北星論集(経) 第 56 巻 第 1 号(通巻第70号) September 2016

研究ノート

キャンベラ距離とニュートン法

吉 田 知 行

目次

［要旨］

キャンベラ距離の性質と二次無理数の連分数・ニュートン近似への応用について論ずる。日本の考古学では，新井宏が鉛同位体法に導入した類似度指数として知られている。キャンベラ距離は，２次無理数の連分数近似や，ニュートン近似の収束の見積もりにも使える。本論文では，キャンベラ距離が距離の公理を満たすことを証明する。さらにこの距離は，二次無理数のニュートン近似や連分数近似の具体的表示としても現れる。また，キャンベラ距離（新井の距離）の誤差評価について見積もりを与えた。これは鉛同位体法に使われる。

1 新井による鉛同位体の類似度指数

現代の考古学においては，単に遺跡を発掘し，遺跡や遺物の記録を取るだけではなく，科学的方法の導入が著しい。鉛同位体法もその一つで，青銅に含まれる鉛(Pb)の安定同位体（質量数204，206，207，208）の割合を質量分析計で測定し，その割合から青銅に含まれる鉛の産地，青銅の製造地や製造法を解明するする方法である。

ある鉛 A に含まれる同位体の割合を，質量数の順に (a_0, a_1, a_2, a_3) とする $(a_0 + a_1 + a_2 + a_3 = 1)$。これを同位体率という。実際には同位体比 $[a1/a_0, a_2/a_0, a_3/a_0]$ が使われることが多い。同位体率と同位体比は片方からもう一方を容易に計算出来る。

鉛同位体法では，青銅器や鉱山の鉛同位体比の散布図から鉛の来歴を考察する。まず問題になるのは，鉛の一致判定法である。もしふたつの鉛の同位体率(あるいは比)が一致していれば，２つの鉛は同じ産地（あるいは同じ素材）から来たと考えられる。しかしこの単純な問題でさえ，鉛同位体法特有の困難がある。

1. 鉛同位体ごとに存在比が相当違う。とくに鉛 204 は数パーセントしか含まれていない。
2. 同位体比の測定誤差が大きく，同位体比ごとに差がある。古いものでは a_1/a_0,

キーワード：新井の類似度指数，キャンベラ距離，二次無理数，ニュートン法，連分数近似

a_2/a_1, a_3/a_0 の順に，0.06%， 0.02%，0.02% といわれている。

3. 貴重な青銅器の鉛同位体比は，一般に再測定ができない。そのためしばしば信頼性の劣る古い測定値を使わざるを得ない。
4. 一致の判定基準が明瞭でない。

このような状況で使える距離関数として，新井宏が提唱した**類似度指数**(非類似度指数，あるいは新井の距離と呼んだ方がよい)がある。

$$D_A(\boldsymbol{a}, \boldsymbol{b}) = \frac{1}{4}\sum_{\alpha=0}^{3}\left|\frac{a_\alpha - b_\alpha}{a_\alpha + b_\alpha}\right| \tag{1}$$

標準偏差で割って基準化する代わりに，$a_\alpha + b_\alpha$ で割って，同位体ごとの同位体率の違いを消している。新井はこの類似度指数を用いて，鉛の一致の問題を研究した。

本論文の著者の吉田は，類似度数の数学的な性質に関心を持った。後述するように，この量が二次無理数の連分数近似やニュートン法の収束に関して現れることを知っていたからである。まったく異なる分野に同じ新井の距離が表れることに驚いたものである。鉛同位体法との関連は「数学セミナー」で論じた。また，新井の距離が，データ解析の分野でキャンベラ距離と呼ばれていることも最近知った。

2 三角不等式

新井の距離は，キャンベラ(Canberra)距離とも呼ばれ，クラスター分析で使われる距離関数のオプションになっている。重み付きマンハッタン距離の一種である。

ふたつの n 次元データ $\boldsymbol{x} = (x_1, \ldots, x_n)$，$\boldsymbol{y} = (y_1, \ldots, y_n)$ $(x_i, y_i > 0)$ の(n 次元)キャンベラ距離は

$$\boldsymbol{D}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n}\left|\frac{x_i - y_i}{x_i + y_i}\right| \tag{2}$$

