OBORO の収束判定条件の設定

杉山耕一朗 (北大理・地惑)
sugiyama@gfd-dennou.org

2004/07/13

目 次

1 誤差 **2**

1.1 数値計算における誤差 2

1.2 許容誤差 2

1.3 ニュートン法の収束判定 3

2 OBORO の収束判定条件 **3**

2.1 許容相対誤差の設定 4

2.2 許容絶対誤差の設定 4

2.3 許容絶対モル数の設定 4

3 参考文献 **5**

1　誤差

1.1　数値計算における誤差

数値誤差を生じることのない計算を理論計算と呼ぶことにしよう. 理論計算とは, いわば無限桁計算である. 数値計算においては浮動小数点 10 進 6 桁 (単精度), もしくは 10 進 15 桁 (倍精度) 程度の有効桁数値を用いて計算するので, 誤差は避けがたい. 誤差が小さいうちは実用上困難が生じることは無いが, 大きくなると数値計算の結果が信頼できないものとなる.

ある数の真値を a, 近似解を x とするとき,

$$e = x - a \tag{1}$$

を誤差 (error) という. 誤差の絶対値

$$|e| = |x - a| \tag{2}$$

を絶対誤差 (absolute error) という.

誤差 e が真の値 a と何桁まであっているかを示すために, 誤差と真の値との比

$$e_R = \frac{e}{a} = \frac{x - a}{a} \tag{3}$$

を相対誤差 (relative error) という. e_R が 10^{-5} であるならば, x は 10 進 5 桁の精度を持つことを意味する. ただし, 真の値が 0 あるいは 0 に近い時には誤差の程度を相対誤差で言い表すことができず, 絶対誤差で言い表さねばならない.

1.2　許容誤差

絶対誤差がある小さな値 ε_A よりも小さいとき, 近似値 x を真の値 a の代わりに用いてもよいと判断できる. この ε_A を許容絶対誤差 (tollerable absolute error) と言う. また相対誤差の絶対値がある小さな値 ε_R よりも小さいとき, 近似値 x を真の値 a の代わりに用いてもよいと判断できる. この ε_R を許容相対誤差 (tollerable relative error) と言う. これらの許容誤差は通常はじめからわかっている値である. 以上をまとめると以下のようになる.

$$|x - a| < \varepsilon_A, \qquad \text{許容絶対誤差} \tag{4}$$

$$|x - a| < \varepsilon_R |a|. \qquad \text{許容相対誤差} \tag{5}$$

しかし, $|a|$ が 0 もしくは 0 に近い時に遠いときにも利用できる誤差の基準としては,

$$|x - a| < \varepsilon_A + \varepsilon_R |a| \tag{6}$$

が実用的である. 一般に a はわかっていないから, $|a|$ の代わりに $|x|$ を使えば,

$$|x - a| < \varepsilon_A + \varepsilon_R |x| \tag{7}$$

となり, 右辺が計算できる.

許容誤差は, 小さければ小さいほど高い精度を要求する条件になっている. 最も高い精度で計算する場合,

許容絶対誤差
: 計算機が記憶できる最小の絶対値. 7 ビット指数部の計算機では $10^{-65} = 5.31890^{-79}$ 程度である.

許容相対誤差 計算機の 16 進桁数の逆数 (マシン・イプシロン) で, 仮数部が 16 進 6 桁 (単精度) ならば, $16^{-(6-1)} = 9.537\times10^{-7}$, 仮数部が 16 進 14 桁 (倍精度) ならば, $16^{-(14-1)} = 2.220\times10^{-16}$ である.

である. これらの許容誤差よりも小さな許容誤差を選ぶことには意味がない.

1.3　ニュートン法の収束判定

ニュートン法は適当な初期値 x^0 から出発し, 反復的に近似解 x^k を求めるものである. ニュートン法を利用する際には, 適当な収束条件を設置し, 近似解が真の値に近くなった場合には反復計算を止めるという操作が必要となる.

真の値が 0 に近くても遠くても利用できる最も汎用的は収束判定条件は,

$$|x^{k+1} - x^k| < \varepsilon_A + \varepsilon_R(|x^k| + |x^{k+1}|) \tag{8}$$

である.

