

代数方程式の数値解法

日本建設コンサルタント 平野 管 保

数値計算は有限桁で行なわれる。この点が非常に重要な点であり、その結果従来常識では考えられない様な解を得る事がある。

まずその第1は、与えられた代数方程式の係数の精度(桁数)に比較して、得られる解の精度(桁数)が悪い事がある点である。これは数値計算の計算方法の善し悪しには関係のない事で、誤差を含む代数方程式の係数が与えられた時、すでに解に入ってしまう誤差である。例えば、与えられた代数方程式の係数の精度(桁数)がすべて m 桁あり、その代数方程式が P 重根を含んでいれば、その P 重根の解は m/P 桁の精度しか得ることが出来ない。又近接根を含む代数方程式の場合も重根と同様な傾向がある。但し、代数方程式の係数の誤差間に何らかの関係がある場合には、解の精度が m/P 桁よりよくなる事があり、この事が桁数を多くして計算をする

理由の1つでもある。

第2は、1つの解が得られて、その解を除いた次数の1次下がった代数方程式の係数を求める時、計算方法が悪いと求められた解の誤差により1次下がった代数方程式の係数に誤差が入ってしまう。これは各解によつて係数に要求する精度(桁数)が係数毎に異なるからである。もし一度でも、次数を下げた代数方程式の係数に誤差が入ると、その代数方程式によつて得られる解は、初めに与えられた代数方程式の解とは異なつた解、即ち本来得られる筈の解の精度(桁数)が得られなくなつてしまう。しかしこれも又桁数を多くして計算を行なえば防ぐ事が出来る。

次に、解へ段々と収束させていく解法、ニュートン法、ベアストー法(ヒッチコック法)では、補正值を求めるために1次微分しか用いていない点であるが、これは補正值の絶対値が十分に小さいと仮定した場合の事であつて、実際の計算での補正值の絶対値はその様に充分小さいとは限らない。従つて2次微分以上の項を考慮して補正值を求めないと、いわゆるニュートン法に於ける飛ぶとゆう現象が起こる。この現象はせつかく解の近くまで収束していながら、補正值を加えると、かえつて解より離れに近似解になつてしまふ。時によつては異なつた値の他の解に近づいてしまふ事もある。

(1) 代数方程式の係数の誤差と解の誤差

係数に誤差を含まない時でも、有限固定桁で計算を行なうと、必ず数値に誤差が入ってしまう。これは計算に用いる桁数より下位の桁は零でなく任意の数値でも同じ数値になる。即ち誤差である。従つて与えられた代数方程式の係数には必ず誤差が含まれている。

$$f(z) = \sum_{i=0}^n (C_i \pm \Delta C_i) z^i = 0 \quad (1-1)$$

n 次の代数方程式を(1-1)式で表わし、近似解 \bar{z} を(1-1)式に代入し、 $f(\bar{z})$ が誤差のみになつたならば、近似解 \bar{z} をそれ以上真の解へ近づける事は出来なくなる。例えば

$$\max_{i=0,1,2,\dots,n} (|\Delta C_i \bar{z}^i|) \geq |f(\bar{z})| \quad (1-2)$$

(1-2)式の条件を満足したならば、更に計算を行なつても解の精度は良くなり、意味がない。しかし

$$(z - z_1)^n = 0 \quad (1-3)$$

(1-3)式の z_1 に誤差を入れて

$$(z - (z_1 + \Delta z_1))^n = 0 \quad (1-4)$$

(1-4)式で桁数を多くとつて計算し、(1-1)式の形にすると

$$C_n \pm \Delta C_n = 1.0$$

$$C_{n-1} \pm \Delta C_{n-1} = -n(z_1 + \Delta z_1)$$

i

$$\begin{aligned} C_1 \pm \Delta C_1 &= (-1)^{n-1} n (z_1 + \Delta z_1)^{n-1} \\ C_0 \pm \Delta C_0 &= (-1)^n (z_1 + \Delta z_1)^n \end{aligned} \quad (1-5)$$

(1-5) 式の係数の誤差間は独立ではなく、(1-2) 式よりも更に解の精度をあげる事が出来る。

次に、(1-1) 式の解 z_1 が P 重根であるとする。(1-1) 式を点 z_1 でテーラー展開してみると

$$f(z) = \sum_{i=p}^n \left\{ \frac{f^{(i)}(z_1)}{i!} \right\} (z - z_1)^i = 0 \quad (1-6)$$

[z_1 は P 重根であるから $f(z_1) = f'(z_1) = \dots = f^{(p-1)}(z_1) = 0$]
 $|z - z_1|$ が充分小さい値の時、(1-6) 式より、 $|f(z)|$ は $|z - z_1|$ に対して P 乗に比例している。一方、 $|z - z_1|$ が充分小さい時、 z と z_1 の数値を比較してみると、上位の桁の数値は等しく、下位の桁の数値のみ異なっているから、(1-1) 式の各項の絶対値は、 $|z - z_1|$ が $1/10$, $1/100$ と小さくなっても、ほとんど変化しない。従って (1-1) 式と (1-6) 式に z を代入した時は等しいので、 $|z - z_1|$ が 1 桁小さくなる様な z を (1-1) 式に代入すれば、即ち近似解の精度が 1 桁良くなれば、 $|f(z)|$ は P 桁小さくなり、有効桁数は P 桁失われる。この事から「与えられた代数方程式の係数がすべて m 桁の精度をもつならば、 P 重根の解の精度は m/P 桁ある。」といえる。

例として、2重根の解を含む3次代数方程式を示す。

$$f(z) = C_3 z^3 + C_2 z^2 + C_1 z + C_0$$

$$C_3 = 1.0000000$$

$$C_1 = 15.0000000$$

$$C_2 = -7.0000000$$

$$C_0 = -9.0000000$$

真の解は $z_1 = 1.0000000$

$z_2 = z_3 = 3.0000000$

$z = 3.0003000$ $C_0 = -9.0000000$ $C_1 z = 45.004500$ $C_2 z^2 = -63.012601$ $C_3 z^3 = 27.008101$ $f(z) = 0.0000000$	$z = 3.0030000$ $C_0 = -9.0000000$ $C_1 z = 45.045000$ $C_2 z^2 = -63.126063$ $C_3 z^3 = 27.081081$ $f(z) = 0.0000180$
$z = 3.0300000$ $C_0 = -9.0000000$ $C_1 z = 45.450000$ $C_2 z^2 = -64.266300$ $C_3 z^3 = 27.818127$ $f(z) = 0.0018270$	$z = 3.3000000$ $C_0 = -9.0000000$ $C_1 z = 49.500000$ $C_2 z^2 = -76.230000$ $C_3 z^3 = 35.937000$ $f(z) = 0.2070000$

関数値 $f(z)$ は 10^{-7} の桁が誤差—になっている。又、解 3.0 は 2重根であるから、 z の精度が 1 桁良くなると、関数値 $f(z)$ の有効桁数は 2 桁減少している事がわかる。従って 8 桁で計算をすると、解 3.0 は 4 桁目まで決定出来る。

