カタラン数

山上 滋

1. カタラン数とは

次の式で定義される自然数を、ベルギーの数学者E.C. Catalan (1814–1894) に因んで**カタラン数** (Catalan number) と呼ぶ。

$$C_n = \frac{1}{n+1}\binom{2n}{n}, \qquad n = 1, 2, 3, \dots.$$

ここで、$\binom{m}{n}$ は二項係数 ${}_mC_n$ を表す。カタラン数が自然数であることは、

$$C_n = \binom{2n}{n} - \binom{2n}{n+1}$$

なる表式からわかる。

$$C_1 = 1, C_2 = 2, C_3 = 5, C_4 = 14, C_5 = 42, C_6 = 132, C_7 = 429, C_8 = 1430, \dots$$

である。

自然数 $0 \leq m \leq n$ に対して、数

$$C_{n,m} = \frac{(n-m+1)(n+m)!}{m!(n+1)!}.$$

を並べることにより、パスカルの三角形と類似のものを得る。

1								
1	1							
1	2	2						
1	3	5	5					
1	4	9	14	14				
1	5	14	28	42	42			
1	6	20	48	90	132	132		
1	7	27	75	165	297	429	429	
1	8	35	110	275	572	1001	1430	1430

2. 三角形分割数

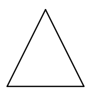
このような数の列は、多面体定理等で有名な L. Euler (1707–1783) にまで遡ることができる。オイラーは、次のような問題を考え、その解答を得た。

「与えられた凸 n 角形 $(n = 3, 4, 5, \dots)$ に対して、それを三角形に分割する方法の総数 E_n を求めよ。」

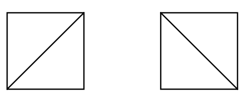
オイラーが与えた公式は、次の通り。

$$E_n = C_{n-2} = \frac{(2n-4)!}{(n-2)!(n-1)!}.$$

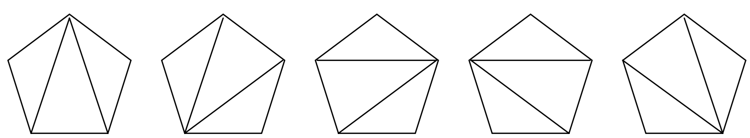
$E_3 = 1$



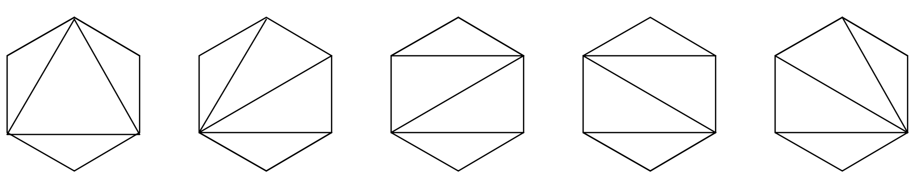
$E_4 = 2$

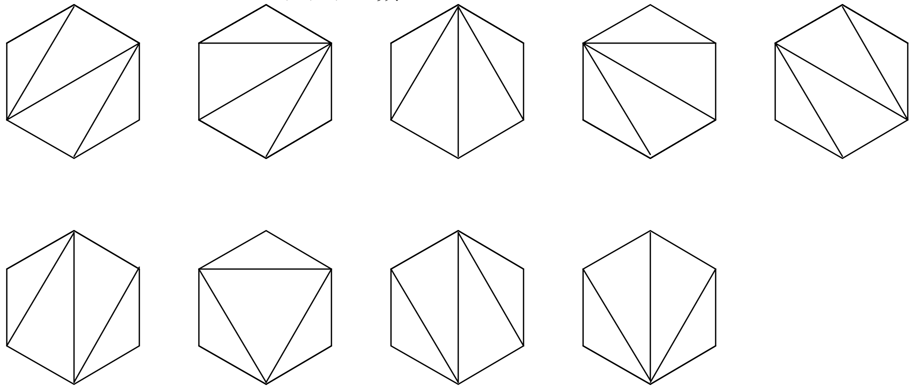


$E_5 = 5$



$E_6 = 14$





公式の証明の前に、$\{E_n\}$ の充たす**漸化式** (recurrence relation)

$$E_n = \sum_{k=2}^{n-1} E_k E_{n-k+1}, \qquad E_2 = 1$$

を導いておこう。

これは、凸 n 角形のある一辺に注目し、その辺を含む三角形の頂点の位置で場合わけをするとわかる (FIGURE 1)。

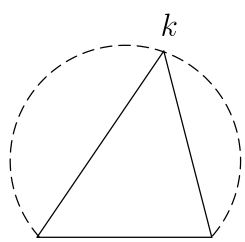


Figure 1

漸化式を使えば、

$$E_3 = 1, E_4 = 2, E_5 = 5, E_6 = 14, E_7 = 42, \ldots$$

であることがわかり、上記オイラーの公式の妥当性が確認できる。

漸化式から公式を導くために、**母関数** (generating function) の方法を使う。数列 $\{E_n\}_{n\geq 2}$ に対して、z の関数 $F(z)$ を次の式で定義する (数列 $\{E_n\}$ の母関数)。

