数理解析研究所講究録
第 42 巻 1968 年 23-35

計数型計算の特性について

日大理工　宇野利雄

計数型計算の実行に当って大きな影響力を持つ「桁落ち」について以下に若干の考察を試みる．桁数を指定された浮動小数点式計算において、計算が乗除算のみならば桁落ちはおこらない．加減算が入っていと同じ程度の大きさの２数の差を作るときに桁落ちがおこる．これによって計算が不正確になる原因は通常次のような場合であることが多い．

２数 a, b はともに小数 ν 位にはじまる有効 k 桁の正確数である．$a-b$ を必要とする数であるとする場合、この計算がそれまでの経過から $(M+a)-(M+b)$ のような形になって行なわれるとする．このとき M が a, b にくらべて非常に大きい小数 μ 位 $(\mu>\nu)$ からはじまる k 桁の数であったとしよう．いかなる計算も「有効 k 桁で実行する」とすると、引き算 $(M+a)-(M+b)$ の実行のとき、$a-b$ の k 桁の情報のうちこの引き算に参加するのは最初の $[k-(\mu-\nu)]$ 桁のみとなり，後部 $(\mu-\nu)$ 桁の情報は脱落してしまう．ここに計算の不正確性が発生するのである．

さてかようにして発生した情報の脱落は計算全体にいかなる影響力を持つであろうか. 計算の構造ともからみあって多様な特徴を示すことの例を以下の2つの場合についてしらべてみる.

才一はごく単純な桁落ちによる誤差発生だけの事柄である. その適例を分散、共分散の計算において見る.

§1. 分散、共分散の計算.

分散、共分散は定義は

$$(1) \qquad S_{xx}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2$$

$$(2) \qquad S_{xy}=\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})$$

であるが、各標本値を最後まで記憶にとどめておくのを必要としない簡単さのために通例次式によって計算している.

$$(3) \qquad S_{xx}=\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2$$

$$(4) \qquad S_{xy}=\frac{1}{n}\sum_{i=1}^{n}x_i y_i-\bar{x}\bar{y}.$$

(1), (2) は序にのべた $a-b$ の形式のものに、(3), (4) は特に標本の変動係数の小さい場合、多分に $(M+a)-(M+b)$ の形式のものに該当している. そのしくみをしらべるために、はじめにまず2数 A, B が接近しているとき $A-B$ における有効桁数の喪失が $|A/(A-B)| \fallingdotseq |B/(A-B)|$ ではから

れることを見ておく．すなわちこれらのものの整数部の桁数をmとすると、|A−B|の頭の数字が|A|の頭の数字より大なるか又は小なるかによって喪失の桁数はm又はm−1である．次例でこれがうかがえるであろう．

$$\begin{array}{r} 21.7652 \\ 21.6841 \\ \hline 00.0811 \end{array} (- \qquad \begin{array}{r} 3.14286 \\ 3.14159 \\ \hline 0.00127 \end{array} (-$$

$$21.7/8.11\times10^{-2}=267 \qquad 3.14/1.27\times10^{-3}=2455$$

上の法則をあてはめれば、(3)による分散の計算では

$$(5) \qquad \bar{x}^2\Big/\Big(\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2\Big),$$

(4)による共分散の計算では

$$(6) \qquad \bar{x}\bar{y}\Big/\Big(\frac{1}{n}\sum_{i=1}^{n}x_iy_i-\bar{x}\bar{y}\Big)$$

によって、それぞれ桁落ちの桁数がはかられる．ところでx_i, y_iなどは確率変数の実現値であって定数ではない．そこで(5), (6)はx_i, y_iの統計的分布によって推定する必要がある．(5)の場合、分子分母の期待値の計算によって

$$\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2=\sigma_x^2+O\Big(\frac{1}{\sqrt{n}}\Big),$$

$$\bar{x}^2=\mu_x^2+O\Big(\frac{1}{\sqrt{n}}\Big)$$

となることがわかるゆえ、

$$(7) \qquad \bar{x}^2\Big/\Big(\frac{1}{n}\sum_{i=1}^{n}x_i^2-\bar{x}^2\Big)=\frac{1}{V_x^2}+O\Big(\frac{1}{\sqrt{n}}\Big)$$

