\boldsymbol{f}}^{(100)}$

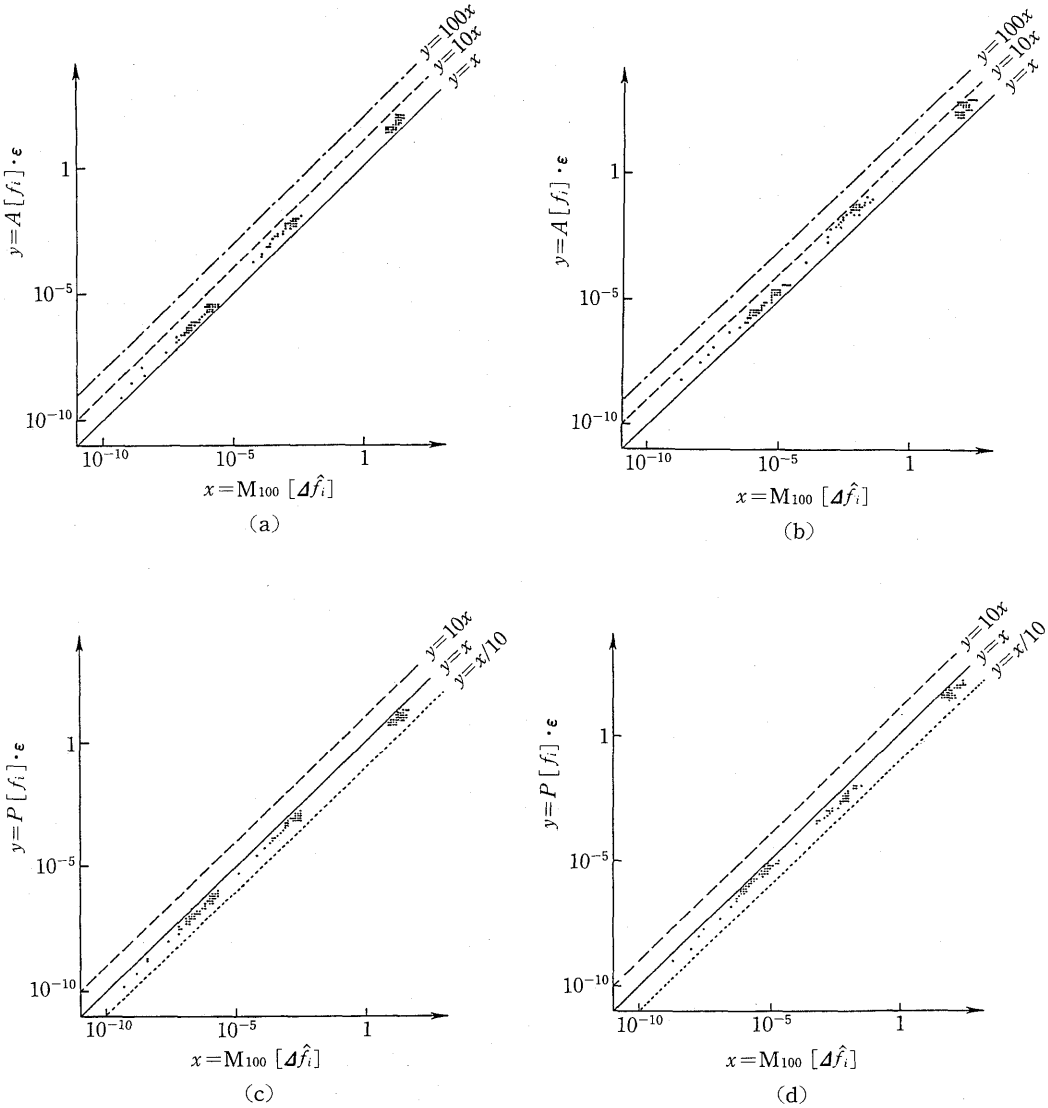


図 2. 108 個の関数の最大丸め誤差（100 個の標本）と評価値との関係（一点鎖線：評価値＝実測値×100，破線：評価値＝実測値×10，実線：評価値＝実測値，点線：評価値＝実測値/10）．(a) VAX-11/780, 絶対評価．(b) M-280H, 絶対評価．(c) VAX-11/780, 確率評価．(d) M-280H, 確率評価．

Fig. 2. Comparison of the observed maximum rounding errors (each among 100 samples) of the 108 functions with the theoretical estimates (chained line: estimate=(observed value)×100, broken line: estimate=(observed value)×10, solid line: estimate=observed value, dotted line: estimate=(observed value)/10). (a) VAX-11/780, absolute bound. (b) M-280H, absolute bound. (c) VAX-11/780, probabilistic bound. (d) M-280H, probabilistic bound.

を求める.

東京大学大型計算機センターの VAX-11/780, M-280H の上で，上記の手順に従って各関数 f_i につき 100 個の丸め誤差標本を作り，以下のようにして丸め誤差 Δf_i の統計的諸量の実測値を計算した.

〈四捨五入（VAX-11/780）の場合〉

標本分散 $V_S[\Delta \hat{f}_i]$:

$$V_S[\Delta \hat{f}_i]=\frac{1}{100}\sum_{x=1}^{100}(\Delta \hat{f}_i^{(x)})^2. \tag{4.3}$$

（(2.15) より $E[\Delta f_i]=0$ であるから，平均は 0 と見なしている.）

最大丸め誤差 $M_{10}[\Delta \hat{f}_i]$（10 個の最大値），$M_{100}[\Delta \hat{f}_i]$（100 個の最大値）:

$$M_{10}[\Delta \hat{f}_i]=\max_{x=1,\ldots,10}|\Delta \hat{f}_i^{(x)}|, \tag{4.4}$$

$$M_{100}[\Delta \hat{f}_i]=\max_{x=1,\ldots,100}|\Delta \hat{f}_i^{(x)}|. \tag{4.5}$$

標本共分散 $Cov_S[\Delta \hat{f}_i, \Delta \hat{f}_j]$:

$$Cov_S[\Delta \hat{f}_i, \Delta \hat{f}_j]=\frac{1}{100}\sum_{x=1}^{100}\Delta \hat{f}_i^{(x)}\Delta \hat{f}_j^{(x)}. \tag{4.6}$$

〈切捨て（M-280H）の場合〉

標本平均 $E_S[\Delta \hat{f}_i]$:

$$E_S[\Delta \hat{f}_i]=\frac{1}{100}\sum_{x=1}^{100}\Delta \hat{f}_i^{(x)}. \tag{4.7}$$

標本分散 $V_S[\Delta \hat{f}_i]$:

$$V_S[\Delta \hat{f}_i]=\frac{1}{100}\sum_{x=1}^{100}(\Delta \hat{f}_i^{(x)})^2-E_S[\Delta \hat{f}_i]^2. \tag{4.8}$$

以下では，これらの実測値と理論的な評価との比較を行う．結果のうちで，特に記していないものは，$(\boldsymbol{x}_1, \boldsymbol{c})$ における丸め誤差の標本によるものである.