$$F(z) = \sum_{n=2}^{\infty} E_n z^{n-1}.$$

形式的な計算により、上の漸化式から

$$F(z)^2 = F(z) - z$$

なる関係式を導くことができる。この二次方程式を $F(0) = 0$ に注意して解くと、

$$F(z) = \frac{1 - \sqrt{1 - 4z}}{2}$$

という表式にたどりつく。

二項定理

$$(1 + t)^l = 1 + \binom{l}{1} t + \cdots + \binom{l}{l} t^l$$

は、l が自然数でなくても、

$$\binom{l}{k} = \frac{l(l-1)\dots(l-k+1)}{k!}$$

と解釈すれば、

$$(1 + t)^l = \sum_{k=0}^{\infty} \binom{l}{k} t^k, \qquad |t| < 1$$

の形で成立するので、

$$\binom{1/2}{k} = (-1)^{k-1} \frac{(2k-2)!}{2^{2k-1} k!(k-1)!}, \qquad k \geq 1$$

を代入して

$$\sqrt{1 - 4z} = 1 - 2 \sum_{k=1}^{\infty} \frac{(2k-2)!}{k!(k-1)!} z^k$$

である。これから、

$$F(z) = \sum_{k=1}^{\infty} \frac{(2k-2)!}{k!(k-1)!} z^k$$

となり、$F(z)$ の定義式と比較すると、

$$E_n = \frac{(2n-4)!}{(n-1)!(n-2)!} = C_{n-2}$$

であることがわかる。

以上のいささか形式的な計算の正当性については、後程少し補う。

3. Catalan's Bracketing Problem

一方、E.C. Catalan は次のような問題から、カタラン数に到達した。

「n 個の量 $a_1, a_2, \ldots, a_n$ の積 $a_1a_2\ldots a_n$ の計算を 2 つの積の計算の繰り返しで行うとして、何通りの計算方法があるか。」

言い換えると、計算の順序を指定するための括弧のつけ方が何通りあるか？

$n = 2, 3, 4$ については

$$a_1a_2$$

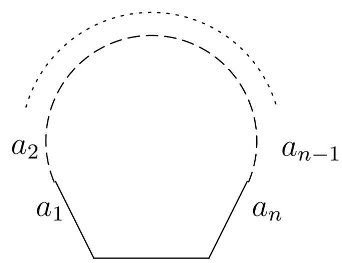
$$(a_1a_2)a_3, \qquad a_1(a_2a_3)$$

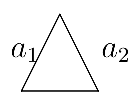
$$((a_1a_2)a_3)a_4, \quad (a_1a_2)(a_3a_4), \quad a_1(a_2(a_3a_4)), \quad a_1((a_2a_3)a_4), \quad (a_1(a_2a_3))a_4$$

である。これから推測されるように、問題の答は、C_{n-1} である。

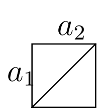
4. オイラー・カタラン対応

与えられた n 個の量 $a_1, a_2, \ldots, a_n$ を図のように凸 $n+1$ 角形の辺に順次貼り付けると、凸多角形の三角形分割と括弧のつけ方が一対一に対応することがわかる。

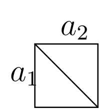




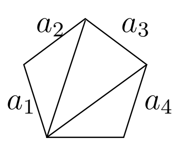
$= a_1a_2$



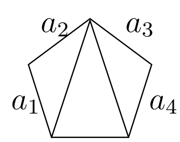
$= (a_1a_2)a_3,$



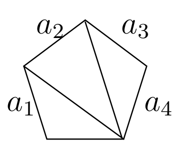
$= a_1(a_2a_3),$



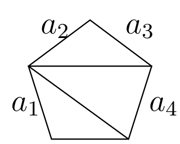
$= ((a_1a_2)a_3)a_4,$



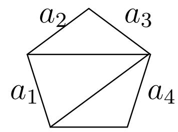
$= (a_1a_2)(a_3a_4),$



$= a_1(a_2(a_3a_4)),$



$= a_1((a_2a_3)a_4),$

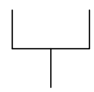


$$= (a_1(a_2a_3))a_4.$$

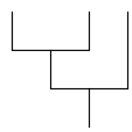
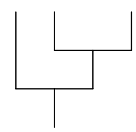
5. 二分木

カタランの括弧づけの問題は、また、つぎのように**二分木** (binary tree) による対応づけも可能である。

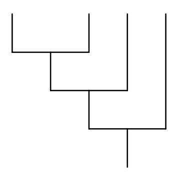
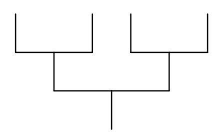
1-node