となる．ここにμ_x, σ_x^2, $V_x=\sigma_x/\mu_x$はそれぞれxの

母平均、母分散および母変動係数である．すなわち

(3)式による分散の計算のときには桁落ちの桁数が変動係数の2乗の逆数の整数部またはそれより1だけ少ない程度であることがわかる．

(4)による共分散の計算のときにも $Var(xy) \neq 0$ ならば、上と同様の考察が成立つが、今度のとき特に注意すべきは $Var(xy)=0$ の場合である．このときは(6)の分母において $E(\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y})=0$ となるので、さらに2次積率の計算により

$$\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y} = \frac{1}{\sqrt{n}}\sigma_x\sigma_y + O\left(\frac{1}{n}\right)$$

$$\bar{x}\bar{y} = \mu_x\mu_y + O\left(\frac{1}{\sqrt{n}}\right)$$

が得られるので

$$(8) \quad \bar{x}\bar{y}\Big/\left(\frac{1}{n}\sum_{i=1}^{n} x_i y_i - \bar{x}\bar{y}\right) = \frac{\sqrt{n}}{V_x V_y}\left(1+O\left(\frac{1}{\sqrt{n}}\right)\right)$$

となる．ここに V_x, V_y はそれぞれ x, y の母集団変動係数である．すなわち

x, y 無相関の場合、共分散を(4)式で計算すれば $\sqrt{n}/(V_x V_y)$ の整数部又はそれより1少ない程度の桁数だけ桁落ちのおこることがわかる．

前のときとちがって桁落ちが $\sqrt{n}$ の程度で n とともにふえ

る可能性がおこる．

以上のような共分散、分散の桁落ちがおこったとすると，それが当然相関係数の計算にひびく．特に母集無相関の場合、これら桁落ちによって幻影的相関係数を得る可能性のあることをも一応戒心しておく必要がある．

以下の例は上の考察を説明するためのものである　例1，2，3 は有効8桁、例4は有効5桁の計算によった．なお x，y はいずれも同じ母平均、母分散を持つ正規母集団からの標本をとった．主データは(3)，(4)による計算、わきに書いた括弧内の数字は(1)，(2)による計算である．

例1.　$N(20,\ 0.5^2)$，　$\dfrac{1}{V_x} = \dfrac{1}{V_y} = 40$

	$n=50$	$n=500$
S_{xx}	2.4857000 −01 (2.4858484 −01)	2.2947000 −01 (2.2940212 −01)
S_{yy}	2.2262000 −01 (2.2263552 −01)	2.4578000 −01 (2.4588692 −01)
S_{xy}	−3.6270000 −02 (−3.6287546 −02)	5.2400000 −03 (5.2411290 −03)
r_{xy}	−1.5418464 −01 (−1.5424925 −01)	2.2064550 −02 (2.2067768 −02)

無相関の場合（以下の例も皆そうである）
r_{xy}＝標本相関係数

例2. $N(200, 0.5^2)$,　$\frac{1}{V_x} = \frac{1}{V_y} = 400$

	$n=500$	$n=1000$
S_{xx}	2.1300000 −01	2.6600000 −01
	(2.2940212 −01)	(2.3441126 −01)
S_{yy}	2.2800000 −01	2.0700000 −01
	(2.4588662 −01)	(2.5526385 −01)
S_{xy}	−1.8000000 −02	−2.8000000 −02
	(5.2410708 −03)	(−1.3905698 −02)
r_{xy}	−8.1679916 −02	−1.1932516 −01
	(2.2067537 −02)	(−5.6847176 −02)
		相関係数が少々おかしくなっている.

例3. $N(2000, 0.5^2)$,　$\frac{1}{V_x} = \frac{1}{V_y} = 4000$

	$n=100$	$n=400$
S_{xx}	9.0000000 −01	13.000000 −01
	(2.4052661 −01)	(2.3335893 −01)
S_{yy}	−5.0000000 −01	11.000000 −01
	(2.7325886 −01)	(2.4088834 −01)
S_{xy}	−6.0000000 −01	17.000000 −01
	(4.2844143 −03)	(2.8016783 −03)
r_{xy}	−8.9442720 −01	1.4216115 +00
	(1.6711782 −02)	(1.1816753 −02)
	相関係数は全然ちがっている	相関係数は背理

例4. 有効5桁の計算

$$N(20,\ 0.5^2),\qquad \frac{1}{V_x}=\frac{1}{V_y}=40$$

	$n=50$		$n=800$	
S_{xx}	2.3000	−01	2.6000	−01
S_{yy}	1.3000	−01	2.1000	−01
S_{xy}	−1.0000	−01	−3.1000	−01
r_{xy}	−5.7830	−01	−1.3267	+00

計算の桁数をへらすと、同じ桁数の桁落ちであっても残留桁数が少なくなり、特にr_{xy}への影響が大きく、μ_x, σ_xの構造において例4は例1と同じであるにかかわらず、例4においては全く出たらめの相関係数が出てくるのを注意すべきである.