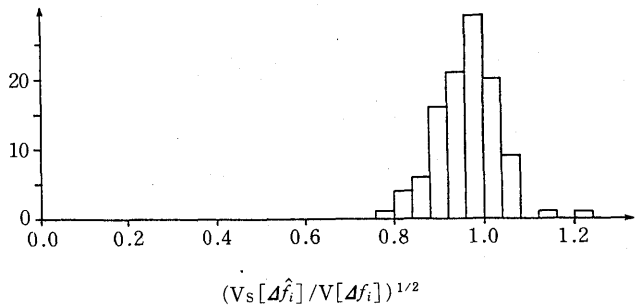


図 3. 108 個の関数の丸め誤差の "標本標準偏差/理論標準偏差" の分布（VAX-11/780）.
Fig. 3. Distribution of "(observed standard deviation)/(theoretical standard deviation)" of the rounding errors of the 108 functions (VAX-11/780).

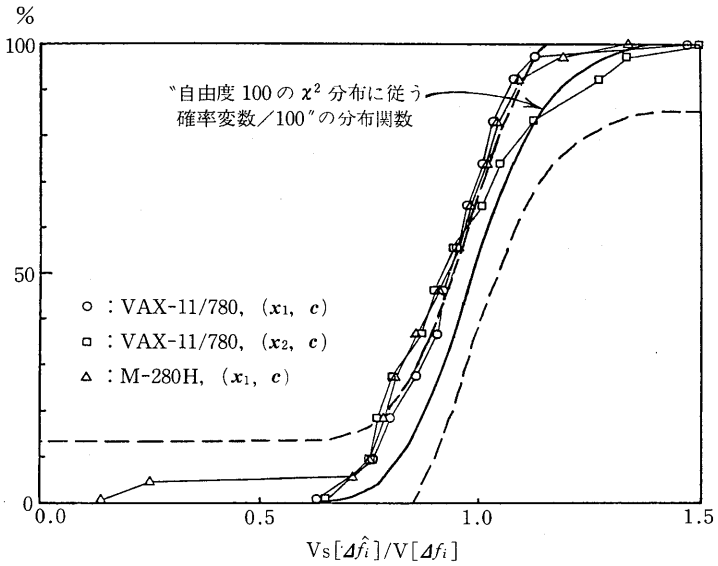


図 4.　108 個の関数の丸め誤差の〝標本分散/理論分散〞の累積分布曲線（VAX-11/780, M-280H）（破線は〝標本分散/理論分散〞が〝自由度 100 の χ^2 分布に従う確率変数/100〞の実現値であると仮定したときの Kolmogorov-Smirnov の検定統計量（両側検定）の分布の 95% 限界）.

Fig. 4.　Cumulative distribution curve of "(observed variance)/(theoretical variance)" of the rounding errors of the 108 functions (VAX-11/780, M-280H).

5.　108 個の関数の丸め誤差の実測値と確率モデルに基づく理論値の比較

108 個の関数のおのおのについて最大丸め誤差 $M_{100}[\Delta\hat{f}_i]$ と絶対評価 (2.11), 確率評価 (2.22), (2.23) との関係を示したのが図 2 である．注目すべきことは丸め誤差の大きさが関数ごとに極端に異なっていること，そして，108 個の点は，傾き 45° の実線の近くに細長く分布していることである．このことは，各関数の丸め誤差の理論的な評価値が，ほぼ最大丸め誤差と同程度の大きさであり，〝絶対評価〞，〝確率評価〞が十分に良い評価となっていることを示している．

以下では，実測した丸め誤差に関する統計的諸量と，第 2 節で述べた丸め誤差の確率モデルによって計算された統計的諸量の比較をより定量的に行い，確率モデルが実際の丸め誤差の挙動を十分良く説明していることを検証する．また，各関数の実際の丸め誤差の分布形が一様分布と正規分布の中間にあることも確かめる．

〈四捨五入（VAX-11/780）の場合〉

（1）　分散および標準偏差の評価：——108 個の関数のおのおのについて，実測された 100 個の丸め誤差標本の標本分散 (4.3) と理論分散 (2.17) の比 V_S/V を計算し，〝標本標準偏差/理論標準偏差〞$(V_S/V)^{1/2}$ のヒストグラムを作ったのが図 3 である．分布が 1.0 の近くに集中していることから丸め誤差の実測値の標準偏差が確率モデルによってほぼ正しく評価されていることが分かる．さらに詳しく調べるために，それら 108 個の比の累積分布曲線を描いたのが図 4 である．各関数の丸め誤差が正規分布に従うものと仮定すると，〝標本分散/理論分散〞V_S/V は〝自由度 100 の χ^2 分布に従う確率変数/100〞と同じ分布をす

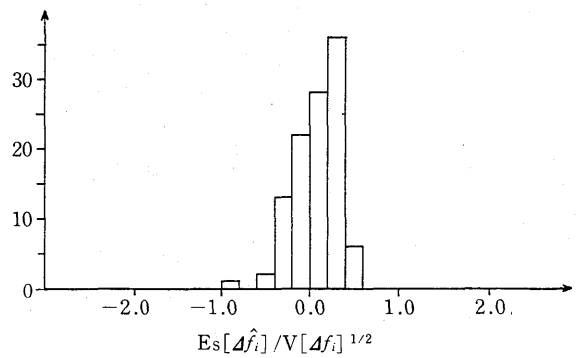


図 5. 108 個の関数の丸め誤差の "標本平均/理論標準偏差" の分布 (VAX-11/780).
Fig. 5. Distribution of "(observed mean)/(theoretical standard deviation)" of the rounding errors of the 108 functions (VAX-11/780).

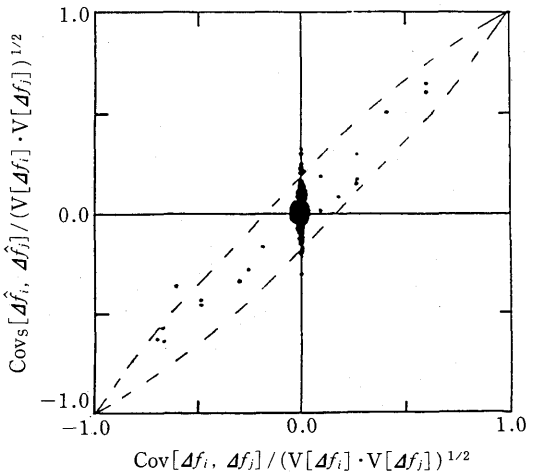


図 6. 丸め誤差の標本相関係数と理論相関係数との比較 (VAX-11/780).
Fig. 6. Comparison of the observed correlation coefficients with the theoretical of the rounding errors of the functions (VAX-11/780).