次に、係数の誤差間が独立でない場合の例(1-5)式を考えてみる。

$$n = 3 \qquad Z_1 = 3.14159265$$

$$\Delta Z_1 = -0.00100000 \qquad Z_1 + \Delta Z_1 = 3.14059265$$

とすると、 $Z_1 + \Delta Z_1$ が Z_1 に対して3桁の精度であるから、3桁で(1-5)式の係数を求め、次に $Z_1 + \Delta Z_1$ を求める計算を行なう。 $Z_1 + \Delta Z_1$ は3重根であるから、 $Z_1 + \Delta Z_1$ の精度は1桁であり、真の解 Z_1 に対しては1桁の精度しか得られない。

$$C_3 \pm \Delta C_3 = 1.0 \qquad = 1.00$$

$$C_2 \pm \Delta C_2 = -3.0 (Z_1 + \Delta Z_1) = -9.42$$

$$C_1 \pm \Delta C_1 = 3.0 (Z_1 + \Delta Z_1)^2 = 29.6$$

$$C_0 \pm \Delta C_0 = -1.0 (Z_1 + \Delta Z_1)^3 = -31.0$$

Z	2.60	3.10
$(C_3 \pm \Delta C_3) Z^3$	17.6	29.8
$(C_2 \pm \Delta C_2) Z^2$	-63.7	-90.5
$(C_1 \pm \Delta C_1) Z$	77.0	91.8
$(C_0 \pm \Delta C_0)$	-31.0	-31.0
$f(Z)$	0.1	0.1

Z が 2.6 あるいは 3.1 でも $f(Z)$ は殆んど誤差のみになる。ところが(1-5)式の係数を求めるのに次の様に9桁で計算し、次に $Z_1 + \Delta Z_1$ を求めると、 $Z_1 + \Delta Z_1$ の精度は3桁であり

真の解 Z_1 に対しても3桁の精度が得られるのである。

$$C_3 \pm \Delta C_3 = 1.0 = 1.00000000$$

$$C_2 \pm \Delta C_2 = -3.0 (Z_1 + \Delta Z_1) = -9.42177795$$

$$C_1 \pm \Delta C_1 = 3.0 (Z_1 + \Delta Z_1)^2 = 29.5899666$$

$$C_0 \pm \Delta C_0 = -1.0 (Z_1 + \Delta Z_1)^3 = -30.9766772$$

Z	3.13500000	3.14400000
$(C_3 \pm \Delta C_3) Z^3$	30.8114854	31.0776100
$(C_2 \pm \Delta C_2) Z^2$	-92.5993536	-93.1317877
$(C_1 \pm \Delta C_1) Z$	92.7645453	93.0308550
$(C_0 \pm \Delta C_0)$	-30.9766772	-30.9766772
$f(Z)$	-0.0000001	0.0000001

計算して得られる解の誤差は、どうしたら得られるか。真の解 Z_{Tj} を (1-1) 式に代入してみる。

$$f(Z_{Tj}) + \varepsilon_j = \sum_{i=0}^n (C_i \pm \Delta C_i) Z_{Tj}^i + \varepsilon_j = 0 \quad (1-8)$$

$j = 1, 2, \dots, n$

ε_j なる誤差項を加えないと、等号は成り立たない。 ε_j として

$$\begin{aligned} \varepsilon_j &= \max_{i=0, 1, 2, \dots, n} \{ |(\pm \Delta C_i) Z_{Tj}^i| \} \\ &= |(\pm \Delta C_m) Z_{Tj}^m| \quad j = 1, 2, \dots, n \quad (1-9) \end{aligned}$$

$$\varepsilon_j = \sum_{i=0}^n |(\pm \Delta C_i) Z_{Tj}^i| \quad j = 1, 2, \dots, n \quad (1-10)$$

(1-10) 式で ε_j を求めると、これ以下で述べる解の誤差を求める方法では、解の誤差が大きめになるため、(1-9) 式を

採用する。又 (1-1) 式を解くと

$$f(Z_{Tj} \pm \Delta Z_{Tj}) = \sum_{i=0}^n (C_i \pm \Delta C_i) (Z_{Tj} \pm \Delta Z_{Tj})^i = 0$$

$$j = 1, 2, \dots, n \quad (1-11)$$

$Z_{Tj} \pm \Delta Z_{Tj}$ が得られる。次に

$$y = Z / Z_{Tj}, \quad \Delta y_{Tj} = \Delta Z_{Tj} / Z_{Tj}$$

$$j = 1, 2, \dots, n \quad (1-12)$$

と変数変換をして (1-8), (1-11) 式をそれぞれ

$$f(Z_{Tj} \cdot y) + \varepsilon_j = \sum_{i=0}^n (C_i \pm \Delta C_i) (Z_{Tj} \cdot y)^i + \varepsilon_j = 0$$

$$j = 1, 2, \dots, n \quad (1-13)$$

$$f(Z_{Tj} \cdot (y \pm \Delta y_{Tj})) = \sum_{i=0}^n (C_i \pm \Delta C_i) (Z_{Tj} \cdot (y \pm \Delta y_{Tj}))^i = 0$$

$$j = 1, 2, \dots, n \quad (1-14)$$

で表わし、 $y = 1.0$ であるから、(1-14) 式から (1-13) 式を減ずると

$$\sum_{i=1}^n \left(\frac{f^{(i)}(Z_{Tj})}{i!} \right) (\pm \Delta y_{Tj})^i = \varepsilon_j$$

$$j = 1, 2, \dots, n \quad (1-15)$$

(1-15) 式の左辺の各項は、どの項も誤差であるから

$$\left| \left(\frac{f^{(i)}(Z_{Tj})}{i!} \right) (\pm \Delta y_{Tj})^i \right| \leq \varepsilon_j$$

$$i = 1, 2, \dots, n; \quad j = 1, 2, \dots, n \quad (1-16)$$

であると考え、(1-16) 式を満足する $|\pm \Delta y_{Tj}|$ を求める

ために $|\pm \Delta y_{Tj}| = \min_{i=1, 2, \dots, n} \left(\left| \sqrt[i]{\frac{\varepsilon_j \cdot i!}{f^{(i)}(Z_{Tj})}} \right| \right)$

$$= \left| \sqrt{\frac{\varepsilon_j \cdot \bar{m}!}{f^{(\bar{m})}(Z_{Tj})}} \right| \quad j=1, 2, \dots, n \quad (1-17)$$

(1-17)式で求めた $|\pm \Delta y_{Tj}|$ を (1-16) 式の左辺に代入す

る。まず $i = \bar{m}$ の項は

$$\left| \frac{f^{(\bar{m})}(Z_{Tj})}{\bar{m}!} \left(\sqrt{\frac{\varepsilon_j \cdot \bar{m}!}{f^{(\bar{m})}(Z_{Tj})}} \right)^{\bar{m}} \right| = \varepsilon_j$$

$$j=1, 2, \dots, n \quad (1-18)$$

他の項は

$$\left| \frac{f^{(i)}(Z_{Tj})}{i!} \left(\sqrt{\frac{\varepsilon_j \cdot \bar{m}!}{f^{(\bar{m})}(Z_{Tj})}} \right)^i \right|$$

$$\leq \left| \frac{f^{(i)}(Z_{Tj})}{i!} \left(\sqrt{\frac{\varepsilon_j \cdot i!}{f^{(i)}(Z_{Tj})}} \right)^i \right| = \varepsilon_j$$

$$j=1, 2, \dots, n \quad (1-19)$$

即ち、(1-17)式で求めた $|\pm \Delta y_{Tj}|$ は必ず (1-16) 式を満足している。従つて (1-17) 式から解の相対誤差を求める事が出来る。解の絶対誤差は (1-12) 式を用い、 Z_{Tj} の代りに $Z_{Tj} \pm \Delta Z_{Tj}$ を代入し

$$|\pm \Delta Z_{Tj}| = |Z_{Tj} \pm \Delta Z_{Tj}| \min_{i=1, 2, \dots, n} \left(\left| \sqrt{\frac{\varepsilon_j \cdot i!}{f^{(i)}(Z_{Tj} \pm \Delta Z_{Tj})}} \right| \right)$$

$$j=1, 2, \dots, n \quad (1-20)$$

で求められる。又重根である解、あるいは近接根を含む解を求める時は (1-20) 式で $i=1$ が選ばれない。例えば、3重

根の解の場合は $i=3$ が選ばれる. 次に, 3重根の解を含む
3次代数方程式を例として, (1-20)式で計算した誤差をあ
げる.