（本§の研究は鈴木利一君の成果にもとずく）

次に計算内容とのからみ合いが影響を持つ一事例として高次代数方程式の解きにくさの1原因を指摘する.

§2. 多項式の計算についての一考察.

近接根が多い代数方程式について、これらの近接根を一々求めることの困難さがよく言われている. はじめに近接根の一形式を提示し、それについての考察をする.

考える変域は複素数面の全体とする. ここで単連結の領域

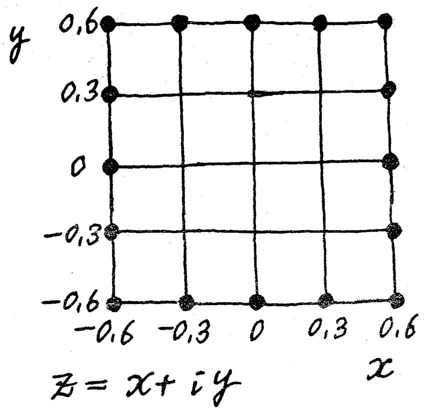
Dを考える．まず静電ポテンシャルのたとえを引くこととし，Dが完全導体でこれが真空中におかれたとする．D上に若干の電気量を与えたとすると，これはただちに釣合いの状態となり，電気はDの境界に一重層となって分布し，これにもとずく釣合いポテンシャルはD内で一定となる．連続一重層にもとずくポテンシャル（対数）が，これによってかこまれる形内で一定になっていることである．上の連続一重層を境界上の discrete な点上における一定点質量の集合でおきかえると，厳格ではないがD内でのポテンシャルがほぼ一定になることはおこらないであろうか．点質量をおく境界上の点を z_k，$(k=1,\cdots,N)$ とし質量 e をこれらに配布すれば，このポテンシャルは $\sum_{k=1}^{N} e \log(z-z_k)$ である．N を増減して考えるときに $\sum_{k=1}^{N} e$ を一定に保つように $e=\frac{1}{N}$ ととっておく．上のポテンシャルがほぼ一定となることは，或いは言いかえて $\left\{\prod_{k=1}^{N}(z-z_k)\right\}^{\frac{1}{N}}$ がほぼ一定となることである．ここに

$$f(z)=\prod_{k=1}^{N}(z-z_k)$$

は z_k に零点をおく N 次の多項式となる．

上のような形内での一定値を近似として得るためには N は相当に大きく，また z_k の配布もDの形に関連してしっかりきめられねばならないであろうが，さほど厳格ではないにし

てもほぼ類似の條件が作れないかと次のような例を作って見た、$z=x+iy$ とし、正方形 $|x|=0.6$, $|y|=0.6$ の上に等間隔（間隔 0.3）で零点を配置した16次多項式



$$(1)\qquad f(z)=\prod_{k=1}^{16}(z-z_k)$$

を作り、(1)式により z の種々の値における $f(z)$ を値を計算した、特に z 実軸上の値が表1になっているが、これによればほぼ見込の状況が作られている。

なおここでは正方形の一辺を1.2にえらんだが、この辺の大きさがこの程度のときは形内での $|f(z)|$ の値はほとんど 10^{-3} 以下に小さい。このような配列はこれらの零点の一とかたまりが一体となって根の捜索を困難にする次にのべるような事情を作る。

(1)の直接計算は乗算のみによるのでまず桁落ちの心配はない、これを展開した形に書いたものは

$$(2)\qquad z^{16}+0.5022\,z^{12}+0.01791153\,z^{8}+0.008324491824\,z^{4}-0.002754990144$$

