研究ノート

等差数列の標準偏差の整数性とペル方程式$x^2-3y^2=1$.

吉 田 知 行

目次

[要旨]

本論文のきっかけは,標準偏差が有理数になるような簡単な数列があれば,統計の講義中の学生の計算練習にふさわしいという問題意識から来た.等差数列の中からそのような数列を探した.この問題を解くには,ペル方程式を解く必要がある.長さ$N=7, 26, 97$ など無数の解が存在する.またこの問題の拡張は,初等整数論のみならず,26次元ローレンツ・リーチ格子の等方的ベクトルの存在にも関係している.2次無理数の近似分数についての新しい公式や,新井の距離との関係も論じた.

1 問題の発端

本学で統計学 I の講義を担当した.標準偏差を計算させる問題をよく出すが,そのとき答えが単純だと学生が答えるのも,教員が採点するのも手間がかかかららず楽である.そこで,整数から成る等差数列で,標準偏差が整数になるものを探した.例えば,

x:1, 2, 3, 4, 5, 6, 7

がそのような例になっている.実際,平均値$\overline{x}=4$ で標準偏差は

$$\sigma(x)=\sqrt{\frac{1}{7}\sum_{k=1}^{7}(k-4)^2}=2.$$

確かに標準偏差は整数である.

この論文では,他に標準偏差が有理数であるような等差数列を探す.さらに関連する数学的な背景を解説する.

2 ペル方程式への問題の書き換え

数列$x_1, x_2, \ldots, x_N$ をサイズNのデータと見なし,単にxと略記する.

xの平均と分散は

$$\overline{x}=\frac{1}{N}\sum_{k=1}^{N}x_k, \tag{2.1}$$

$$\sigma^2(x)=\overline{(x-\overline{x})^2}=\overline{x^2}-(\overline{x})^2 \tag{2.2}$$

で与えられる.定数a, b に対し,

$$\overline{ax+b}=a\overline{x}+b, \tag{2.3}$$

$$\sigma^2(ax+b)=a^2\sigma^2(x) \tag{2.4}$$

である.ここで$ax+b$ は$ax_1+b, \ldots, ax_N+b$から成るデータである.

$x(a, b)$ を公差a,初項$a+b$ の等差数列とする.とくに断らなければ長さはNとする.

キーワード:標準偏差,ペル方程式,連分数

たとえば$x(1, 0)$ は数列$1, 2, \dots, N$ であり，
$x(a, b) = ax(1, 0) + b$
である．したがって数列$x = x(1, 0)$ の平均と分散は，

$$\overline{x(1,0)} = \frac{1}{N}\sum_{k=1}^{N} k = \frac{N+1}{2}, \qquad (2.5)$$

$$\begin{aligned}\sigma^2(x(1,0)) &= \overline{x^2} - (\overline{x})^2 \\ &= \frac{1}{N}\sum_{k=1}^{n} N^2 - \left(\frac{N+1}{2}\right)^2 \\ &= \frac{(N+1)(2N+1)}{6} - \frac{(N+1)^2}{4} \\ &= \frac{(N+1)(N-1)}{12} \qquad (2.6)\end{aligned}$$

で与えられる．これより，一般の等差数列 $x(a, b)$ の平均と分散は，

$$\overline{x(a,b)} = \frac{N+1}{2}a + b, \qquad (2.7)$$

$$\sigma^2(x(a,b)) = \frac{(N+1)(N-1)}{12}a^2 \qquad (2.8)$$

となる．

かくて問題は次のように表される：

問題A．自然数a とN に対し，標準偏差

$$\sqrt{\frac{N^2-1}{12}a^2}$$

はいつ整数になるか？

問題を有理数まで拡張すると，問題Aは，自然数N と有理数a に対し，標準偏差の値

$$\sqrt{\frac{N^2-1}{12}a^2} = \left|\frac{a}{3}\right|\sqrt{\frac{N^2-1}{3}}$$

がいつ有理数に成るかという問題になる．右辺の$|a/3|$は有理数なので，$\sqrt{\frac{N^2-1}{3}}$ がいつ有理数になるかということである．

N をx と書き，$\sqrt{\frac{N^2-1}{3}}$ をy と書けば，問題は次の様になる：

問題B．自然数x と有理数y に対する方程式

$$x^2 - 3y^2 = 1 \qquad (2.9)$$

を解け．

補題2.1　方程式（2.9）のy は整数である．したがって（2.9）はペル方程式である．

（証明）$(3y)^2 = 3(x^2-1)$ なので，$3y$ は代数的整数である．$3y$ は有理数でもあるので，$y' := 3y$ は整数である．$y'^2 = 9y^2 = 3(x^2-1)$ なので，y' は 3 の倍数となる．結局$y = y'/3$も整数である．（証終）