$$f(z) = (\pi + ei) \{ z - (\pi + ei) \}^3 = C_3 z^3 + C_2 z^2 + C_1 z + C_0$$

上式の C_3, C_2, C_1, C_0 を計算し, 誤差を入れる.

$$C_3 \pm \Delta C_3 = 3.14159210 + 2.71828130 i$$

$$C_2 \pm \Delta C_2 = -7.44164200 - 51.2384030 i$$

$$C_1 \pm \Delta C_1 = -115.901801 + 181.198677 i$$

$$C_0 \pm \Delta C_0 = 285.555090 - 84.7328600 i$$

$$\pm \Delta C_3 = 5.5 \times 10^{-7} + 5.2 \times 10^{-7} i \quad |\pm \Delta C_3| = 7.5690 \times 10^{-7}$$

$$\pm \Delta C_2 = 2.9 \times 10^{-6} + 2.3 \times 10^{-6} i \quad |\pm \Delta C_2| = 3.7014 \times 10^{-6}$$

$$\pm \Delta C_1 = 8.1 \times 10^{-6} + 8.8 \times 10^{-6} i \quad |\pm \Delta C_1| = 1.1960 \times 10^{-5}$$

$$\pm \Delta C_0 = 3.2 \times 10^{-5} + 3.2 \times 10^{-5} i \quad |\pm \Delta C_0| = 4.5255 \times 10^{-5}$$

$$Z_1 = 3.15217390 + 2.74294324 i$$

$$Z_2 = 3.15749449 + 2.69689966 i$$

$$Z_3 = 3.11511032 + 2.71500419 i$$

$$|\Delta Z_1| = 0.0072310$$

$$|\Delta Z_2| = 0.0071916$$

$$|\Delta Z_3| = 0.0071117$$

真の誤差より (1-20)式で計算した誤差は小さめにでている
. これは ε_j を (1-9)式により計算しているからであり, ε_j

を(1-10)式より計算すると誤差は大きめにでる。

(2) 解への収束方法

与えられた代数方程式を

$$f(z) = C_0 + \sum_{i=1}^n C_i z^i = 0 \quad (2-1)$$

とする。繰り返し計算の j 回目の近似値を z_j とすると

$$f(z) = C_{0j} + \sum_{i=1}^n C_{ij} (z - z_j)^i = 0 \quad (2-2)$$

$$C_{0j} = f(z_j)$$

$$C_{ij} = f^{(i)}(z_j) / i!$$

で表わせる。もし(2-2)式で $C_{0j} = 0$ であるならば、 z_j は解である。次に $j+1$ 回目の近似値 $z_{j+1} = z_j + \Delta z_j$ を(2-1)式に代入し

$$|f(z_{j+1})| = |f(z_j + \Delta z_j)| < |f(z_j)| \quad (2-3)$$

になる様な Δz_j を求め、これを繰り返して $j \rightarrow \infty$ とすれば $f(z_j) \rightarrow 0$ となり解が求められる。

$\alpha \geq 1.0$ なる実数 α を用いて(2-2)式を

$$f(z) = \left(1 - \frac{1}{\alpha}\right) C_{0j} + \frac{1}{\alpha} C_{0j} + \sum_{i=1}^n C_{ij} (z - z_j)^i = 0 \quad (2-4)$$

と変形し、(2-4)式の右辺の第2項と第3項の合計を

$$f(z) = \frac{1}{\alpha} C_{0j} + \sum_{i=1}^n C_{ij} (z - z_j)^i \quad (2-5)$$

とする。次に

$$\left| \sqrt[m]{\frac{-C_{0j}}{\alpha C_{mj}}} \right| = \min_{i=1, 2, \dots, n} \left(\left| \sqrt[i]{\frac{-C_{0j}}{\alpha C_{ij}}} \right| \right) \quad (2-6)$$

で m を求める。(m 項として 2 項以上ある時は、いずれを選んでよい。) 極座標を用いて

$$-\frac{C_{0j}}{\alpha C_{mj}} = r_{\alpha m} e^{i\theta_{\alpha m}} \quad (2-7)$$

と表わし、(2-7) 式を用いて仮の Δz_j を

$$\Delta z_{jm'} = \sqrt[m]{r_{\alpha m}} e^{i(\theta_{\alpha m} + 2\pi m')/m} \quad m' = 1, 2, \dots, m \quad (2-8)$$

とする。(m' は m 個の中で任意の 1 つを選べばよい。)

ここで、(2-5) 式に $z = z_j + \Delta z_{jm'}$ を代入する。

$$\begin{aligned} f(z_j + \Delta z_{jm'}) &= \frac{1}{\alpha} C_{0j} + \sum_{i=1}^n C_{ij} (z_j + \Delta z_{jm'} - z_j)^i \\ &= \frac{1}{\alpha} C_{0j} + \sum_{i=1}^n C_{ij} (\Delta z_{jm'})^i \end{aligned} \quad (2-9)$$

(2-7) 式と (2-8) 式より

$$\begin{aligned} -C_{0j} / (\alpha \cdot C_{mj}) &= (\Delta z_{jm'})^m \\ \frac{1}{\alpha} C_{0j} + C_{mj} (\Delta z_{jm'})^m &= 0 \end{aligned} \quad (2-10)$$

従って (2-9) 式は

$$f(z_j + \Delta z_{jm'}) = \sum_{i=1}^{n'} C_{ij} (\Delta z_{jm'})^i \quad (2-11)$$

n' : $1, 2, \dots, n$ の中から m を除いたもの

となる。又(2-6)式で最小値を選んであるの。(2-9)

式の ΔZ_{jm} の*i*次の項に ΔZ_{jm} の値を代入すると

$$\begin{aligned} |C_{ij}(\Delta Z_{jm})^i| &= \left| C_{ij} \left(\sqrt[m]{\frac{-C_{0j}}{\alpha \cdot C_{mj}}} \right)^i \right| \\ &\leq \left| C_{ij} \left(\sqrt[i]{\frac{-C_{0j}}{\alpha \cdot C_{ij}}} \right)^i \right| = \left| \frac{-C_{0j}}{\alpha} \right| \quad (2-12) \end{aligned}$$

依って(2-9)式の中には、零次の項の絶対値より、絶対値の大きい項はない。もしここで

$$\left| \frac{1}{\alpha} C_{0j} \right| = |f(z_j)| > |f(z_j + \Delta Z_{jm})| \quad (2-13)$$

ならば ΔZ_{jm} を*j*回目の補正值 ΔZ_j として

$$z_{j+1} = z_j + \Delta Z_j \quad (\Delta Z_j = \Delta Z_{jm}) \quad (2-14)$$

を*j+1*回目の近似値とする。

$$\left| \frac{1}{\alpha} C_{0j} \right| = |f(z_j)| \leq |f(z_j + \Delta Z_{jm})| \quad (2-15)$$

ならば α を更に大きくして(2-4)式より繰り返す。(2-6)

)式で α を大きくすると

$$\left| \sqrt[i]{\frac{-C_{0j}}{\alpha \cdot C_{ij}}} \right| = \left| \sqrt[i]{\frac{1}{\alpha}} \right| \left| \sqrt[i]{\frac{-C_{0j}}{C_{ij}}} \right| \quad (2-16)$$