で定義される。新井の距離のように，全体に $1/n$ を乗ずる流儀もある。

$$-1 < \frac{x-y}{x+y} < 1 \quad (x, y > 0)$$

なので，

$$\boldsymbol{D}(\boldsymbol{x}, \boldsymbol{y}) \le n \tag{3}$$

である。

とくに 1 次元キャンベラ距離は

$$D(x, y) = \left|\frac{x-y}{x+y}\right|, \quad (x, y > 0) \tag{4}$$

で与えられる。あきらかに

$$\boldsymbol{D}(\boldsymbol{x}, \boldsymbol{y}) = \sum_{i=1}^{n} D(x_i, y_i) \tag{5}$$

である。

$$|z| = \begin{cases} D(1+z, 1-z) & (|z| < 1) \\ 1/D(1+z^{-1}, 1-z^{-1}) & (|z| > 1) \end{cases} \tag{6}$$

なので，絶対値を含む多くの距離をキャンベラ距離で書き換えることができる。

また $|x' - y'| < x + y$ なら

$$\left|\frac{x' - y'}{x+y}\right| = D(x+y+x'-y', x+y-x'+y') \tag{7}$$

である。

定理 1 キャンベラ距離 $\boldsymbol{D}$ は距離の公理を満たす。すなわち，任意の n 次元正ベクトル $\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{z}$ に対し，

(D1) $\boldsymbol{D}(\boldsymbol{x}, \boldsymbol{y}) = 0 \iff \boldsymbol{x} = \boldsymbol{y}$;
(D2) $\boldsymbol{D}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{D}(\boldsymbol{y}, \boldsymbol{x})$;
(D3) $\boldsymbol{D}(\boldsymbol{x}, \boldsymbol{z}) \le \boldsymbol{D}(\boldsymbol{x}, \boldsymbol{y}) + \boldsymbol{D}(\boldsymbol{y}, \boldsymbol{z})$.

(証明)(D1), (D2) はあきらかなので，三角不等式 (D3) を示す。さらに $n = 1$ と仮定できる。したがって

$$D(x, y) = \left|\frac{x-y}{x+y}\right|$$

について三角不等式を示せばよい：

$$D(x,y)+D(y,z)\geq D(x,z). \tag{8}$$

（注意：等号は，$y=x$ または $y=z$ のときに限る。）あとの議論でも使うので，$a:=y+z$, $b:=z+x$, $c:=x+y$ と置く。このとき

$$\begin{cases} x=\dfrac{-a+b+c}{2} \\ y=\dfrac{a-b+c}{2} \\ z=\dfrac{a+b-c}{2} \end{cases} \tag{9}$$

$$\begin{cases} y-z &= c-b \\ z-x &= a-c \\ x-y &= b-a \end{cases} \tag{10}$$

さらに三角不等式が成り立つ：

$$\begin{cases} a<b+c \\ b<c+a \\ c<a+b \end{cases} \tag{11}$$

この条件は，辺の長さが a,b,c であるような三角形の存在条件でもある。したがって問題の三角不等式は

$$\frac{|b-a|}{c}+\frac{|c-b|}{a}\geq\frac{|c-a|}{b} \tag{12}$$

に同等である。同じことだが

$$ab|b-a|+bc|c-b|\geq ac|c-a| \tag{13}$$

に同値である。この不等式は a と c の入れ替えで対称なので，初めから $a<c$ と仮定できる（$a=c$ 場合は明かである）。

(i) $a\leq b\leq c$ の場合。

$$\begin{aligned} *:=&\ ab|b-a|+bc|c-b|-ac|c-a| \\ =&\ ab(b-a)+bc(c-b)-ac(c-a) \\ =&\ (b-a)(c-b)(c-a)\geq 0 \end{aligned}$$