3　ペル方程式$x^2-3y^2 = 1$ の解．

一般のペル方程式$x^2 - Dy^2 = 1$（$D > 0$は平方数でない正の整数）に比べると，$D = 3$の場合は「最小解」が$(x_1, y_1) = (2, 1)$ と単純である．ペル方程式の一般論によれば，

$$x_n + y_n\sqrt{3} = (2+\sqrt{3})^n \quad (n = 1, 2, \dots) \qquad (3.10)$$

で整数列$\{x_n\}, \{y_n\}$ を定義すれば，(x_n, y_n) がペル方程式$x^2 - y^2 = 1$ のすべての解を与える．（3.10）の右辺を二項定理で展開することにより次の表現を得る：

$$x_n = \sum_{i=0}^{[n/2]} \binom{n}{2i} 2^{n-2i}3^i$$

$$y_n = \sum_{i=0}^{[(n-1)/2]} \binom{n}{2i+1} 2^{n-2i}3^i$$

しかし(x_n, y_n) を順に求めるなら，漸化式が便利である：

$x_1 = 2,\ y_1 = 1,$
$x_{n+1} = 2x_n + 3y_n,$
$y_{n+1} = x_n + 2y_n$

同じことだが，行列を使って

$$\begin{bmatrix} x_{n+1} \\ y_{n+1} \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}\begin{bmatrix} x_n \\ y_n \end{bmatrix} \qquad (3.11)$$

とも表せる．これより

$$\begin{bmatrix} x_n \\ y_n \end{bmatrix} = \begin{bmatrix} 2 & 3 \\ 1 & 2 \end{bmatrix}^n \begin{bmatrix} 1 \\ 0 \end{bmatrix} \qquad (3.12)$$

も得られる．

今必要なのはx_n（データサイズN）なので，次の漸化式も役に立つ．

$$x_{n+2} = 4x_{n+1} - x_n,\ x_0 = 1,\ x_1 = 2,$$
$$y_{n+2} = 4y_{n+1} - y_n,\ y_0 = 0,\ y_1 = 1.$$

便宜上$x_0 = 1, y_0 = 0$とした．

これによって問題Ａと問題Ｂの解は表１のようになる．nが奇数の場合，標準偏差σは半整数になる．したがって標準偏差を整数にするなら元の数列を2倍する必要がある．例えば，$n=3$の場合を見ると，$2, 4, \cdots, 52$の標準偏差は15と整数になる．実際の手計算の練習問題に出すなら$n=2$（$N=7$）くらいであろう：

$$\frac{(1-4)^2+(2-4)^2+\cdots+(7-4)^2}{7} = 2^2.$$

平方和の公式を知っている学生には$n=3 (N=26)$や$n=4$（$N=97$）も可能であろう．

n	$N=x_n$	y_n	σ
1	2	1	0.5
2	7	4	2.0
3	26	15	7.5
4	97	56	28.0
5	362	209	104.5
6	1351	780	390.0
7	5042	2911	1455.5
8	18817	10864	5932.0
9	70226	40545	20272.5
10	262087	151316	75658.0

表１　標準偏差が有理数になる等差数列
1,2,...;, N．標準偏差　$\sigma = y_n/2$.

４　その他の有限数列の和

（A）$\displaystyle\sum_{k=1}^{N} k = \frac{N(N+1)}{2}$が平方数$y^2$になるのはいつか．$x=2N+1$と置けば，これはペル方程式

$$x^2 - 8y^2 = 1$$

を解く問題に帰着される．

（B）$\displaystyle\sum_{k=1}^{N} k^2 = \frac{N(N+1)(2N+1)}{6}$が平方数$y^2$になるのはいつか．

Wikipedia（英語版）によれば，この問題は，フランス人数学者 Édouard Lucas（1875）による．自明な$N=1$を除けば，解は$N=24$だけである：

$$1^2+2^2+3^2+\cdots+24^2=70^2$$

［CS 99］によれば，この等式は

$$w = (0, 1, 2, \ldots, 24 \mid 70)$$

が26次元Lorentz-Leech 格子　$\mathrm{II}_{25,1}$の等方的ベクトル（isotropic vector）であることを意味する．次の等式もある：

$$1^2+3^2+5^2+\cdots+45^2+47^2+51^2=145^2.$$

（C）等差数列の３乗和が３乗数，４乗和が４乗数になることが自明な場合を除いて存在するかどうかは分からない．３乗和については楕円曲線上の有理点を求める問題に帰着される．$m \geq 3$の場合，等差数列のm次モーメントが有理数のm乗になることがあるかどうかも知られていないようである．さらに尖度や歪度についても同様の問題が考えられる．これらは，不定方程式の問題として興味深い．

（D）標準偏差は母集団標準偏差$\sigma(x)$を用いてきたが，不偏標準偏差$u(x)$でも対応する問題が考えられる．等差数列$x(1,0)$の場合

$$u(x) = \sqrt{\frac{N(N+1)}{12}}$$

となる．これが有理数であるための条件は$x=2N+1$, $y=4u(x)$と置いて，(x, y)がペル方程式$x^2-3y^2=1$の解になることである．表１から，たとえば$(7, 4)$，$(97, 56)$が解である．そのとき$N=3, 48$となる．対応する不偏標準偏差はそれぞれ1, 14である．