ると考えられる（正規分布からかなりずれていても近似的には同様）ので，後者の分布関数も比較のために描いてある．かなり粗い数学モデルに基づいた議論であるにもかかわらず，両者は概ね一致していることが観察される（図中破線は「"標本分散/理論分散" が "自由度 100 の χ^2 分布に従う確率変数/100" の実現値である」と仮定した時の Kolmogorov-Smirnov の検定統計量（両側検定）の分布の 95% 限界である）．このように，確率モデルによる標準偏差と実測値が十分精度良く一致していることが確かめられたので，以下で実測値をその標準偏差によって正規化する必要があるときは，理論標準偏差によって正規化することにする．なお，確率モデルによると四捨五入の場合，各関数の丸め誤差の理論平均は 0 であるが，この例題においては，108 個の "関数の丸め誤差の標本平均 (4.7) の理論標準偏差に対する比" の分布は図 5 のようになった．

(2) 共分散の評価：——2 つの関数の丸め誤差の共分散の理論評価 (2.18) の妥当性を

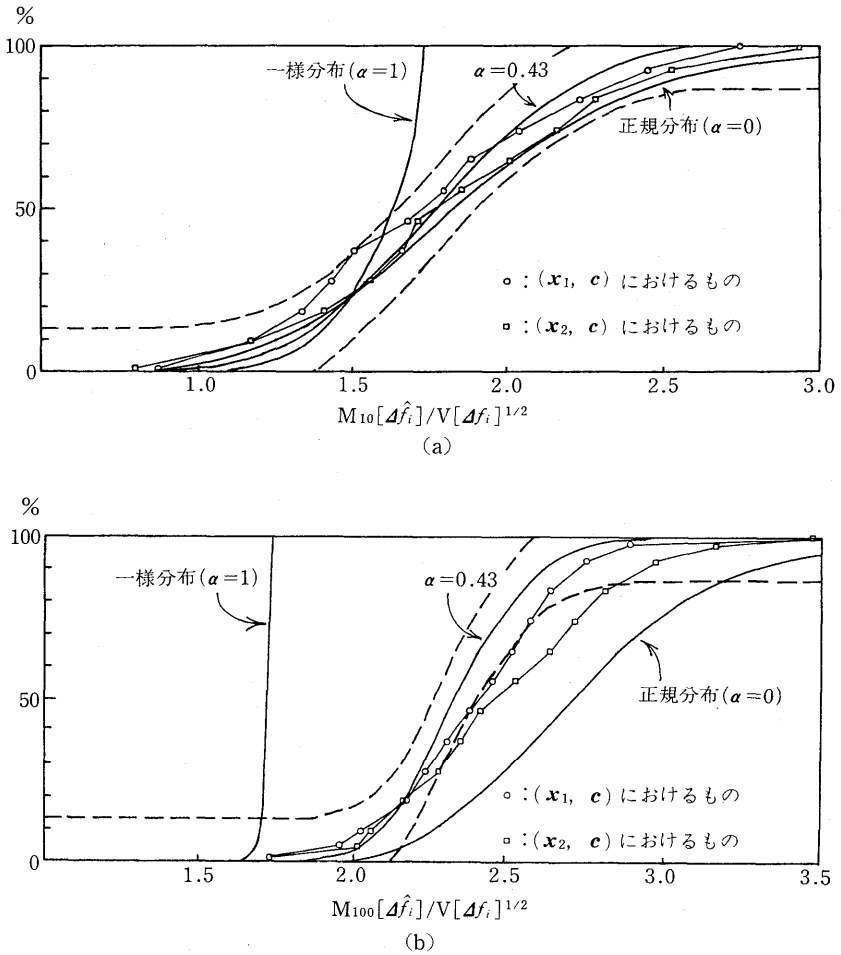


図 7. 108 個の関数の丸め誤差の "最大丸め誤差/理論標準偏差" の累積分布曲線（VAX-11/780）（破線は"標本数が 10 個（あるいは 100 個）の最大丸め誤差"が確率変数 $M_{10}[X_{0.43}]$（あるいは $M_{100}[X_{0.43}]$）に従って分布すると仮定したときの Kolmogorov-Smirnov の検定統計量（両側検定）の分布の 95% 限界）. (a) 標本数が 10 個の最大丸め誤差の場合. (b) 標本数が 100 個の最大丸め誤差の場合.

Fig. 7. Cumulative distribution curve of "(observed maximum rounding error)/(theoretical standard deviation)" of the rounding errors of the 108 functions (VAX-11/780). (a) observed maximum rounding errors among 10 samples. (b) observed maximum rounding errors among 100 samples.

検証するために，108 個の関数の丸め誤差の分散共分散行列（108 行×108 列）の上三角部分の 108×107/2＝5778個の要素から 200 個をランダムに選び，おのおのについて，理論相関係数

$$\frac{\mathrm{Cov}[\Delta f_i, \Delta f_j]}{(\mathrm{V}[\Delta f_i]\mathrm{V}[\Delta f_j])^{1/2}} \tag{5.1}$$

と実測値相関係数

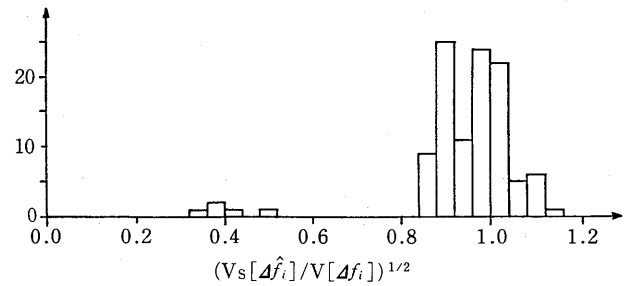


図 8. 108 個の関数の丸め誤差の "標本標準偏差/理論標準偏差" の分布 (M-280H).
Fig. 8. Distribution of "(observed standard deviation)/(theoretical standard deviation)" of the rounding errors of the 108 functions (M-280H).

$$\frac{\mathrm{Covs}[\Delta\hat{f}_i,\ \Delta\hat{f}_j]}{(\mathrm{V}[\Delta f_i]\mathrm{V}[\Delta f_j])^{1/2}} \tag{5.2}$$

を計算して比較したのが図 6 である．破線は，横軸の相関係数を持つ平均 0 の 2 次元正規分布に従って 2 つの関数の丸め誤差が分布していると仮定して，標本数 100 の場合の標本相関係数の分布の 95% 限界を結んだものである．ほとんどの点 (200 点中 183 点) がこの破線の中に入っており，理論値と標本値が良く合致していることが観察される．

(3) 最大丸め誤差：—— 標本数 10, 100 の 2 つの場合について，108 個の関数の "最大丸め誤差/理論標準偏差" の累積分布曲線を描いたのが図 7 である．丸め誤差の確率モデルによると，各関数の丸め誤差の分布形は，一様分布よりは裾が長く，正規分布よりは裾が短い分布となるはずである．そこで，比較のために，

(i) 108 個の関数の丸め誤差が全て独立な正規分布に従って分布すると仮定したときの "最大丸め誤差/標準偏差" の累積分布曲線 ($\alpha=0$ の曲線),

(ii) 同じく，丸め誤差が全て独立な一様分布に従って分布すると仮定したときの "最大丸め誤差/標準偏差" の累積分布曲線 ($\alpha=1$ の曲線)

も重ねて描いてある (α の意味と $\alpha=0.43$ の曲線については後述する)．標本数 10 の場合 (図 7(a)) も 100 の場合 (図 7(b)) も実測値の累積分布曲線がほぼ (i), (ii) の曲線の間を通っているのが見られる．

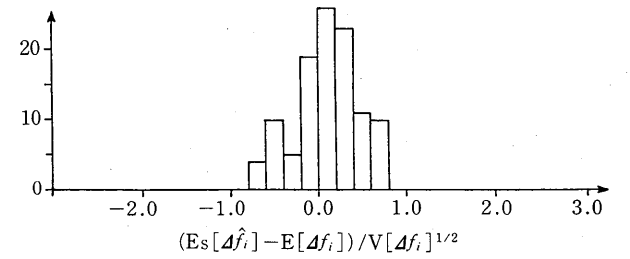


図 9. 108 個の関数の丸め誤差の "(標本平均−理論平均)/理論標準偏差" の分布 (M-280H).
Fig. 9. Distribution of "((observed mean)−(theoretical mean))/(theoretical standard deviation)" of the rounding errors of the 108 functions (M-280H).