となる。ここで今、全体の零点を配列をそのままにしておいて原点を少し移動して見る。例としては原点を−1だけ移動し多項式 $f(z-1)$ を作った。これを z で展開したものは

$$
\begin{aligned}
(3)\quad & z^{16} && -16z^{15}\\
& +120z^{14} && -560z^{13}\\
& +1820{,}5022z^{12} && -4374{,}0264z^{11}\\
& +8041{,}1452z^{10} && -11550{,}484z^{9}\\
& +13118{,}60691153z^{8} && -11837{,}88569224z^{7}\\
& +8472{,}53432284z^{6} && -4766{,}74544568z^{5}\\
& +2069{,}851131591824z^{4} &&\\
& && -671{,}52034364 7296z^{3}\\
& +153{,}696669790944z^{2} &&\\
& && -22{,}202990207296z\\
& +1{,}525681031680.
\end{aligned}
$$

ここでは念の為に1桁の丸めを行なわい計算を行なった。

(2) 又は (3) による多項式の計算はその中に加減算をふくむ。z を与えて普通の Horner 方式により値を出す場合、(2) の形式すなわち原点 $z=0$ が零点のかたまりの中に位置する場合はさほど桁落ちはおこらず、$f(z)$ の値はほとんど正鵠を得た値を与える。これに対し (3) による場合は万事の状況がただ z を1だけ移動しただけであるにもかかわらず表2で示すような不規則な正負交代を示し、たと

えば根の探索などを迷わす原因を与える．これは序にのべたような事情、正確な値 $a-b$ を求めるのに引く数、引かれる数の両方に大きな数 M が加わって真の $a-b$ をくらます状態を作るためであるように思われる．

表 1

z	$f(z)$ ‥ (1) による
0.00	−2.7549901 −03
0.05	−2.7549382 −03
0.10	−2.7541576 −03
0.15	−2.7507712 −03
0.20	−2.7416230 −03
0.25	−2.7221694 −03
0.30	−2.6861154 −03
0.35	−2.6242899 −03
0.40	−2.5212891 −03
0.45	−2.3460600 −03
0.50	−2.0268762 −03
0.55	−1.3883585 −03
0.60	0
0.65	+3.1735886 −03
0.70	+1.0550672 −02
0.75	+2.7602567 −02

展開した形 (2) で計算した $f(z)$，ただし計算は浮動 8 桁で行なったので、これに合わせて (2) の係数も終りの 2 項だけではあるが 8 桁に丸めたまの，これによる数値は本表とほとんど変るところがない．最後の数字が 1，2 ちがう程度

表2. (3)による $f(z-1)$

z	$f(z-1)$
1.50	−4.5065000 −03
1.51	−4.9652000 −03
1.52	−9.2332000 −03
1.53	−4.3970000 −03
1.54	+1.1874900 −02
1.55	−4.5906000 −03
1.56	−1.5387000 −03
1.57	−2.1244700 −03
1.58	−7.7959000 −03
1.59	+2.1430000 −04
1.60	+3.2522000 −03
1.61	−7.7290000 −04
1.62	+2.2703000 −03
1.63	+1.6625400 −02
1.64	−2.8419100 −02
1.65	−6.5899000 −03
1.66	−1.0194200 −02
1.67	+1.7295200 −02
1.68	−1.9146200 −02

ここでの $z=1.6$ が表1での $z=0.6$ に対応し、本来この附近ではここだけで多項式が 0 となるところである. 表1で 0.6 と対称な −0.6 で状況はすべて対称になっているはずである. これに対応するところはここでは $z=0.4$ であるが, ここでは 1.6 の附近ほどの不規則性はあらわれない.

$f(z-1)=0$ の根がはっきり $z=0.4$ とさがせるほどに値の計算ができ

ている．序にのべた $(M+a)-(M+b)=a-b$ となる，Mに相当するものがずっと小さくなるのであろう．

なお表2における(3)の計算も浮動8桁の立て前で行なったので，これに合わせて(3)の係数は8桁に丸めた．そのことのために特に正方形の右側の方で零点が大きく移動してしまったというおそれもなくはない．

上記は要するに根の cluster の内部で多項式がほとんど一定で且つ0に近い値をとるために，原点 $z=0$ がこの cluster からはなれている場合，多項式の計算が桁落ちのために支離滅裂となって，このために根の捜索がいちじるしく困難になるという状況を示す例である．