5 若干のコメント

一般のペル方程式 $x^2 - Dy^2 = 1$（D は平方数でない正整数）の整数解を求めるには，$\sqrt{D}$ の連分数展開を使うのが標準的である．ここでは，標準とは少し違う方法で，近似分数を効率よく求める方法を紹介する．

無理数 $\omega > 0$ に対し，$\omega_0 = \omega$ と置き，以下

$$\omega_n = k_n + \frac{1}{\omega_{n+1}},\ k_n = [\omega_n] \qquad (5.13)$$

で自然数列 $\{k_n\}_{n=0}^{\infty}$ と実数列 $\{\omega_n\}_{n=0}^{\infty}$ を定義する．ここで $[x]$ は実数 x の整数部分（ガウスの記号）である．したがって $x - [x]$ は x の小数部分である．このとき ω は

$$\omega = k_0 + \cfrac{1}{k_1 + \cfrac{1}{\ddots \cfrac{1}{k_n + \cfrac{1}{\omega_{n+1}}}}} \qquad (5.14)$$

と連分数展開される．この連分数を

$$\omega = [k_0, k_1, \ldots, k_n, \omega_{n+1}]$$

と表現する．$P_n/Q_n = [k_0, k_1, \ldots, k_n]$ を第 n 近似分数という．

$$\lim_{n\to\infty}[k_0, k_1, \ldots, k_n] = \omega$$

である．

$\omega = \sqrt{D}$ のときは次の等式で第 n 近似分数を求めることができる．

$$\frac{P_n/Q_n + \sqrt{D}}{P_n/Q_n - \sqrt{D}} = \left(\frac{k_0 + \sqrt{D}}{k_0 - \sqrt{D}}\right)^{n+1} \qquad (5.15)$$

この式はあまり知られていないようだが，数学的帰納法から直接，あるいは連分数に関する知られた式から容易に証明できる．

(5. 15) を使えば連分数に関する様々なことが分かる．$k_0 - \sqrt{D} < 0$ から

$$k_0 = P_0/Q_0 < P_2/Q_2 < \cdots < \omega < \cdots < P_3/Q_3 < P_1/Q_1 = k_1$$

が分かるし，$P_n/Q_n \to \omega$ も分かる．漸化式を求めることもできる．詳細は専門書に譲る．

ペル方程式 $x^2 - Dy^2 = 1$ の解と連分数との関係では，以下のことが成り立つ．

(ｉ) ペル方程式は必ず自然数解を持つ．

(ii) そのような解は，$\sqrt{D}$ の奇数番目の近似分数 P_n/Q_n のどれかである．

(iv) そのような解 (x,y) の中で，x が最小（同じことだが $x + y\sqrt{D}$ が最小の解（最小解という）を (x_1,y_1) とする．

$$x_n + y_n\sqrt{D} = (x_1 + y_1\sqrt{D})^n$$

で自然数 (x_n, y_n) を定義すれば，これはペル方程式 $x^2 - Dy^2 = 1$ の解であり，逆にこのペル方程式の解は必ずこの様にして得られる．

$D = 3$ の場合は議論が簡単である．２次体 $\mathbb{Q}(\sqrt{3})$ が単項イデアル環，したがって素元分解環であることが問題を簡単にするし，最小解が初項の $k_1 = 2$ から得られる (2, 1) である．

最後のコメントとして，(5.15) に現れた

$$D_A(x,y) = \left|\frac{x-y}{x+y}\right| \qquad (5.16)$$

について注意しておく．この量は正の実数 x, y 間の距離を与える（吉田2014）．例えば三角不等式

$$\left|\frac{x-z}{x+z}\right| \le \left|\frac{x-y}{x+y}\right| + \left|\frac{y-z}{y+z}\right|$$

が成り立つ．鉛同位体法に関する新井宏の研究に現れたので，新井の距離と呼ぶ．違った形でキャンベラ距離と言うこともある．

結論として，$\sqrt{D}$ と第 n 近似分数 Pn/Qn との間の新井の距離は次を満たす:

$$D_A(\sqrt{D}, P_n/Q_n) = D_A(\sqrt{D}, P_1/Q_1)^n$$

この式は，二次の無理数でも成り立ち，２次体の整数論における新たな知見をもたらす．

新井の距離について，同様の結果は，$\sqrt{D}$ のニュートン近似でも登場する．初等整数論や数値解析の一部では，このように新井の距

離が強力な威力を発揮する分野があるようだ.

[謝辞]

この研究は部分的に科学研究費基盤(C) #25400001(代表者.吉田知行) の支援を受けた.

参考文献

[Ar 07] 新井宏「理系の視点からみた「考古学」の論争点」大和書房 (2007).
[Ba 03] E.J.Barbeaqu, "Pell's Equation," Springer (2003).
[CS 99] J.H.Conway-N.J.A.Sloane, "Sphere packeings, lattices, and groups (Ver 3)," Springer (1999).
[Ta 71] 高木貞治「初等整数論講義」(第2版) 共立出版 (1971)
[Yo 14] 吉田知行「新井の距離と関連する三角不等式」数学セミナー2014年11月号