〈切捨て（M-280）の場合〉

（1）　分散および標準偏差の評価：――108個の関数のおのおのについて，実測された100個の丸め誤差標本の標本分散（4.8）と理論分散（2.20）の比 V_S/V を計算し，"標本標準偏差/理論標準偏差" $(V_S/V)^{1/2}$ のヒストグラムを作ったのが図8である．いくつかの外れ値はあるものの，全体としては1.0付近に集中して分布しており，実測値の標準偏差が確率モデルによってほぼ正しく評価されていることが分かる．四捨五入の場合と同様の解析を行うために，108個の "標本分散/理論分散" の累積分布曲線を図4に描き加えた．各関数の丸め誤差が正規分布に従うものと仮定すると，"標本分散/理論分散" V_S/V は "自由度99の χ^2 分布に従う確率変数/99" と同じ分布をすると考えられる（これは，(4.8)が平均の実測値を使った標本分散であるため）．"自由度99の χ^2 分布に従う確率変数/99"の分布関数は既に描いてある "自由度100の χ^2 分布に従う確率変数/100" の分布関数とほぼ同じなので，その曲線と "標本分散/理論分散" の累積分布曲線を比較すると，四捨五入の場合と同様，両者は概ね一致していることが観察される．このように，切捨ての場合も確率モデルによる標準偏差と実測値が十分精度良く一致していることが確かめられたので，以下で実測値をその標準偏差によって正規化する必要があるときは，理論標準偏差によって正規化することにする．

（2）　平均の評価：――108個の関数のおのおのの丸め誤差 Δf_i について，標本平均(4.7)と確率モデルに基づく平均の理論値(2.19)との差を理論標準偏差によって正規化した量

$$\frac{E_S[\Delta \hat{f}_i]-E[\Delta f_i]}{V[\Delta f_i]^{1/2}} \tag{5.3}$$

の分布のヒストグラムが図9である．$E_S[\Delta \hat{f}_i]$ が100個の標本値の平均なので，確率モデルが正しければヒストグラムの分布の標準偏差は0.1程度になるはずである．実際に得られたヒストグラムの標準偏差は0.3程度で，予想されるよりも若干バラツキが大きいが，モデルおよび実験の精度を考慮すると，はなはだしくくい違っているとはいえないであろう．

以上のことから，第2節で述べた丸め誤差の確率モデルによって，各関数の丸め誤差の標準偏差や平均，共分散がほぼ正しく推定できており，$\Delta \hat{f}_i$ は(2.15)で表される確率変数の実現値と考えて差し支えないこと，108個の関数の丸め誤差の分布形は，一様分布よりは裾が長く，正規分布よりは裾が短いことが確かめられた．

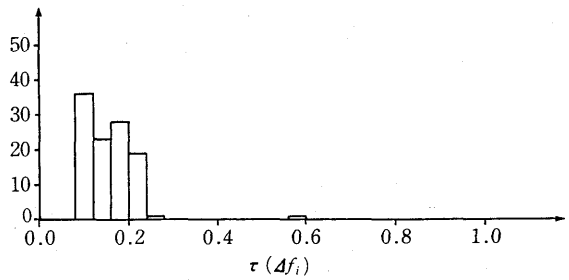


図10.　108個の関数の丸め誤差の "確率評価/絶対評価"（$=\tau(\Delta f_i)$）の分布（VAX-11/780）．
Fig. 10.　Distribution of "(probabilistic bound)/(absolute bound)" ($=\tau(\Delta f_i)$) of the rounding errors of the 108 functions (VAX-11/780).

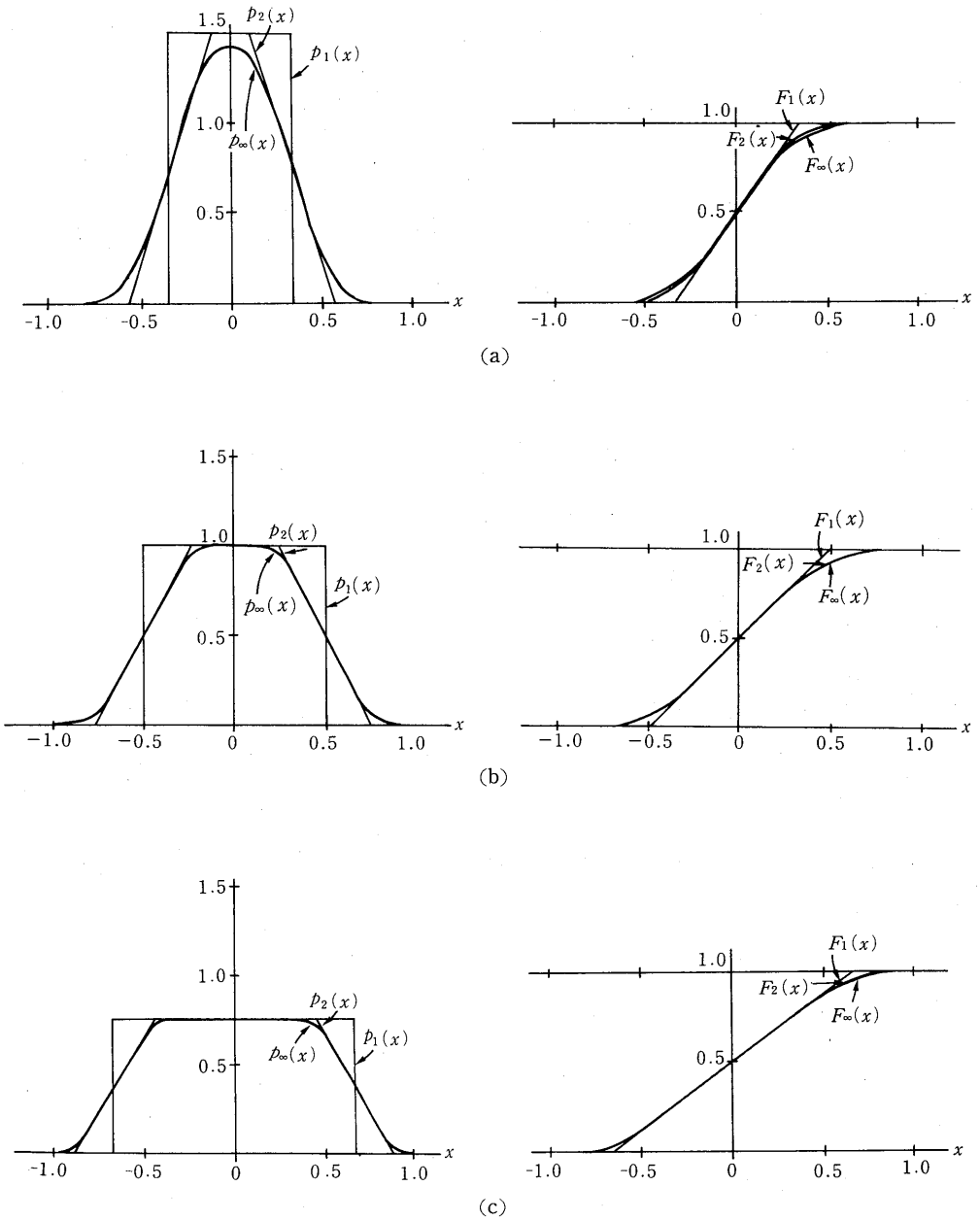


図 11. 丸め誤差の分布を近似する確率変数 (6.4) の分布関数および密度関数の形. ($F_n(x)$ および $p_n(x)$ は (6.4) の和を n 項目までで打ち切ったときの確率変数の分布関数および密度関数を表す) (Kabaya and Iri (1987a) より引用). (a) $\alpha=1/3$ の場合. (b) $\alpha=1/2$ の場合. (c) $\alpha=2/3$ の場合.