となって正しい。等号成立は $a=b$ または $b=c$ の場合に限る。

(ii) $a\leq c\leq b$ の場合。$ac\leq ab\leq bc$ なので，

$$\begin{aligned} *:=&\ ab|b-a|+bc|c-b|-ac|c-a| \\ \geq&\ ac|b-a|+ac|c-b|-ac|c-a| \\ =&\ ac(|b-a|+|c-b|-|c-a|) \\ =&\ 2ac(b-c)\geq 0 \end{aligned}$$

で正しい。等号成立は $b=c$ の場合に限る。

(iii) $b\leq a\leq c$ の場合。

$$\begin{aligned} *:=&\ ab|b-a|+bc|c-b|-ac|c-a| \\ =&\ ab(a-b)+bc(c-b)-ac(c-a) \\ =&\ (a-b)\,(ab+(a+b-c)c)\geq 0 \end{aligned}$$

等号成立は，$b=a$ の場合に限る。

以上により問題の不等式 (13) が示された。また等号は $b=a$ または $b=c$ のときに限る。(8) とそこでの等号成立の条件もこれからしたがう。

注意． n 次元キャンベラ距離に関する三角不等式 (定理 1(3)) で等号成立の条件は，任意の i で，$x_i=y_i$ または $y_i=z_i$ が成り立つことである。

3　三角形に関する不等式

キャンベラ距離の三角不等式からしたがう不等式を紹介する。

定理 2　開区間 $(0,\pi/2)$ において

$$D(\alpha,\beta)=\left|\frac{\sin(\alpha-\beta)}{\sin(\alpha+\beta)}\right| \tag{14}$$

は距離関数である。

(証明)$x=\tan\alpha$, $y=\tan\beta$ と置けば，

$$\begin{aligned} \frac{x-y}{x+y}&=\frac{\tan\alpha-\tan\beta}{\tan\alpha+\tan\beta} \\ &=\frac{\sin(\alpha-\beta)/\cos\alpha\cos\beta}{\sin(\alpha+\beta)/\cos\alpha\cos\beta} \\ &=\frac{\sin(\alpha-\beta)}{\sin(\alpha+\beta)} \end{aligned}$$

したがって $D(\alpha,\beta)$ はキャンベラ距離 $D(x,y)$ に等しい。

正数 x, y, z から，$a := y+z$, $b := z+x$, $c := x+y$ と置けば，a, b, c は三角形 ABC の３辺の長さであり，逆に三角形 ABC の３辺の長さはこのように表示される。

定理３ a, b, c が三角形 ABC の３辺の長さなら，

$$ab|a-b| + bc|b-c| \geq ac|a-c|. \qquad (15)$$

この不等式が，キャンベラ距離の三角不等式に同値であることはすでに示した。

４ 等長変換

$\boldsymbol{D}(\boldsymbol{x}, \boldsymbol{y})$ は n 次元キャンベラ距離，D は一次元キャンベラ距離を表す．正ベクトル $\boldsymbol{x} = (x_1, \ldots, x_n)$, $\boldsymbol{y} = (y_1, \ldots, y_n)$ に対し

$$\boldsymbol{x} \cdot \boldsymbol{y} := (x_1 y_1, \ldots, x_n y_n),$$
$$\boldsymbol{y}/\boldsymbol{x} := (y_1/x_1, \ldots, y_n/x_n)$$

と置く。また，$\mathbf{1} = (1, \ldots, 1)$ によってオールワンベクトルを表す。

定義. $\mathbb{R}^n_{>0}$ 上のキャンベラノルムを

$$\boldsymbol{D}(\boldsymbol{z}) := D(\mathbf{1}, \boldsymbol{z}) = \sum_{i=1}^{n} \left| \frac{1-z_i}{1+z_i} \right| \qquad (16)$$

で定義する。とくに１次元キャンベラノルムは

$$D(z) := D(1, z) = \left| \frac{1-z}{1+z} \right| \qquad (17)$$

である．キャンベラノルムはキャンベラ距離と同じ記号 $\boldsymbol{D}(\boldsymbol{z})$ と $D(z)$ で表す。

補題４ (1) $\boldsymbol{D}(\boldsymbol{x}, \boldsymbol{y}) = D(\boldsymbol{y}/\boldsymbol{x})$.
(2) $D(D(z)) = \min(z, 1/z)$ $(z > 0)$.