*i*の小さい項は $\left| \sqrt[i]{1/\alpha} \right|$ の小さくなるなり方が、*i*の大きい項よりも速いので、段々低次の項が選ばれ、極限では零でない一番*i*の小さい項が選ばれる。一方、 ΔZ_{jm} の絶対値も小さくなるので、(2-9)式で考えればわかる様に、(2-13)

)式を満足する Δz_{jm} が存在する. (2-13)式が満足されれば

$$\begin{aligned}
 |f(z_{j+1})| &= |f(z_j + \Delta z_j)| = \left| \left(1 - \frac{1}{\alpha}\right) C_{0j} + \bar{f}(z_j + \Delta z_j) \right| \\
 &< \left| \left(1 - \frac{1}{\alpha}\right) C_{0j} + \frac{1}{\alpha} C_{0j} \right| = |C_{0j}| \\
 &= |f(z_j)| \qquad (2-17)
 \end{aligned}$$

従つて (2-3)式が満足されるから $j \rightarrow \infty$ で $C_{0j} \rightarrow 0$ となり、解が1つ求められる。

次に、この収束方法を用いた例をあげる。

$$\begin{aligned}
 z^6 + 26.417062 z^5 + 162.32544 z^4 \\
 + 681.48974 z^3 + 477.60822 z^2 \\
 + 1088.9984 z + 74.312797 = 0
 \end{aligned}$$

解は $z_1 = -0.070187322$

$$z_2 = -0.13285772 - 1.3416340 i$$

$$z_3 = -0.13285772 + 1.3416340 i$$

$$z_4 = -3.0725744 - 4.4472419 i$$

$$z_5 = -3.0725744 + 4.4472419 i$$

$$z_6 = -19.936010$$

初期値を 0.0 とすると、一般のニュートン法と同様に先ず解 z_1 が求められる。解 z_1 を除いた5次の代数方程式の係数を求め、初期値を 0.0 とし、次の解 z_2 の収束計算を繰り返す。

繰り返し回数

1

$$-1.2568739 i$$

2	$-0.13701416 - 1.3309150 i$
3	$-0.13285435 - 1.3416929 i$
4	$-0.13285772 - 1.3416340 i$

4回で収束する。5次の代数方程式の係数は解 Z_1 が実数であるからすべて実数である。故に、前記の様に初期値 0.0 では一般のニュートン法を用いると複素解 Z_2 は求められない。

(3) 多項式の除算

与えられた代数方程式を

$$f(z) = \sum_{i=0}^n C_i z^i = 0 \quad (3-1)$$

とし、得られた解を Z_1 とすると、代数方程式の次数を1次下げる計算をしなければならない。

$$\bar{f}(z) = \left(\sum_{i=0}^n C_i z^i \right) / (z - Z_1) = \sum_{i=0}^{n-1} \bar{C}_i z^i = 0 \quad (3-2)$$

ところが、得られた解 Z_1 が与えられた代数方程式の係数に要求する精度(桁数)は、値の異なる他の解がその係数に要求する精度(桁数)とは異なる。今、2つの解を $|Z_1| < |Z_2|$ として、(1-2)式により、解 Z_1, Z_2 の精度に最も影響する項を求める。

$$|\pm \Delta C_i Z_1^i| = \max_{i=0,1,2,\dots,n} (|\pm \Delta C_i Z_1^i|)$$

$$|\pm \Delta C_{i_2} Z_2^{i_2}| = \max_{i=0,1,2,\dots,n} (|\pm \Delta C_i Z_2^i|) \quad (3-3)$$

(3-3)式中の $\pm \Delta C_i$ は(3-1)式中の係数 C_i の含む誤差である。(3-3)式より

$$\begin{aligned} |\pm \Delta C_{i_1} Z_1^{i_1}| &\geq |\pm \Delta C_{i_2} Z_1^{i_2}| \\ |\pm \Delta C_{i_2} Z_2^{i_2}| &\geq |\pm \Delta C_{i_1} Z_2^{i_1}| \end{aligned} \quad (3-4)$$

(3-4)式より

$$|Z_1^{i_2-i_1}| \leq |\pm \Delta C_{i_1} / \pm \Delta C_{i_2}| \leq |Z_2^{i_2-i_1}| \quad (3-5)$$

$$1.0 \leq |Z_2 / Z_1|^{i_2-i_1} \quad (3-6)$$

一方、仮定より $1.0 < |Z_2 / Z_1|$ であるから

$$i_2 - i_1 \geq 0 \quad \text{即ち} \quad i_2 \geq i_1 \quad (3-7)$$

従つて、絶対値の大きい解は、絶対値の小さい解よりも、与えられた代数方程式の次数の高い項の係数の誤差により、誤差が決定される。次に、 $\pi + ei$ の2重根及び $0.1(\pi + ei)$ の2重根を含む4次の代数方程式の例をあげる。

$$f(Z) = C_4 Z^4 + C_3 Z^3 + C_2 Z^2 + C_1 Z + C_0$$

まず、次数の低い項の係数に誤差が多く入っている場合は

$$C_4 = 3.1415921 + 2.7182813 i$$

$$C_3 = -5.4572040 - 37.574828 i$$

$$C_2 = -54.473842 + 85.163374 i$$

$$C_1 = 62.822090 - 18.641200 i$$

$$C_0 = -11.274100 - 5.1001000 i$$

$$Z_1 = 3.1425427 + 2.7207709 i$$

$$Z_2 = 3.1406429 + 2.7157922 i$$

$$Z_3 = 0.31530804 + 0.27323534 i$$

$$Z_4 = 0.31301054 + 0.27042248 i$$

(— の部分は誤差)

絶対値の小さい解 Z_3, Z_4 は解 Z_1, Z_2 に比較して精度が悪い。

これに反し、次数の高い項の係数に誤差が多く入っていると

$$C_4 = 3.1414000 + 2.7181000 i$$

$$C_3 = -5.4571800 - 37.574810 i$$

$$C_2 = -54.473842 + 85.163374 i$$

$$C_1 = 62.822122 - 18.641232 i$$

$$C_0 = -11.274256 - 5.1002296 i$$

$$Z_1 = 3.1686124 + 2.7432769 i$$

$$Z_2 = 3.1149880 + 2.6936860 i$$

$$Z_3 = 0.31436722 + 0.27214006 i$$

$$Z_4 = 0.31395197 + 0.27151739 i$$

絶対値の大きい解 Z_1, Z_2 は解 Z_3, Z_4 に比較して精度が悪い。

ここで実際に (3-2) 式を用いて高次の項より除算する。

$$\bar{C}_{n-1} = C_n$$

$$\bar{C}_{n-2} = C_n Z_1 + C_{n-1}$$

$$\vdots$$

$$\begin{aligned}\bar{C}_1 &= C_n Z_1^{n-2} + C_{n-1} Z_1^{n-3} + \dots + C_3 Z_1 + C_2 \\ \bar{C}_0 &= C_n Z_1^{n-1} + C_{n-1} Z_1^{n-2} + \dots + C_3 Z_1^2 + C_2 Z_1 + C_1 \\ \bar{C}_{-1} &= C_n Z_1^n + C_{n-1} Z_1^{n-1} + \dots + C_3 Z_1^3 + C_2 Z_1^2 + C_1 Z_1 + C_0 = 0.0\end{aligned}\quad (3-8)$$