Fig. 11. The density function and the cumulative function of the random variable (6.4) to approximate the distribution of rounding errors. ($F_n(x)$ and $p_n(x)$ are the distribution function and the density function, respectively, of the random variable defined by truncating the summation (6.4) at the n-th term) (from Kabaya and Iri (1987a)). (a) $\alpha=1/3$, (b) $\alpha=1/2$, (c) $\alpha=2/3$.

6. 各関数の丸め誤差の分布形に関する考察とその類型化

本節では各関数の丸め誤差の分布について，より詳しい解析と類型化を行う．以下，四捨五入の場合に話を限ることにする．関数 f_i の丸め誤差の表現式 (2.15) における加重 $|\partial f_i/\partial v_j| \cdot m(v_j)\varepsilon$ $(j=1, \ldots)$ を大きさの順に並べ換えたものを $T[f_i, k]$ $(k=1, \ldots)$ と記すことにすると，(2.15) は次のように書き換えられる：

$$\Delta f_i = \sum_{k=1} T[f_i, k] U_k \tag{6.1}$$

($T[f_i, 1] \geqq T[f_i, 2] \geqq \cdots$; U_k は互いに独立な $[-1, 1]$ 上の一様分布). 関数 f_i の丸め誤差の分布が一様分布に近くなるかそれとも正規分布に近くなるかは，数列 $\{T[f_i, k]\}$ の性質による．$\{T[f_i, k]\}$ が，k が増加するにつれて急激に減少する場合，すなわち，(6.1) において，少数の大きな項だけが和に利いている場合，分布は一様分布に近く，$\{T[f_i, k]\}$ が，k が増加してもあまり速く減少しない場合，すなわち多数の同程度の大きさの項が和に寄与している場合，分布は正規分布に近くなると考えられる．式 (6.1) が与える確率変数の分布について，上述のような差を表現する量として

$$\tau(Y) = \frac{|Y| \text{の標準偏差}}{|Y| \text{のとりうる最大値}} \qquad (Y: \text{確率変数}) \tag{6.2}$$

なる統計量を考える．$\tau(Y)$ は，〝$|Y|$ がそのとりうる最大値に近い大きさの値をとる頻度の目安〟であるといえる．正規分布に従う確率変数に対しては $\tau=0$，そして，一様分布に従う確率変数では $\tau=(1/3)^{1/2}$ となる．式 (2.15) あるいは (6.1) で定義される丸め誤差の確率モデルの確率変数 Δf_i に対しては

$$\tau(\Delta f_i) = \frac{P[f_i]\varepsilon}{A[f_i]\varepsilon} = \frac{P[f_i]}{A[f_i]} \tag{6.3}$$

となる．図 10 は 108 個の関数の丸め誤差に関する τ の値の分布である．$\tau(\Delta f_i)$ の値が，0.08 から 0.24 付近に分布しているのが見られる．$\tau(\Delta f_i)$ の値の違いは，$|\Delta f_i|$ のその最大値 $A[f_i]\varepsilon$ に近い大きさの値の現れやすさの違いを反映している．

ここで，かなり大胆に，数列 $\{T[f_i, k]\}$ を $T[f_i, k] \approx T[f_i, 1]\ (1-\alpha)^{k-1}$ と近似することができると考えて，区間 $[-1, 1]$ 上に値をとる次の確率変数 X_α $(0<\alpha<1)$ を導入する：

$$X_\alpha \equiv \alpha \sum_{k=0}^{\infty} (1-\alpha)^k U_k. \tag{6.4}$$

この確率変数族に関しては，$0<\alpha<1$ でその密度関数，分布関数が存在することが証明されており，様々な性質が調べられている (Kabaya and Iri (1987a))．$\alpha=1/3, 1/2, 2/3$ の場合について，X_α の密度関数，分布関数の形を図 11 に示す．X_α は $\alpha=0$ では定義されないが，$\alpha \to 0$ のとき $\sqrt{3(2-\alpha)/\alpha}\ X_\alpha$ が標準正規分布に近づく．

いま，確率変数 X_α を関数の丸め誤差の分布形の雛型として採用し，各 Δf_i に対し適当な α を対応させ，丸め誤差の実測値 $\Delta \hat{f}_i$ を X_α に適当な尺度因子を掛けたものの実現値と見なすことにしてみる．確率変数 X_α に対しては

$$\tau(X_\alpha) = \sqrt{\frac{\alpha}{3(2-\alpha)}} \tag{6.5}$$

と書けるので，各関数の丸め誤差 Δf_i に対する α (以下 $\alpha[\Delta f_i]$ と記す) を，$\tau(\Delta f_i) = \tau(X_\alpha)$ となるように定めると，(6.3) より，