(証明) 直接計算による。

キャンベラ距離の元になった分数式

$$\frac{1-z}{1+z}$$

は射影直線 $\boldsymbol{P}^1 = (\mathbb{R}^2 - \{(0,0)\})/\sim$ 上の一次変換である。上の補題の (2) はこの一次変換の位数が２であることを意味する:

$$\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}^2 = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

ただしキャンベラ距離の値は１未満なので，計算途中で１を超えた場合は逆数で置き換えることをする。

キャンベラ距離の等長変換を定義する。例えば，$\boldsymbol{c} = (c_1, \cdots, c_n)$ に対し，成分ごとの積 $(\boldsymbol{x}, \boldsymbol{y}) \mapsto (\boldsymbol{c} \cdot \boldsymbol{x}, \boldsymbol{c} \cdot \boldsymbol{y})$ はキャンベラ距離を保存する。したがってこの変換は等長変換である。しかし，$\boldsymbol{D}(\boldsymbol{x}, \boldsymbol{y}) = \boldsymbol{D}(\boldsymbol{y}/\boldsymbol{x})$ なので，このような変換は自明な等長変換と言える。ここでは，より本質的な等長変換の定義を与える。

定義. n 次元キャンベラノルム $\boldsymbol{D}$ に関する**等長変換** (isometry) とは，$\mathbb{R}^n_{>0}$ 上の位相変換 (自分自身への同相写像) $\boldsymbol{f}$ でノルムを不変にするものである。

$$\boldsymbol{D}(\boldsymbol{f}(\boldsymbol{z})) = \boldsymbol{D}(\boldsymbol{z}) \qquad (18)$$

キャンベラノルムに関する等長変換 $\boldsymbol{f}$ があれば，

$$(\boldsymbol{x}, \boldsymbol{y}) \mapsto (\boldsymbol{x}, \boldsymbol{f}(\boldsymbol{y}/\boldsymbol{x}) \cdot \boldsymbol{x})$$

はキャンベラ距離を保存する (キャンベラ距離に関する等長変換)。

例. $z \mapsto 1/z$ は，１次元キャンベラノルムに関する等長変換である。対応するキャンベラ距離に関する等長変換は $(x, y) \mapsto (y, x)$ (あるいは $(x, y) \mapsto (1/x, 1/y)$) である。この変換は，$\mathbb{R}_{>0}$ の乗法群の自己同形写像である。

n 次元キャンベラ距離 (またはノルム) に関する等長変換全体は合成に関して群を成す。これを**等長変換群**という。２次元以上の等張変換群の構造は分からない。

定理５ 乗法群 $\mathbb{R}^n_{>0}$ の自己同形写像になっているような (ノルムに関する) 等長変換の群は，B_n 型ワイル群に同型である．

用語の説明と証明には準備が必要なので，ここでは省略する。代わりに１次元の場合に群同形の仮定なしに証明を与えておく。

補題６　１次元キャンベラノルムに関する等長変換群は位数２で，$z \mapsto 1/z$ で生成される。

(証明)α が $\mathbb{R}_{>0}$ からそれ自身への等長写像なら，

$$\left|\frac{1-\alpha(z)}{1+\alpha(z)}\right| = \left|\frac{1-z}{1+z}\right| \quad (z>0)$$

これを解くと

$$\alpha(z) = z \quad \text{または} \quad 1/z \tag{19}$$

となる。とくに $\alpha(1)=1$ である。

$$A := \{z>0 \mid \alpha(z)=z\},$$
$$B := \{z>0 \mid \alpha(z)=1/z\}$$

はどちらも $\mathbb{R}_{>0}$ の閉集合で，

$$A \cup B = \mathbb{R}_{>0}, \quad A \cap B = \{1\}.$$

開区間 $(0,1)$ と $(1,\infty)$ はどちらも連結なので，次のどちらかが一方が成り立つ:
(a) $(0,1) \subset A$ かつ $(0,1) \cap B = \emptyset$.
(a') $(0,1) \subset B$ かつ $(0,1) \cap A = \emptyset$.