(3-8)式の \bar{C}_{-1} は完全な零ではなく、収束判定(1-2)式より

$$|\pm \Delta \bar{C}_{-1}| = |\pm \Delta C_{i1} Z_1^{i1}| = \max_{i=0,1,2,\dots,n} (|\pm \Delta C_i Z_1^i|) \quad (3-8')$$

程度の誤差をもち、又、(3-8)式の係数の誤差は

$$|\pm \Delta \bar{C}_i| = \max_{j=i+1, i+2, \dots, n} (|\pm \Delta C_j Z_1^{j-(i+1)}|) \quad (3-8'')$$

$$i = 0, 1, 2, \dots, n-1$$

である。次に、(3-8)式の係数を用いて、(3-2)式により Z_2 を求めるのであるが、(3-1)式により求められる Z_2 と比較して誤差はどうであろうか。誤差を比較するために、(3-1)式に Z_2 を代入した時に得られる項と同じ項を含む次式を考えてみる。

$$\bar{f}(Z) \cdot Z_2 + \bar{C}_{-1} = 0 \quad (3-9)$$

(3-9)式に Z_2 を代入し、各項を(3-8)式で書き換え、

$$\bar{C}_{n-1} Z_2^n = C_n \cdot Z_2^n$$

$$\bar{C}_{n-2} Z_2^{n-1} = C_n Z_1 Z_2^{n-1} + C_{n-1} \cdot Z_2^{n-1}$$

⋮

$$\bar{C}_1 Z_2^2 = C_n Z_1^{n-2} Z_2^2 + C_{n-1} Z_1^{n-3} Z_2^2 + \dots + C_2 \cdot Z_2^2$$

$$\begin{aligned}\bar{C}_0 Z_2 &= C_n Z_1^{n-1} Z_2 + C_{n-1} Z_1^{n-2} Z_2 + \dots + C_2 Z_1 Z_2 + C_1 \cdot Z_2 \\ \bar{C}_{-1} &= C_n Z_1^n + C_{n-1} Z_1^{n-1} + \dots + C_2 Z_1^2 + C_1 Z_1 + C_0.\end{aligned}\quad (3-10)$$

(3-10)式の右辺の最後の項のみを集めると(3-1)式に Z_2 を代入した場合の各項と同じになる。又(3-9)式では定数としての Z_2 を(3-2)式に乘じ、零(誤差)である \bar{C}_{-1} を加えている。従つて、もし(3-1)式で Z_1 を求める時、無限桁の収束計算をすれば $\bar{C}_{-1}=0$ であり、(3-1)式で求められる Z_2 と、(3-9)式即ち(3-2)式で求められる Z_2 は等しくなる。ところが、(3-2)式で Z_2 を求める時の収束判定は

$$|\pm \Delta \bar{C}_{j_2} Z_2^{j_2+1}| = \max_{j=0,1,2,\dots,n} (|\pm \Delta \bar{C}_{j-1} Z_2^j|) \quad (3-11)$$

であるから、 \bar{C}_{-1} は零でなく誤差である。一方(3-1)式で Z_2 を求める時の収束判定は

$$|\pm \Delta C_{i_2} Z_2^{i_2}| = \max_{i=0,1,2,\dots,n} (|\pm \Delta C_i Z_2^i|) \quad (3-12)$$

である。ここで、(3-11)式と(3-12)式との間に

$$|\pm \Delta \bar{C}_{j_2} Z_2^{j_2+1}| > |\pm \Delta C_{i_2} Z_2^{i_2}| \quad (3-13)$$

の関係があると、(3-2)式で得られる Z_2 は、(3-1)式から得られる Z_2 よりも大きい誤差を含む事になる。又、(3-8)式で定めた $\pm \Delta C_{i_1}$ を含む項をもつ(3-8)式で、誤差の絶対値が最大の項は $\pm \Delta C_{i_1}$ を含む項である。依つて(3-8')式、(3-8'')式中 $\pm \Delta C_{i_1}$ を含む項をもつ $|\pm \Delta \bar{C}_i|$ は

$$|\pm \Delta \bar{C}_i| = \max_{j=i+1, i+2, \dots, n} (|\pm \Delta C_j Z_1^{j-(i+1)}|) = |\pm \Delta C_{i_1} Z_1^{i_1-(i+1)}|$$

$$i = -1, 0, 1, \dots, i_1-1 \quad (3-14)$$

になる。 Z_1 と Z_2 の関係を

$$Z_1 > Z_2 \quad (3-15)$$

とすると、(3-7)式より $i_1 \geq i_2$ であるから $\pm \Delta C_{i_2}$ を含む項をもつ(3-8)式は必ず $\pm \Delta C_{i_1}$ を含む項をもつ。従つて、 $i_1 > i_2$ の場合は(3-14)式で $\pm \Delta C_{i_2}$ を含む項は選ばれず、この $|\pm \Delta \bar{C}_i|$ を用いている(3-11)式では $\pm \Delta C_{i_2}$ を含む項は勿論選ばれなく、(3-13)式の関係が成り立つ。又、 $i_1 = i_2$ の場合は(3-14)式を(3-11)式に代入し、 $\pm \Delta C_{i_1} = \pm \Delta C_{i_2}$ を含む項の中で誤差の絶対値が最も大きい項を求めると

$$|\pm \Delta C_{i_1} Z_1^{i_1}| = \max_{i=-1, 0, 1, \dots, i_1-1} (|\pm \Delta C_{i_1} Z_1^{i_1-(i+1)} Z_2^{i+1}|)$$

$$(3-16)$$

であるから、やはり(3-13)式の関係が成り立つ。即ち、

1. 得られた解の絶対値より絶対値の小さい解を、高次の項から(3-2)式を除いて得られた $f(Z)$ から求めると、直接 $f(Z)$ から得られる解よりも、解の精度が悪くなる。

次に、 Z_1 と Z_2 の関係が

$$Z_1 \leq Z_2 \quad (3-17)$$

である時はどうなるであろうか。(3-10)式中で誤差の絶対

値が最大の項は

$$|\pm 4C_i Z_2^i| = \max_{j=0,1,2,\dots,i} (|\pm 4C_i Z_1^{i-j} Z_2^j|) \\ i = 0, 1, 2, \dots, n \quad (3-18)$$

の中にあるから (3-13) 式は成立せず。 (3-2) 式で求められる Z_2 が (3-1) 式から求められる Z_2 と同じ精度で得られる。次に例をあげる。(計算は 10 進法 8 桁で行なう。)

$$f(Z) = \sum_{i=0}^3 C_i Z^i = 0 \quad (3-19)$$

$$C_3 = 1.0000000 \quad C_1 = 9.9692877 \times 10^6$$

$$C_2 = -3.1733228 \times 10^4 \quad C_0 = -3.1006278 \times 10^7$$

真の解は $Z_2 = \pi$, $Z_1 = \pi \times 10^2$, $Z_3 = \pi \times 10^4$

まず $Z_1 = 3.1415927 \times 10^2$ を求め、(3-8) 式により

$$\bar{C}_2 = C_3 = 1.0000000$$

$$\bar{C}_1 = C_3 Z_1 + C_2 = -3.1419069 \times 10^4$$

$$\bar{C}_0 = (C_3 Z_1 + C_2) Z_1 + C_1 = 9.8695900 \times 10^4$$

次いで得られる解は

$$Z_2 = 3.1415880, \quad Z_3 = 3.1415927 \times 10^4$$

Z_2 の絶対値は Z_1 の絶対値より小さいから解の精度は悪くなり

Z_3 の絶対値は Z_1 の絶対値より大きいから解の精度は悪くならない。

先に (3-2) 式を高次の項より除算したが、低次の項より除算すると

$$\bar{C}_n = (C_0/Z_1^n + C_1/Z_1^{n-1} + \dots + C_{n-2}/Z_1^2 + C_{n-1}/Z_1 + C_n) = 0.0$$

$$\bar{C}_{n-1} = -(C_0/Z_1^n + C_1/Z_1^{n-1} + \dots + C_{n-2}/Z_1^2 + C_{n-1}/Z_1)$$

$$\bar{C}_{n-2} = -(C_0/Z_1^{n-1} + C_1/Z_1^{n-2} + \dots + C_{n-2}/Z_1)$$

$$\bar{C}_1 = -(C_0/Z_1^2 + C_1/Z_1)$$

$$\bar{C}_0 = -(C_0/Z_1)$$

(3-20)