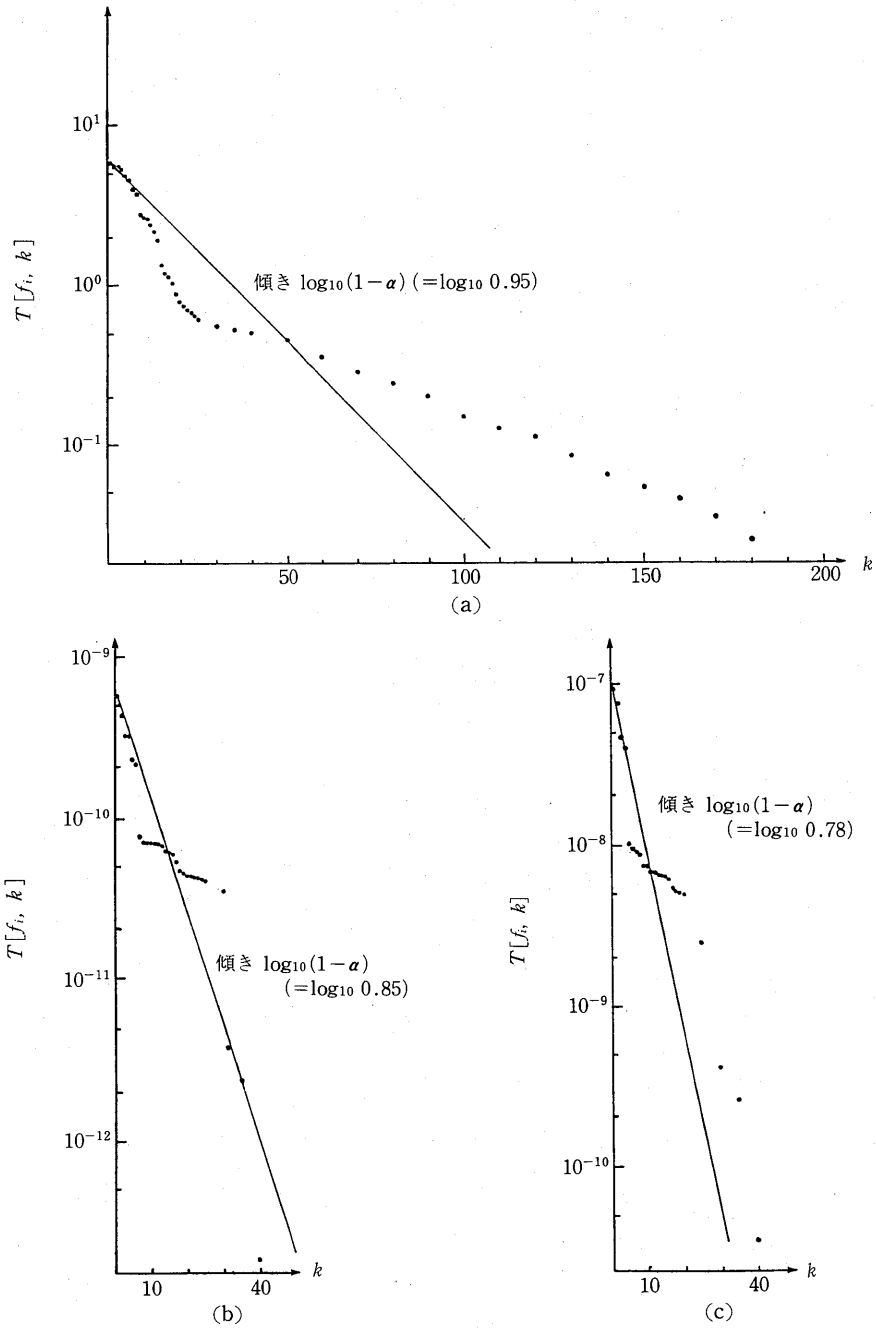
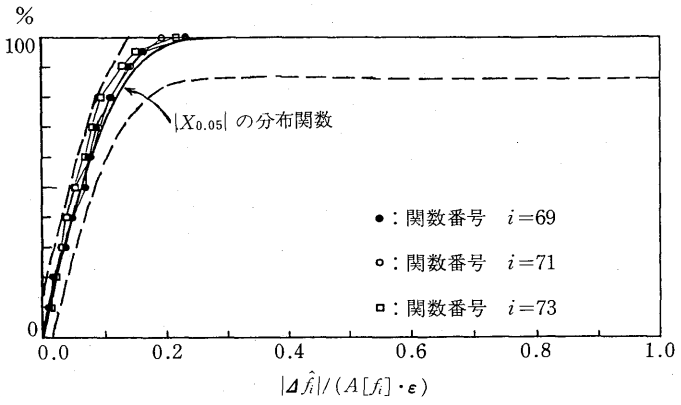
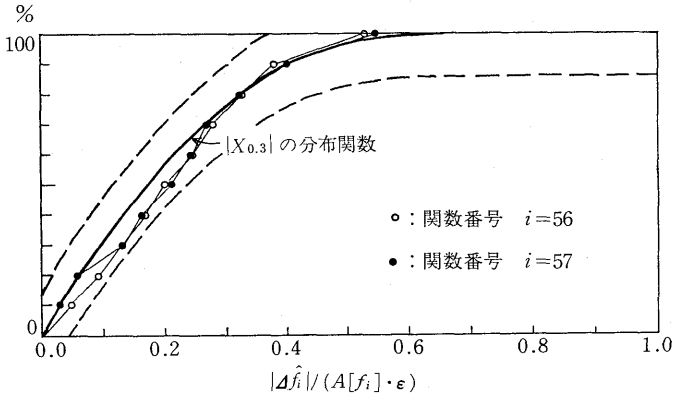


図 12. 関数の丸め誤差の理論式 (6.1) における U_k の係数の値の大きさ (VAX-11/780) (k は大きさの順位). (a) 関数番号 $i=90$, $\tau=0.095$, $\alpha=0.05$. (b) 関数番号 $i=19$, $\tau=0.16$, $\alpha=0.15$. (c) 関数番号 $i=9$, $\tau=0.20$, $\alpha=0.22$.

Fig. 12. The magnitudes of the coefficients of U_k in the probabilistic model of the rounding errors of functions (6.1) (VAX-11/780) (k is the rank in magnitude). (a) Function number $i=90$, $\tau=0.095$, $\alpha=0.05$. (b) Function number $i=19$, $\tau=0.16$, $\alpha=0.15$. (c) Function number $i=9$, $\tau=0.20$, $\alpha=0.22$.



(a)



(b)

図 13. 個々の関数の丸め誤差標本の累積分布曲線と $|X_\alpha|$ の分布関数との比較(VAX-11/780)(破線は丸め誤差標本が $|X_\alpha|$ の実現値であると仮定したときの Kolmogorov-Smirnov の検定統計量(両側検定)の分布の 95% 限界). (a) $\tau(\Delta f_i)\fallingdotseq 0.09$ ($\alpha[\Delta f_i]\fallingdotseq 0.05$) となる Δf_i の場合. (b) $\tau(\Delta f_i)\fallingdotseq 0.24$ ($\alpha[\Delta f_i]\fallingdotseq 0.3$)となる Δf_i の場合.

Fig. 13. Comparison of the cumulative distribution curve of observed rounding errors of each function with the distribution function of $|X_\alpha|$ (VAX-11/780). (a) Δf_i with $\tau(\Delta f_i)\fallingdotseq 0.09$ ($\alpha[\Delta f_i]\fallingdotseq 0.05$). (b) Δf_i with $\tau(\Delta f_i)\fallingdotseq 0.24$ ($\alpha[\Delta f_i]\fallingdotseq 0.3$).

$$\alpha[\Delta f_i]=\frac{6\tau(\Delta f_i)^2}{1+3\tau(\Delta f_i)^2}=\frac{6P[f_i]^2}{A[f_i]^2+3P[f_i]^2} \tag{6.6}$$

を得る. このようにして, パラメータ α を決定することの妥当性を調べるために, $\tau(\Delta f_i)$ の異なる3つの関数について, $\log_{10}T[f_i, k]$ ($k=1, 2, \ldots$) と, (6.6) によって定められる傾き $\log_{10}(1-\alpha)$ の直線をプロットしてみたのが図12である. 式(6.6)によって定めた α を用いた数列 $\{(1-\alpha)^{k-1}\ T[f_i, 1]\}$ が数列 $\{T[f_i, k]\}$ の主要な部分の減少の様子をかなり良く代表していることが分かる. このことは, $\tau(\Delta f_i)$ の値が, $\{T[f_i, k]\}$ の "減衰率" に関する情報を集約した良いパラメータであり, (6.6) によって α を定めれば, 略々

$$\frac{\Delta f_i}{A[f_i]\varepsilon}=\frac{1}{A[f_i]\varepsilon}\sum_{k=0}^{\infty}T[f_i, k+1]U_k\approx\alpha\sum_{k=0}^{\infty}(1-\alpha)^kU_k=X_\alpha \tag{6.7}$$