同様に次のどちらか一方が成り立つ:
(b) $(1,\infty) \subset A$ かつ $(1,\infty) \cap B = \emptyset$.
(b') $(1,\infty) \subset A$ かつ $(1,\infty) \cap B = \emptyset$.

結局以下の４つの条件のひとつが成り立つ:

(ab) $\alpha(z) = z \quad (z>0)$.
(a'b') $\alpha(z) = 1/z \quad (z>0)$.
(ab') $\alpha(z) = \begin{cases} z & (0<z\le 1) \\ 1/z & (z \ge 1) \end{cases}$
(a'b) $\alpha(z) = \begin{cases} 1/z & (0<z\le 1) \\ z & (z \ge 1) \end{cases}$

(ab') と (a'b) の場合は，距離を保つが，同相でないので，条件に会わない。結局 (ab) と (a'b') のどちらかが起こる。すなわち自己同形群は位数２である。

５　双曲線関数による表示

キャンベラ距離の定義は，双曲線正接関数 tanh を使えば，マンハッタン距離に似た形になる。まず n 次元正値ベクトル $\boldsymbol{x} = (x_1,\ldots,x_n)$ は，普通の n 次元数ベクトル $\boldsymbol{\theta} = (\theta_1,\ldots,\theta_n)$ によって $\boldsymbol{x} = e^{\boldsymbol{\theta}}$，ここで $x_i = e^{\theta_i}$ と表せることに注意する。このとき n 次元キャンベラ距離は

$$\begin{aligned}\boldsymbol{D}(\boldsymbol{x},\boldsymbol{y}) &= \boldsymbol{D}(e^{\boldsymbol{\phi}}, e^{\boldsymbol{\psi}}) \\ &= \sum_{i=1}^{n}\left|\tanh\frac{\phi_i-\psi_i}{2}\right| \\ &= \sum_{i=1}^{n}\tanh\left|\frac{\phi_i-\psi_i}{2}\right| \end{aligned} \tag{20}$$

ここで $\boldsymbol{\psi} = (\psi_i), \boldsymbol{\phi} = (\phi_i)$ は数ベクトルで，

$$\tanh\theta = \frac{e^{\theta}-e^{-\theta}}{e^{\theta}+e^{-\theta}} = \frac{e^{2\theta}-1}{e^{2\theta}+1}$$

とくに，１次元キャンベラ距離について

$$D(e^{\phi}, e^{\psi}) = \tanh\left|\frac{\phi-\psi}{2}\right|. \tag{21}$$

６　二次無理数のニュートン近似

連続微分可能な関数 $f(x)$ に対し，適当な x_0 から始まるニュートン列 $\{x_n\}$ を漸化式

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x_n)} \tag{22}$$

で定義する。このとき，$\{x_n\}$ は $f(x)$ の零点 ω に収束すると期待できる。

$D>0$ に対する２次無理数 $\omega = a+b\sqrt{D}$ $(a,b \in \mathbb{R})$ のニュートン近似を考える。ω の共役を $\omega' = a-b\sqrt{D}$ とすれば，ω, ω' は

$$f(x) := (x-\omega)(x-\omega') = x^2-2ax+a^2-b^2D$$

の解である。とくに

$$\omega+\omega' = 2a, \quad \omega\omega' = a^2-b^2D.$$

したがって ω のニュートン近似列 $\{x_n\}$ の漸

化式は

$$\begin{aligned} x_{n+1} &= x_n - \frac{x_n^2 - 2ax_n + \omega\omega'}{2(x_n - a)} \\ &= \frac{{x_n}^2 - \omega\omega'}{2(x_n - a)} \end{aligned} \tag{23}$$

で与えられる。今の場合はニュートン近似列 $\{x_n\}$ の具体的式が簡単に得られる。

定理 7

$$\frac{x_n - \omega}{x_n - \omega'} = \left(\frac{x_0 - \omega}{x_0 - \omega'}\right)^{2^n} \tag{24}$$