2　OBORO の収束判定条件

oboro モデルでは上昇する空気塊の擬断熱を仮定しているため, 上昇するにつれて空気塊に含まれる凝結性成分の存在量が 0 に近くなる. 言い替えると真の値が 0 に近くなる温度・圧力・元素量でも計算する必要があるため, 0 に近い場合でも適用できる汎用的な収束判定条件を設定する必要がある.

$$|x^{k+1} - x^k| < \varepsilon_A + \varepsilon_R(|x^k| + |x^{k+1}|) \tag{9}$$

この収束判定条件を利用するためには, $\varepsilon_A, \varepsilon_R$ の 2 つを設定する必要がある.

具体例を考える. 数値解が十分に 0 より大きい場合には, 相対誤差

$$\frac{n^{k+1} - n^k}{n^k} < \varepsilon_R \tag{10}$$

を収束条件とするのは問題ない. しかし問題となるのは十分に 0 に近いような数値解の場合で, その場合には例えば

$$\begin{aligned} \frac{n^{k+1} - n^k}{n^k} &= \frac{1.11111111111111\times10^{-12} - 1.11111111111110\times10^{-12}}{1.0\times10^{-12}} \\ &= \frac{1.0\times10^{-27}}{1.0\times10^{-12}} < \varepsilon_R \end{aligned} \tag{11}$$

というように, 誤差が限りなく 0 に近くなるまで計算せねばならなくなる. ところが, 上述の例のように 10^{27} 桁まで計算することには意味がない. なぜならば重要なのは系のエントロピーであり, 系のエントロピーに影響を与えないような誤差は計算する必要がない. したがって, 許容絶対誤差を設定し, 系全体の熱力学状態に影響を与えないような誤差を無視するように設定する.

2.1 許容相対誤差の設定

許容相対誤差は恣意的に決める必要がある. 現在の計算は倍精度なので, 許容相対誤差としてはマシン・イプシロン程度 (10^{13}) の値を与えることにする.

2.2 許容絶対誤差の設定

oboro モデルで計算するのはモル数であり, 計算する空気塊において重要な物理量はエントロピーである. エントロピーの値に寄与しないほど化学種のモル数の誤差が小さいならば, その化学種のモル数を精度よく計算する必要はない. そこで許容絶対誤差として, 系のエントロピーに影響を与える程度の量と設定した.

許容絶対誤差は任意に与えることができず, モデルの内部で計算せねばならない.

$$\begin{aligned}
&\frac{s_i \Delta n_i}{\sum_i s_i n_i} < \varepsilon_R \\
&\frac{s_i \Delta n_i}{\bar{s} \sum_i n_i} < \varepsilon_R \\
&\frac{\Delta n_i}{\sum_i n_i} < \varepsilon_R \times 0.1 \\
&\varepsilon_A = \varepsilon_R \times \sum_i n_i \times 0.1
\end{aligned} \tag{12}$$

ただし, 添字 i は化学種を意味し, $\bar{s}$ は比エントロピーの平均である. 化学種間のエントロピーの値はおよそオーダー 3 程度にしか違わないので, ここでは $s_i \approx \bar{s}$ とし, 念のために $1/10$ 倍してある.

2.3 許容絶対モル数の設定

oboro モデルで計算するのはモル数であり, 計算する空気塊において重要な物理量はエントロピーである. エントロピーの値に寄与しないほど化学種のモル数が小さいならば, その化学種のモル数を計算する必要はない. 微少量の化学種を系から除外して計算を続けた方が効率が良い.

そこで許容絶対モル数を設定した. これはある化学種のエントロピーの寄与が全エントロピーの許容相対誤差程度か否かによって判断する.

$$\begin{aligned}
&\frac{s_i n_i}{\sum_i s_i n_i} < \varepsilon_R, \\
&N_i \equiv \frac{\varepsilon_R \times \sum_i s_i n_i}{s_i}.
\end{aligned} \tag{13}$$

ここで N_i は化学種 i に関する許容絶対モル数である．この N_i は反復計算の中で繰り返し計算する．

3 参考文献

川上一郎 著, 数値計算の基礎, 日本大学原子力研究所 (日本大学理工学部) 計算物理教室, 1997.