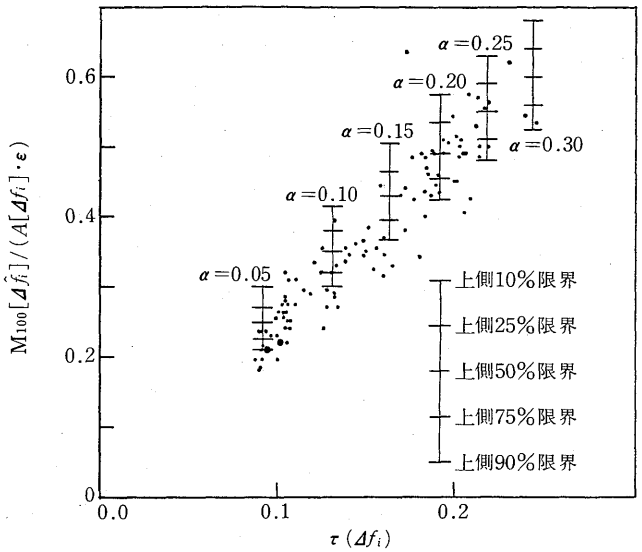


図 14. 108個の関数の丸め誤差の $\tau(\Delta f_i)$ (= "確率評価/絶対評価") と "最大丸め誤差/絶対評価" の関係 (VAX-11/780).

Fig. 14. Relation between $\tau(\Delta f_i)$ (="(probabilistic bound)/(absolute bound)") and "(observed maximum rounding error)/(absolute bound)" of the rounding errors of the 108 functions (VAX-11/780).

が成立し，X_α によって各関数の丸め誤差の分布の特徴を良く捉えられることを示している．108個の関数の丸め誤差に対する τ の値は 0.08～0.24 に分布しているが，(6.6) より，対応する α の値は 0.05～0.3 となる．

以上の考察に基づいて，各関数の丸め誤差の分布の違い，そして，$X_{\alpha[\Delta f_i]}$ がどの程度 $\Delta\hat{f}_i/(A[f_i]\varepsilon)$ の分布の特徴を捉えているかについて調べてみる．

(1) 各関数ごとの $|X_{\alpha[\Delta f_i]}|$ の分布関数と，100個の丸め誤差標本の累積分布曲線の比較：—— 108個の関数の丸め誤差の中で，

(i) 正規分布に近い，$\tau(\Delta f_i)\fallingdotseq 0.09$ (すなわち，$\alpha\fallingdotseq 0.05$) となる，$\Delta f_i$ と

(ii) 一様分布に近い，$\tau(\Delta f_i)\fallingdotseq 0.24$ (すなわち，$\alpha\fallingdotseq 0.3$) となる，Δf_i と

をいくつか選び，それぞれ100個の Δf_i の標本に対する $|\Delta\hat{f}_i|/(A[f_i]\varepsilon)$ の累積分布曲線と $|X_\alpha|$ の分布関数とを図13に描いた．まず，(i)，(ii)の Δf_i の累積分布曲線を比較すると，両者の差異は明瞭である — 正規分布により近い (i) の場合の方が累積分布曲線が急激に立ち上がっている — ことが分かる．このことは，関数の丸め誤差の分布には様々なタイプがあることを示している．一方，(i)，(ii)どちらの場合も，$|\Delta\hat{f}_i|/(A[f_i]\varepsilon)$ の累積分布曲線と (6.6) で定められた $|X_\alpha|$ の分布関数が十分に良く一致していることが観察される．これは，$|X_{\alpha[\Delta f_i]}|$ が $|\Delta\hat{f}_i|/(A[f_i]\varepsilon)$ の十分に良い近似になっていることを示している．

(2) 最大丸め誤差の累積分布曲線による比較：—— 第5節では，実際の108個の "最大丸め誤差/標準偏差" の累積分布曲線が，"全ての関数の丸め誤差が一様分布に従うと想定した場合" と "全ての関数の丸め誤差が正規分布に従うと想定した場合" のほぼ中間を通っていることを確かめた．図10を見ると，各関数の丸め誤差は，$\tau=0.3$ よりは小さいことがわかる．式 (6.6) によると，$\tau\leqq 0.3$ ならば $\alpha\leqq 0.43$ であるから，各関数の丸め誤差の分布は，$X_{0.43}$ よりも正規分布に近いはずである．そこで，108個の関数の丸め誤差が，$X_{0.43}$ に従って分布するとした

ときの "最大丸め誤差/標準偏差" の累積曲線を図7に描き加えてみた．標本数＝10，100の各場合について，実際の108個の "最大丸め誤差/標準偏差" の累積分布曲線が "全ての関数の丸め誤差が $X_{0.43}$ に従って分布すると想定した場合" と "全ての関数の丸め誤差が正規分布に従って分布すると想定した場合" の間を通っており，この結果は X_α が丸め誤差の実際を良く捉えていることを示している．

(3) "最大丸め誤差/絶対評価" の分布による比較：—— $\Delta\hat{f}_i$ を(2.15)あるいは(6.1)に従う Δf_i の実現値と考えるならば，τ が大きい Δf_i ほど，$|\Delta\hat{f}_i|/(A[f_i]\varepsilon)$ の値も大きくなる傾向があると考えられる．そこで，$\tau(\Delta f_i)$ の値を横軸に，"(標本数100の場合の確率変数の標本最大値)/(確率変数がとりうる値の最大値)" を縦軸にとり，108個の関数について，$M_{100}[\Delta\hat{f}_i]/(A[f_i]\varepsilon)$ をプロットしたのが図14である．108個の点が左下から右上にかけて分布しており，"τ の値が大きい Δf_i ほど，$|\Delta f_i|$ がとりうる最大値に比して相対的に大きい値が出やすい" 傾向が観察される．このことは，定性的な意味で，各関数の丸め誤差の実測値が(2.15)あるいは(6.1)の分布に従うことの検証になっている．

また，さらに，$\Delta\hat{f}_i$ を(6.7)の $X_{\alpha[\Delta f_i]}$ の実現値と見なしうるならば，$M_{100}[\Delta\hat{f}_i]/(A[f_i]\varepsilon)$ の値は，"(標本数100の場合の $|X_{\alpha[\Delta f_i]}|$ の標本最大値)/($|X_{\alpha[\Delta f_i]}|$ のとりうる最大値)" の分布に従うはずである．これを検証するために，図14に，併せて $\alpha=0.05, 0.1, 0.15, 0.20, 0.25, 0.30$ について"(標本数100の場合の $|X_{\alpha[\Delta f_i]}|$ の標本最大値)/($|X_{\alpha[\Delta f_i]}|$ のとりうる最大値)" の分布の上側10%点，25%点，50%点，75%点，90%点を描き加えてある．実測値の最大値の分布範囲が対応する $|X_{\alpha[\Delta f_i]}|$ の最大値の分布範囲に比べて全体として若干下に偏っているが，大部分重なっており，おのおのの関数の丸め誤差の分布の特徴が $X_{\alpha[\Delta f_i]}$ によって十分に良く捉えられていることが分かる．