$x_0' = (x_0 - \omega)/(x_0 - \omega')$ と置けば

$$x_n = \frac{\omega - \omega' {x_0'}^{2^n}}{1 - {x_0'}^{2^n}} \tag{25}$$

(証明)

$$\begin{aligned} \frac{x_{n+1} - \omega}{x_{n+1} - \omega'} &= \frac{\dfrac{{x_n}^2 - \omega\omega'}{2(x_n - a)} - \omega}{\dfrac{{x_n}^2 - \omega\omega'}{2(x_n - a)} - \omega'} \\ &= \frac{{x_n}^2 - \omega\omega' - 2\omega(x_n - a)}{{x_n}^2 - \omega\omega' + 2\omega'(x_n - a)} \\ &= \frac{(x_n - \omega)^2}{(x_n - \omega')^2} = \left(\frac{x_n - \omega}{x_n - \omega'}\right)^2. \end{aligned}$$

この等式を繰り返し使うことによって定理の結果が得られる。

この公式はひょっとすると知られていないかもしれない。

定理 8 簡単のため $b > 0$ と仮定する。$x_0 > a$ なら，ニュートン列 $\{x_n\}$ は $\omega = a + b\sqrt{D}$ に2次収束する。$x_0 < a$ なら，$\{x_n\}$ は $\omega' = a - b\sqrt{D}$ に2次収束する。$x_0 = a$ なら，収束しない。

(証明)$\{x_n\}$ が ω に収束するための条件は

$$-1 < \frac{x_0 - \omega}{x_0 - \omega'} < 1$$

である。$b > 0$ なので $\omega > \omega'$。したがって $x_0 > a$ なら

$$x_0 > a = (\omega + \omega')/2 > \omega'$$

今の場合，ω への収束条件

$$-x_0 + \omega' < x_0 - \omega < x_0 - \omega'$$

は満たされている。$x_0 < a$ の場合も同様である。前定理により収束は2次収束である。最後に $x_0 = a$ の場合，x_1 が定義されない。

定理 9 $\omega = \sqrt{D}$ の場合のニュートン列の満たす漸化式は

$$\frac{x_n - \sqrt{D}}{x_n + \sqrt{D}} = \left(\frac{x_0 - \sqrt{D}}{x_0 + \sqrt{D}}\right)^{2^n} \tag{26}$$

とくにキャンベラ距離について

$$D(x_n, \sqrt{D}) = D(x_0, \sqrt{D})^{2^n}$$

すなわち $\{x_m\}$ は，キャンベラ距離に関して $\sqrt{D}$ に2次収束する。

7 記号ニュートン法

$D(t) = 1 + t$ とし，$\omega(t) = \sqrt{1+t}$ を有理数係数形式的べき級数環 $\mathbb{Q}[[t]]$ の中で考える。記号ニュートン法によれば，漸化式

$$x_{n+1}(t) \quad = \frac{x_n(t)^2 + 1 + t}{2x_n(t)}$$

で定義されるニュートン近似多項式列は，初期多項式 $x_0(t)$ に関する適当な条件の下で $\sqrt{1+t}$ に収束する。

定理 10 任意の $x_0(t) \in 1 + t\mathbb{Q}[[t]]$ に対し

$$\frac{x_n(t) - \sqrt{1+t}}{x_n(t) + \sqrt{1+t}} = y_0(t)^{2^n} \tag{27}$$

である。ここで

$$y_0(t) := \frac{x_0(t) - \sqrt{1+t}}{x_0(t) + \sqrt{1+t}} \tag{28}$$

さらに $\{x_n(t)\}$ は形式的べき級数環 $\mathbb{Q}[[t]]$ おいて $\sqrt{1+t}$ に「2次収束」する:

$$x_n(t) \equiv \sqrt{1+t} \pmod{t^{2^n}}$$

(証明) 漸化式は普通のニュートン法と同様に得られる。他は形式的べき級数関連の定義から直ちに得られる。

8　実数の連分数展開

モジュラー変換とは射影直線 $\boldsymbol{P}(\mathbb{R}) = \mathbb{R} \cup \{\infty\}$ 上の変換

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \cdot z := \frac{az+b}{cz+d}$$