(3-20) 式の \bar{C}_n は完全な零ではなく、(1-2) 式より

$$|\pm \Delta \bar{C}_n| = |\pm \Delta C_i Z_1^{i-n}| = \max_{i=0,1,2,\dots,n} (|\pm \Delta C_i Z_1^{i-n}|)$$

(3-20')

程度の誤差をもち、(3-20) 式の係数の誤差は

$$|\pm \Delta \bar{C}_i| = \max_{j=0,1,2,\dots,i} (|\pm \Delta C_j Z_1^{j-(i+1)}|)$$

$$i = 0, 1, 2, \dots, n-1$$

(3-20'')

である。次に、(3-9) 式と同じ理由で

$$F(Z) \cdot Z_1 + \bar{C}_n Z_2^n = 0$$

(3-21)

(3-21) 式を考え、(3-20) 式の係数を用いて書き換える。

$$\bar{C}_n Z_2^n = (C_0 Z_2^n / Z_1^n + C_1 Z_2^n / Z_1^{n-1} + \dots + C_{n-2} Z_2^n / Z_1^2 + C_{n-1} Z_2^n / Z_1 + C_n Z_2^n)$$

$$\bar{C}_{n-1} Z_2^{n-1} Z_1 = -(C_0 Z_2^{n-1} / Z_1^{n-1} + C_1 Z_2^{n-1} / Z_1^{n-2} + \dots + C_{n-2} Z_2^{n-1} / Z_1 + C_{n-1} Z_2^{n-1})$$

$$\bar{C}_{n-2} Z_2^{n-2} Z_1^2 = -(C_0 Z_2^{n-2} / Z_1^{n-2} + C_1 Z_2^{n-2} / Z_1^{n-3} + \dots + C_{n-2} Z_2^{n-2})$$

$$\bar{c}_1 z_2 z_1 = -(C_0 z_2 / z_1 + C_1 z_2)$$

$$\bar{c}_0 z_1 = -(C_0)$$

(3-22)

(3-22)式の右辺の最後の項のみを集めると(3-1)式に z_2 を代入した場合の各項と同じになる。(3-22)式の係数を用いた場合、(3-2)式で z_2 を求める時の収束判定は

$$|\varepsilon| = \max \left\{ \max_{j=0,1,2,\dots,n-1} (|\pm \Delta \bar{c}_j z_2^j z_1|), |\pm \Delta \bar{c}_n z_2^n| \right\}$$

(3-23)

である。(3-23)式と(3-12)式との間に

$$|\varepsilon| > |\pm \Delta C_{i_2} z_2^{i_2}| \quad (3-24)$$

の関係があれば、(3-2)式で得られる z_2 は、(3-1)式から得られる z_2 よりも大きい誤差を含む。一方、(3-20)式で定めた $\pm \Delta C_{i_1}$ を含む項をもつ(3-20)式で、誤差の絶対値が最大の項は $\pm \Delta C_{i_1}$ を含む項である。依って(3-20)式の中で $\pm \Delta C_{i_1}$ を含む項をもつ $|\pm \Delta \bar{c}_i|$ は

$$|\pm \Delta \bar{c}_i| = \max_{j=0,1,2,\dots,i} (|\pm \Delta C_j z_1^{j-(i+1)}|) = |\pm \Delta C_{i_1} z_1^{i_1-(i+1)}|$$

$$i = i_1, i_1+1, \dots, n-1 \quad (3-25)$$

になる。 z_1 と z_2 の関係を

$$z_1 < z_2 \quad (3-26)$$

とすると、(3-7)式より $i_1 \leq i_2$ であるから $\pm \Delta C_{i_2}$ を含む

項をもつ (3-20) 式は必ず $\pm \Delta C_{i_1}$ を含む項をもつ。従つて、
 $i_1 < i_2$ の場合は (3-25) 式で $\pm \Delta C_{i_2}$ を含む項は選ばれず、
 この $|\pm \Delta C_{i_1}|$ を用いている (3-23) 式では $\pm \Delta C_{i_2}$ を含む項は勿
 論選ばれなく、(3-24) 式が成り立つ。又、 $i_1 = i_2$ の場合
 は (3-25) 式を (3-23) 式に代入し、 $\pm \Delta C_{i_1} = \pm \Delta C_{i_2}$ を含む項
 の中で誤差の絶対値が最も大きい項を求めると

$$|\pm \Delta C_{i_1} z_2^n / z_1^{n-i_1}| = \max_{i=i_1, i_1+1, \dots, n} (|\pm \Delta C_{i_1} z_2^i / z_1^{i-i_1}|) \quad (3-27)$$

であるから、やはり (3-24) 式の関係が成り立つ。即ち

2. 得られた解の絶対値より絶対値の大きい解を、低次の
 項から (3-2) 式を除して得られた子 (Z) から求めると、
 直接 $f(z)$ から得られる解よりも、解の精度が悪くなる。

次に、 z_1 と z_2 の関係が

$$z_1 \geq z_2 \quad (3-28)$$

である時はどうなるであろうか。(3-22) 式中で誤差の絶対
 値が最大の項は

$$|\pm \Delta C_i z_2^i| = \max_{j=i, i+1, \dots, n} (|\pm \Delta C_i z_2^j / z_1^{j-i}|) \quad (3-29)$$

$$i = 0, 1, 2, \dots, n$$

の中にあるから (3-24) 式は成立せず、(3-2) 式で求めら
 れる z_2 が (3-1) 式から求められる z_2 と同じ精度で得られる。
 次に、(3-19) 式を用いた例をあげる。

$$\bar{c}_2 = -[(C_0/Z_1 + C_1)/Z_1 + C_2]/Z_1 = 9.9999914 \times 10^4$$

$$\bar{c}_1 = -(C_0/Z_1 + C_1)/Z_1 = -3.1419069 \times 10^4$$

$$\bar{c}_0 = -(C_0/Z_1) = 9.8696047 \times 10^4$$

$$Z_2 = 3.1415927, \quad Z_3 = 3.1415954 \times 10^4$$

Z_2 の絶対値は Z_1 の絶対値より小さいから解の精度は悪くならない。 Z_3 の絶対値は Z_1 の絶対値より大きいから解の精度は悪くなる。

しからば、(3-2)式の係数は、どうして計算したら Z_1 の誤差を入れずに求められるか。これには2通りの計算方法が考えられる。その1つは非常に多くの桁をとり、(3-1)式で Z_1 を求める時に収束の桁数を多くする方法であり、他の1つは高次の項、低次の項の両方から $(Z-Z_1)$ で除算し、(3-3)式により得られるこの項をその境とする方法である。では、第1の方法、桁を多くとって、収束桁数を多くする方法について述べる。まず、桁数をどれだけとればよいか。(3-1)式を用いて Z_1 を求める時、高次の項より除算して(3-2)式の係数を求めるとすれば、 Z_1 の絶対値より絶対値の小さい解についてのみ考えればよい。今、(3-1)式で Z_2 を求めると、収束判定は(3-12)式であるから、(3-10)式の各項の計算を行なう時、その各項に入る4捨5入による誤差が(3-12)式より小さくなれば、 Z_2 の誤差は、(3-2)式で Z_2