以上のことから，"各 $\Delta\hat{f}_i$ が Δf_i の可能な最大値 $A[f_i]\varepsilon$ に近い大きさの値をとる頻度の違い" に着目することによって関数の丸め誤差の分布形の差異が明らかにできること，そして，それらの特徴が(2.15)あるいは(6.1)をさらに単純化して得られる確率変数 $X_{\alpha[\Delta f_i]}$ によってかなり良く捉えられることが検証された．

7. ま と め

本論文では，高速自動微分法において丸め誤差推定に用いられている丸め誤差の確率モデルの検証を中心に，実用規模の例題における丸め誤差の振舞いを調べた．そして，

(1) 計算された関数値に含まれる丸め誤差を"多数の独立な一様分布に従う確率変数の加重和"として表される確率変数の実現値であると見なす確率モデルによって，実際の丸め誤差の持つ統計的諸性質を定量的にもかなり良く説明できること；
(2) 各関数の丸め誤差の分布形の違いが "絶対評価に近い大きさの丸め誤差の値の出やすさの違い" として把握できること；
(3) (1)で述べた確率モデルをさらに単純化して得られる確率変数 "分布幅が等比級数的に減少する互いに独立な一様分布に従う確率変数の無限和"は，関数の計算値に含まれる丸め誤差の分布形の良いモデルになっており，実際の丸め誤差の分布は，等比級数の公比を適当に定めることによって，その特徴が十分に良く捉えられること

を明らかにした．

第6節で導入した確率変数 X_α およびその分布関数，密度関数に関しては，それら自身が多く

の興味深い数学的性質を持つことが明らかにされており，数値積分などへの応用の研究も進められている（Kabaya and Iri (1987a, 1987b)；Moriguti et al. (1987)；西井 他 (1987)）．

本論文で用いた評価式は，発生誤差 δv_i に関する高次の項を無視している，影響係数を計算するときの丸め誤差が考慮に入れられていない，等の点において厳密なものではないが，高速自動微分法を用いた精密かつ厳密な誤差評価法の研究も進められている（久保田，伊理(1987)）．

今後の課題としては，線形計算など様々な計算過程における丸め誤差の振舞いの解析，そして，丸め誤差の情報を積極的に数値計算の制御に取り入れたアルゴリズムの工夫，等々があげられる．

謝　辞

日頃研究全般にわたってご相談にのっていただいている統計数理研究所 田辺國士教授，そしてご助言をいただいた同研究所 安芸重雄助手に心より感謝いたします．

参 考 文 献

Alefeld, G., and Herzberger, J. (1985). *Introduction to Interval Computations*, Academic Press, New York.

Iri, M. (1984). Simultaneous computation of functions, partial derivatives and estimates of rounding errors — Complexity and practicality, *Japan Journal of Applied Mathematics*, **1**, 223-252.

伊理正夫，久保田光一 (1986). 高速微分法とその応用，第7回数理計画シンポジウム論文集，159-184.

伊理正夫，土谷 隆，星 守 (1985). 偏導関数計算と丸め誤差推定の自動化の大規模非線形方程式系への応用，情報処理，**26**, 1411-1420.

Iri, M., Tsuchiya, T., and Hoshi, M. (1988). Automatic computation of partial derivatives and rounding error estimates with applications to large-scale systems of nonlinear equations, Research Memorandum RMI 88-01, Department of Mathematical Engineering and Information Physics, University of Tokyo (to appear in *J. Comput. Appl. Math.*, **23**, 1988).

Kabaya, K., and Iri, M. (1987a). Sum of uniformly distributed random variables and a family of nonanalytic C^∞-functions, *Japan Journal of Applied Mathematics*, **4**, 1-22.

Kabaya, K., and Iri, M. (1987b). On operators defining a family of nonanalytic C^∞-functions, Research Memorandum RMI 87-01, Department of Mathematical Engineering and Instrumentation Physics, University of Tokyo (to appear in *Japan Journal of Applied Mathematics*, **5**, 1988).

久保田光一，伊理正夫 (1987). 高速自動微分法と区間解析とを用いた丸め誤差推定の厳密化，応用数学合同シンポジウム研究報告集，1987年12月21日-23日，325-336.

Moriguti, S., Iri, M., and Kabaya, K. (1987). On asymptotic properties of the eigenfunctions of a linear operator Φ^*: $q(x) \mapsto \frac{1}{2\alpha}\int_{(1-\alpha)x-\alpha}^{(1-\alpha)x+\alpha} q(\xi)d\xi$, Research Memorandum RMI 87-04, Department of Mathematical Engineering and Information Physics, Faculty of Engineering, University of Tokyo.

日本科学技術研修所(1980). DPS (V2)利用者マニュアル，情報処理振興事業協会，日本科学技術研修所，東京．

西井 修，室田一雄，伊理正夫 (1987). 非解析的な関数を用いた変数変換型数値積分公式，情報処理学会論文誌，**28**, 799-806. (英語版：Nishii, O., Murota, K., and Iri, M. (1987). A quadrature formula using a nonanalytic transformation of the integration variable, Research Memorandum RMI 87-03, Department of Mathematical Engineering and Instrumentation Physics, University of Tokyo).

土谷 隆 (1986). 高速微分法および丸め誤差推定法とその応用，東京大学大学院工学系研究科計数工学専門課程修士論文．

Wilkinson, J. H. (1963). *Rounding Errors in Algebraic Processes*, Prentice-Hall Inc., Englewood Cliffs, New Jersey.

Analysis of Rounding Errors in Large Systems of Nonlinear Equations

Takashi Tsuchiya
(The Institute of Statistical Mathematics)

Masao Iri
(Faculty of Engineering, University of Tokyo)

In this paper we analyze the behavior of rounding errors in computing complicated functions in a large system of nonlinear equations. We investigate theoretical properties as well as the adequacy of the probabilistic model of rounding errors based on the following two assumptions:

(i) The rounding error incurred in the computed value of a function is represented as the weighted sum of the rounding errors, each generated at a step in the computational procedure for the function value, with the partial derivative of the function with respect to the intermediate variable corresponding to the step as the weight;

(ii) the generated rounding error at each step of the procedure can be regarded as an instance of a uniformly distributed independent random variable where the width of distribution is determined by the result of the computational step and the floating-point representation employed.

The adequacy of the model is checked through many computational experiments.

We take as an example a system of nonlinear equations with 108 variables for the equilibrium state of a water-methanol distillation tower in a chemical plant. The fast automatic differentiation algorithm, which is an efficient and exact method for computing the gradient of a function with a number of variables, is employed to evaluate the theoretical estimates of rounding errors. The estimated statistical parameters of the rounding errors given by the proposed model are compared with the experimental values and are shown to be in good agreement. In order to characterize the distribution of rounding errors of each function, we also introduce a one-parameter family of random variables defined by the weighted sum of an infinite number of independent random variables obeying the uniform distribution with a geometrical progression as the weights. Comparison with the experimental values of the rounding errors shows that the one-parameter family is a plausible model for the distribution of the rounding errors in computing complicated functions.

Key words: Fast automatic differentiation, rounding errors, probabilistic model.