のことである。ここで a, b, c, d は整数で，$ad - bc = \pm 1$ とする。モジュラー変換の合成は行列の積と同じである。

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \circ \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix} = \begin{bmatrix} aa' + bc' & ab' + bd' \\ ca' + dc' & cb' + dd' \end{bmatrix}$$

実数 x に対し，$[x]$ と $\{x\}$ はそれぞれ x の整数部分と小数部分を表す。したがって，$[x]$ は x を超えない最大の整数である。このとき

$$x = [x] + \{x\}, \quad 0 \leq \{x\} < 1$$

ここまでの準備の元で，実数の連分数展開について述べる。正の実数 ω から初めて，数列 $\{\omega_n\}$, $\{q_n\}$, $\{u_n\}$ を以下のように帰納的に定める。

(1) $\omega_0 := \omega$.

(ii) ある ω_n まで決まったとして，

$$q_n := [\omega_n], u_n := \{\omega_n\}, \omega_{n+1} := 1/u_n$$

と定義する。

このとき

$$q_n \in \mathbb{N}, 0 < u_n < 1, \omega_{n+1} > 1.$$

以下では簡単のため ω は正の無理数とする。こうすれば，上の過程は止まることがない。すなわち ω_n は整数でなく，$u_n \neq 0$，したがって $\omega_{n+1} = 1/u_n$ が定義される。

この過程はモジュラー変換を使って表せる。上の数列の構成を見ると，

$$\begin{aligned} \omega_n &= a_n + u_n = q_n + \frac{1}{\omega_{n+1}} \\ &= \frac{q_n \omega_{n+1} + 1}{\omega_{n+1}} = \begin{bmatrix} q_n & 1 \\ 1 & 0 \end{bmatrix} \cdot \omega_{n+1} \end{aligned}$$

となっている。これを繰り返し使うと

$$\omega = \begin{bmatrix} q_0 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} q_{n-1} & 1 \\ 1 & 0 \end{bmatrix} \cdot \omega_n$$

となる。また

$$\begin{aligned} \omega &= q_0 + \frac{1}{\omega_1} \\ &= q_0 + \cfrac{1}{q_1 + \cfrac{1}{\omega_2}} \\ &= q_0 + \cfrac{1}{q_1 + \cfrac{1}{q_2 + \ddots}} \end{aligned}$$

と連分数でも表せる。最後の連分数を $[q_0; q_1, q_2, \cdots]$ と表す。

$$\begin{bmatrix} q_0 & 1 \\ 1 & 0 \end{bmatrix} \cdots \begin{bmatrix} q_{n-1} & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} P_n & P_{n-1} \\ Q_n & Q_{n-1} \end{bmatrix}$$

で数列 $\{P_n\}, \{Q_n\}$ を定めると

$$\begin{bmatrix} P_{n+1} \\ Q_{n+1} \end{bmatrix} = \begin{bmatrix} P_n q_n + P_{n-1} \\ Q_n q_n + Q_{n-1} \end{bmatrix} \tag{29}$$

である。

これから

$$\frac{P_n}{Q_n} - \frac{P_{n-1}}{Q_{n-1}} = \frac{(-1)^{n-1}}{Q_{n-1} Q_n}$$

したがって有理数 $P_n/Q_n = [q_0; q_1, \ldots, q_{n-1}]$ は ω の近似有理数を与える。

詳細は高木「初等整数論講義」を参照。

9　近似有理数とキャンベラ距離

$\omega = \sqrt{D}$ の形の実無理数だけを扱う。ここで $D > 0$ は平方因子を持たない整数とする。

定理 11

$$\frac{\sqrt{D} - P_n/Q_n}{\sqrt{D} - P_n/Q_n} = \left(\frac{\sqrt{D} - P_1/Q_1}{\sqrt{D} - P_1/Q_1} \right)^n$$