を求めても、(3-1)式で Z_2 を求めた時より大きくはならない。又、 Z_1 を求める時、他の解は未知であるから(3-12)式の値はわからない。そこで、(3-12)式で $|\pm 4C_0|$ が選ばれたとすると、(3-10)式のすべての項に入る4捨5入の誤差は $|\pm 4C_0|$ より小さくなる様に桁をとらなければならない事になる。実際に、 Z_1 を求めた後で Z_2 が求められ、

$$|\pm 4C_{i2}Z_2^{i2}| > |\pm 4C_0| \quad (3-30)$$

であったとしても、 $|Z_2| < |Z_1|$ であるから

$$|\pm 4C_i Z_1^i| = \max_{j=0,1,2,\dots,i} (|\pm 4C_i Z_1^j Z_2^{i-j}|) \quad (3-31)$$

$$i = 0, 1, 2, \dots, n$$

が成立し、 $C_i Z_1^i$ ($i=0, 1, \dots, n$)に入る4捨5入の誤差が $|\pm 4C_0|$ より小さい様に、 C_i, Z_1^i ($i=0, 1, \dots, n$)の桁をとっているから、当然 $|\pm 4C_{i2}Z_2^{i2}|$ より小さい4捨5入の誤差しか(3-10)式のすべての項には入らない。

次に、桁数を多くとつて計算した例を示す。

①係数の精度は4桁 ③ — の部分は誤差

②計算桁数は10桁 ④ $\Delta\Delta\Delta$ の部分は4捨5入による誤差

$$\sum_{i=0}^3 C_i Z^i = 0 \quad \text{真の解は } Z_1 = \pi \times 10^4, Z_2 = \pi \times 10^2, Z_3 = \pi$$

$$C_3 = 1.000000000 \quad C_1 = 9.969000000 \times 10^6$$

$$C_2 = -3.173000000 \times 10^4 \quad C_0 = -3.100000000 \times 10^7$$

まず、 Z_1 を10桁計算で求める。 $Z_1 = 3.141267545 \times 10^4$

$$\bar{C}_2 = C_3$$

$$\bar{C}_1 = C_3 Z_1 + C_2 = \bar{C}_2 Z_1 + C_2$$

$$\bar{C}_0 = (C_3 Z_1 + C_2) Z_1 + C_1 = \bar{C}_1 Z_1 + C_1$$

$$\bar{C}_{-1} = [(C_3 Z_1 + C_2) Z_1 + C_1] Z_1 + C_0 = \bar{C}_0 Z_1 + C_0$$

$$\bar{C}_2 Z_1 = 3.141267545 \times 10^4$$

$$C_2 = -3.173000000 \times 10^4$$

$$\bar{C}_1 = 3.173245500 \times 10^2$$

$$\bar{C}_1 Z_1 = -9.968013101 \times 10^6$$

$$C_1 = 9.969000000 \times 10^6$$

$$\bar{C}_0 = 9.868990000 \times 10^2$$

$$\bar{C}_0 Z_1 = 3.100113799 \times 10^7$$

$$C_0 = -3.100000000 \times 10^7$$

$$\bar{C}_{-1} = 1.137990000 \times 10^3$$

この様に $\bar{C}_0 Z_1$ に C_0 を加算する時、 C_0 の — の桁に $\bar{C}_0 Z_1$ の $\Delta\Delta$ の桁が入るように桁数をとつて収束計算を行なえば、解 $Z_2 = \pi \times 10^2$ 、 $Z_3 = \pi$ は $Z_1 = \pi \times 10^4$ の解を除いた 2 次方程式で解いても、初めに与えられた 3 次代数方程式から直接解かれた解 Z_2 、 Z_3 と精度は変わらない。

ベアストー法（ヒッチコック法）で収束計算を行なう時にも、同様に桁数を増加させないと、次数の下げられた代数方程式より得られる解は、初めに与えられた代数方程式から直

接求めた解よりも誤差は大きくなる。次に、4次代数方程式の例を示す。

① 係数の精度は4桁 ③ — の部分は誤差

② 計算桁数は9桁 ④ $\Delta\Delta$ の部分は4捨5入による誤差

$$\sum_{i=0}^4 C_i Z^i = 0$$

真の解は $Z_1 = \pi \times 10^3$, $Z_2 = \pi \times 10^2$, $Z_3 = \pi \times 10$, $Z_4 = \pi$

$$C_4 = 3.14100000$$

$$C_2 = 3.47600000 \times 10^6$$

$$C_3 = -1.09700000 \times 10^4$$

$$C_1 = -1.08200000 \times 10^8$$

$$C_0 = 3.06000000 \times 10^8$$

まず、 Z_1 と Z_2 を9桁計算で求め、 $Z^2 - (Z_1 + Z_2)Z + Z_1 Z_2$ で除

す。 $(Z_1 + Z_2) = 3.45796527 \times 10^3$, $Z_1 Z_2 = 9.87072031 \times 10^5$

$$\bar{C}_2 = C_4$$

$$\bar{C}_2 = C_3$$

$$\bar{C}_1 = \bar{C}_2 (Z_1 + Z_2) + \bar{C}_2$$

$$\bar{C}_1 = C_2 - \bar{C}_2 Z_1 Z_2$$

$$\bar{C}_0 = \bar{C}_1 (Z_1 + Z_2) + \bar{C}_1$$

$$\bar{C}_0 = C_1 - \bar{C}_1 Z_1 Z_2$$

$$\bar{C}_1 = \bar{C}_0 (Z_1 + Z_2) + \bar{C}_0 = 0.0$$

$$\bar{C}_1 = C_0 - \bar{C}_0 Z_1 Z_2 = 0.0$$

$$\bar{C}_2 (Z_1 + Z_2) = 1.08614689 \times 10^4$$

$$\bar{C}_2 = -1.09700000 \times 10^4$$

$$\bar{C}_1 = 1.08531100 \times 10^2$$

$$C_2 = 3.47600000 \times 10^6$$

$$-\bar{C}_2 Z_1 Z_2 = -3.10039325 \times 10^6$$

$$\bar{C}_1 = 3.75606750 \times 10^5$$

$$\bar{c}_1(z_1 + z_2) = -3.75296775 \times 10^5$$

$$\bar{c}_1 = 3.75606750 \times 10^5$$

$$\bar{c}_0 = 3.09975000 \times 10^2$$

$$C_1 = -1.08200000 \times 10^8$$

$$-\bar{c}_1 z_1 z_2 = 1.07128013 \times 10^8$$

$$\bar{c}_0 = -1.07198700 \times 10^6$$

$$\bar{c}_0(z_1 + z_2) = 1.07188278 \times 10^6$$

$$\bar{c}_0 = -1.07198700 \times 10^6$$

$$\bar{c}_{-1} = -1.04220000 \times 10^2$$

$$C_0 = 3.06000000 \times 10^8$$

$$-\bar{c}_0 z_1 z_2 = -3.05967653 \times 10^8$$

$$\bar{c}_{-1} = 3.23470000 \times 10^4$$

C_0 に $-\bar{c}_0 z_1 z_2$ を加算する時、 C_0 の—の桁に $-\bar{c}_0 z_1 z_2$ の $\Delta\Delta$ の桁が入る様に桁数をとつて収束計算をすれば、次数の下げられた2次方程式から得られる解 z_3, z_4 は、初めに与えられた4次方程式から直接求めた解 z_3, z_4 と同じ精度が得られる。

最後に、桁数を増加させずに、且つ得られた解の誤差を入れずに、次数を1つ下げた代数方程式の係数の求め方を述べる。これが、数値解法が有限桁である事をはつきり示している点であり、単なる多項式の除算であつても計算する順序が異なると、同じ式で計算しても解が得られない事がある。

得られた解 Z_1 を用いて

$$|\pm \Delta C_{imax} Z_1^{imax}| = \max_{i=0,1,2,\dots,n} (|\pm \Delta C_i Z_1^i|) \quad (3-32)$$

なる $imax$ を求め、この $imax$ を用いて (3-1) 式を

$$f(Z) = \left(\sum_{i=0}^{imax-1} C_i Z^i \right) + C_{imax} Z^{imax} + \left(\sum_{j=imax+1}^n C_j Z^j \right) \quad (3-33)$$

の様に分離し、(3-9) 式の第1項は $(Z-Z_1)$ で低次の項より除し、第3項は $(Z-Z_1)$ で高次の項より除す。(3-2) 式

中の \bar{C}_i は

中の \bar{C}_i は

$$\bar{C}_{n-1} = C_n$$

$$\bar{C}_{n-2} = C_n Z_1 + C_{n-1}$$

⋮

$$\bar{C}_{imax} = C_n Z_1^{n-(imax+1)} + C_{n-1} Z_1^{n-(imax+2)} + \dots + C_{imax+1}$$

$$\bar{C}_{imax-1} = -(C_0 / Z_1^{imax} + C_1 / Z_1^{imax-1} + \dots + C_{imax-1} / Z_1)$$

⋮

$$\bar{C}_1 = -(C_0 / Z_1^2 + C_1 / Z_1)$$

$$\bar{C}_0 = -C_0 / Z_1$$

$$\bar{C}_{-1} = C_n Z_1^{n-imax} + C_{n-1} Z_1^{n-(imax+1)} + \dots + C_{imax+1} Z_1$$

$$+ C_{imax}$$

$$+ C_0 / Z_1^{imax} + C_1 / Z_1^{imax-1} + \dots + C_{imax-1} / Z_1$$

$$= f(Z_1) / Z_1^{imax} = 0.0$$

(3-34)

\bar{c}_{-1} は低次の項、高次の項の両側より除して、中央に残る項である。ところで (3-2) 式で次の解 Z_2 を求めると

(3-2) 式に定数を乗じ、零である項を加えた式で Z_2 を求めても、(3-2) 式から直接得られる解 Z_2 と等しい解が得られる筈であるから、(3-2) 式の代わりに次の式で解を求める。

$$|Z_2| \geq |Z_1| \quad \bar{f}(Z) \times Z_2 + \bar{c}_{-1} Z_2^{i_{max}} = 0 \quad (3-35)$$

$$|Z_2| < |Z_1| \quad \bar{f}(Z) \times Z_1 + \bar{c}_{-1} Z_2^{i_{max}} = 0 \quad (3-36)$$

(3-35) 式の第 1 項での Z_2 及び (3-36) 式の第 1 項での Z_1 は定数と考え、 $\bar{c}_{-1} = 0$ であるから (3-35), (3-36) 式の第 2 項は零である。ここでは (3-36) 式の説明のみ行なうが、(3-35) 式についても同様に説明が出来る。まず (3-36) 式を Z_1, Z_2 及び (3-1) 式の係数 C_i を用いて書き直してみる。

$$\bar{c}_{n-1} Z_2^{n-1} Z_1 = C_n Z_2^{n-1} Z_1$$

$$\bar{c}_{n-2} Z_2^{n-2} Z_1 = C_n Z_2^{n-2} Z_1^2 + C_{n-1} Z_2^{n-2} Z_1$$

⋮

$$\bar{c}_{i_{max}} Z_2^{i_{max}} Z_1 = C_n Z_2^{i_{max}} Z_1^{n-i_{max}} + C_{n-1} Z_2^{i_{max}} Z_1^{n-(i_{max}+1)} + \dots$$

$$\dots + C_{i_{max}+1} Z_2^{i_{max}} Z_1$$

$$\bar{c}_{-1} Z_2^{i_{max}} = C_n Z_2^{i_{max}} Z_1^{n-i_{max}} + C_{n-1} Z_2^{i_{max}} Z_1^{n-(i_{max}+1)} + \dots$$

$$\dots + C_{i_{max}+1} Z_2^{i_{max}} Z_1 + C_{i_{max}} Z_2^{i_{max}}$$

$$+ C_0 Z_2^{i_{max}} / Z_1^{i_{max}} + C_1 Z_2^{i_{max}} / Z_1^{i_{max}-1} + \dots$$

$$\dots + C_{i_{max}-1} Z_2^{i_{max}} / Z_1$$

$$\bar{C}_{i_{\max}-1} Z_2^{i_{\max}-1} Z_1 = - (C_0 Z_2^{i_{\max}-1} / Z_1^{i_{\max}-1} + C_1 Z_2^{i_{\max}-1} / Z_1^{i_{\max}-2} + \dots + C_{i_{\max}-1} Z_2^{i_{\max}-1})$$

$$\bar{C}_1 Z_2 Z_1 = - (C_0 Z_2 / Z_1 + C_1 Z_2)$$

$$\bar{C}_0 Z_1 = - C_0 \quad (3-37)$$

(3-37)の各式の右辺の合計を零とする様に Z_2 を求めると、(3-36)式即ち(3-2)式で解 Z_2 を求める事と同じになる。従つて(3-37)の各式の右辺で誤差の最も大きい項を求め、その誤差が(3-1)式で解 Z_2 を求める時の最も誤差の大きい項の誤差と等しい数值であるならば、(3-1)式で求める解 Z_2 と、(3-2)式で求める解 Z_2 の精度が同じわけである。(3-13)式の中で $\pm \Delta C_i$ ($i = n, n-1, \dots, i_{\max}$)による最大の誤差はどうであろうか。それは $|Z_2| < |Z_1|$ であり、且つ(3-32)式より $|\pm \Delta C_{i_{\max}} Z_2^{i_{\max}}|$ であることが知れる。又 $\pm \Delta C_i$ ($i = 0, 1, 2, \dots, i_{\max}-1$)による最大の誤差は

$$\max_{i=0, 1, 2, \dots, i_{\max}-1} (|\pm \Delta C_i Z_2^i|)$$

であるから、(3-37)の各式の中で最も誤差の大きい項は

$$\max_{i=0, 1, 2, \dots, i_{\max}} (|\pm \Delta C_i Z_2^i|)$$

一方、(3-1)式に Z_2 を代入して、最も誤差の大きい項は

$$\max_{i=0, 1, 2, \dots, n} (|\pm \Delta C_i Z_2^i|)$$

であり、 $|Z_2| < |Z_1|$ であるから(3-7)式より両者は等しい

。次に(3.19)式を用いた例をあげる。

$$\bar{C}_2 = C_3 = 1.0000000$$

$$\bar{C}_1 = C_3 Z_1 + C_2 = 3.1419069 \times 10^4$$

$$\bar{C}_0 = -C_0 / Z_1 = 9.8696047 \times 10^4$$

$$Z_2 = 3.1415927, \quad Z_3 = 3.1415927 \times 10^4$$

Z_2, Z_3 共に解の精度は悪くならない。