とくに P_n/Q_n は $\sqrt{D}$ に，キャンベラ距離に関しても普通の絶対値に関しても 1 次収束する。

証明は $\{P_n\}$ と $\{Q_n\}$ に関する漸化式から得られる。

10　キャンベラ距離の一般化

キャンベラ距離の一般化はいろいろ考えられる。ここでは 1 次元キャンベラ距離の一般化として，次の形のものを考える。

$$D_\lambda(x,y) = \frac{|x-y|}{(x+y)^\lambda}, \quad x,y>0. \qquad (30)$$

ここで λ は実数である。

予想. $D_\lambda(x,y)$ は距離の公理を満たす．すなわち次の三角不等式が成り立つ。

$$\frac{|x-y|}{(x+y)^\lambda} + \frac{|y-z|}{(y+z)^\lambda} \geq \frac{|x-z|}{(x+z)^\lambda} \qquad (31)$$

同じことだが，a,b,c が三角形の辺の長さなら

$$a^\lambda b^\lambda |a-b| + b^\lambda c^\lambda |b-c| \geq a^\lambda c^\lambda |a-c| \quad (32)$$

$(\lambda=0)$ 正しい。これは絶対値に対する三角不等式に外ならない：

$$|x-y|+|y-z| \geq |x-z|$$

$(\lambda=1)$ 正しい。キャンベラ距離である。
$(\lambda=-1)$ 正しい。$(x+y)|x-y| = |x^2-y^2|$ なので，三角不等式に帰着される。
$(\lambda=2,3)$ 正しい。場合分けと計算が複雑だが，数式処理システムを使って確かめられる。

11　キャンベラ距離の誤差評価

n 次元キャンベラ距離 $\boldsymbol{D}(\boldsymbol{x},\boldsymbol{y})$ の値が，$\boldsymbol{x},\boldsymbol{y}$ の誤差にどう影響されるかについて評価式を与えておく。これは鉛同位体法においてデータの誤差が新井の距離の誤差への影響を調べる必要性から来た問題である。

定理 12　n 次元キャンベラ距離 $\boldsymbol{D}(\boldsymbol{x},\boldsymbol{y})$ の絶対誤差は，データ $x_1,\ldots,x_n;y_1,\ldots,y_n$ の相対誤差の和に近似的に等しい。

$$|d\boldsymbol{D}(\boldsymbol{x},\boldsymbol{y})| \sim \sum_{i=1}^n \frac{|d\log x_i| + |d\log y_i|}{2}$$

（証明）対数微分法を使って 1 次元キャンベラ距離を全微分すると

$$d\log\frac{|x_i-y_i|}{x_i+y_i} = \frac{2y_i dx_i - 2x_i dy_i}{x_i^2-y_i^2}$$

したがって

$$d\boldsymbol{D}(\boldsymbol{x},\boldsymbol{y}) = \sum_i \frac{2\mathrm{sgn}(x_i-y_i)(y_i dx_i - x_i dy_i)}{(x_i+y_i)^2}$$

これの絶対値を取ればよい。

参考文献

[1] 新井宏『理系の視点からみた「考古学」の論争点』大和書房（2007）
[2] 高木貞治『初等整数論講義第 2 版』共立出版（1971）
[3] 吉田知行「鉛同位体法の数理（1）～（4）」数学セミナー（2015/8-11）

校正時の追加。(1) キャンベラ距離の定義を

$$\boldsymbol{D}(\boldsymbol{x},\boldsymbol{y}) = \sum_{i=1}^n \frac{|x_i-y_i|}{|x_i|+|y_i|}$$

とし，定義域を $(x_i,y_i) \neq (0,0), (i=1,2,\cdots,n)$ とする流儀もある。値域は ∞ も許す。
(2) 5 節に関係することだが，速度 u と速度 v の相対論的合成則は

$$u \oplus v = \frac{u+v}{1+\dfrac{uv}{c^2}} \quad (c \text{ は真空中の光速})$$

で与えられる。同じことだが，

$$\frac{c-u\oplus v}{c+u\oplus v} = \left(\frac{c-u}{c+u}\right)\left(\frac{c-u}{c+u}\right)$$

すなわち，キャンベラ距離について，

$$D(c,u\oplus v) = D(c,u)D(c,v).$$