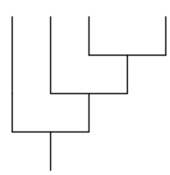
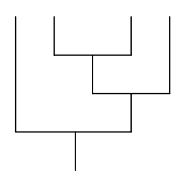
a_1a_2 

2-node

$(a_1a_2)a_3$  $a_1(a_2a_3)$ 

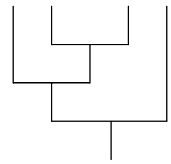
3-node

$((a_1a_2)a_3)a_4$  $(a_1a_2)(a_3a_4)$ 

$a_1(a_2(a_3a_4))$  $a_1((a_2a_3)a_4)$ 

6. Lattice Paths

与えられた binary tree に対して、次の規則で枝に番号 $1, 2, \ldots, 2n$ をつける。(枝の総数は $2n$ であることに注意。) 根から出発して左側の枝に順次番号をつける。枝の末端に達したならば、もとに戻って、ま

$(a_1(a_2a_3))a_4$ 

だ番号のついていない枝が現れたらその枝に進み、以下、最初の規則を繰り返す。

さて、このように番号付けられた枝に対して、k 番目の枝が左右どちらに分岐しているかで、l, r のラベルを付ける。上のふたつの例だと、

$$\begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ l & l & r & l & r & r \end{array} \qquad \begin{array}{cccccc} 1 & 2 & 3 & 4 & 5 & 6 \\ l & l & r & r & l & r \end{array}$$

ラベルは、左優先でつけているので、k 番目以下の r の総数は l の総数を越えない。例えば、

$$l \quad r \quad r \quad l \quad r \quad l$$

という配列は許されない。

このような $2n$ 個の l または r の列に対して、$(0,0)$ から (n,n) への **lattice path** （最短折線）を規則「l は x 座標を $+1$ 進め、r は y 座標を $+1$ 進める」で対応させると、

「$(0,0)$ から (n,n) への lattice path で、半平面 $y \le x$ に含まれるもの」

が得られる。逆にこのような lattice path は、一つの n-node binary tree に対応する。

ここで、Figure 2 における lattice path の総数は、$\binom{m+n}{n}$ であることに注意しておく。

さて、$y \le x$ に含まれる $(0,0)$ から (n,n) への lattice path 全体は、x 軸方向に $+1$ ずらすことにより、$(1,0)$ から $(n+1,n)$ への lattice path で、直線 $y = x$ を通らないもの全体に一致する。このような path の総数を求めるために、その補集合 P を考える。すなわち、$(1,0)$ から $(n+1,n)$ への lattice path で直線 $y = x$ と交点をもつもの全体が P である。さて、P に含まれる path p に対して、直線との最初の交点を

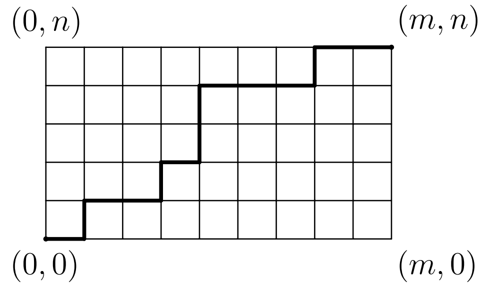


Figure 2

(i,i) として、$(1,0)$ から (i,i) への部分を $y=x$ に関して折り返して得られる $(0,1)$ から $(n+1,n)$ への lattice path を p' で表す (Figure 3)。この対応で、集合 P は、$(1,0)$ から $(n+1,n)$ への free lattice path 全体に写される。従って、P の個数は、$\binom{2n}{n+1}$ であり、binary n-branching tree の総数は、

$$\binom{2n}{n} - \binom{2n}{n+1} = C_n$$

で与えられる。

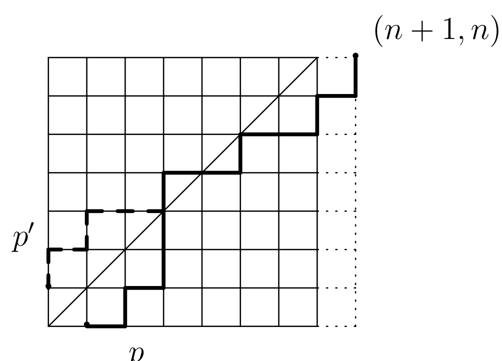
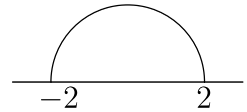


Figure 3

7. 半円分布

確率分布

$$\frac{1}{2\pi}\int_{-2}^{2}\sqrt{4-x^2}dx = 1$$



を**半円分布** (semicircle distribution) とよぶ。半円分布の**積率** (moment) は、積分

$$\mu_n = \frac{1}{2\pi}\int_{-2}^{2} x^n\sqrt{4-x^2}dx, \qquad n = 0, 1, 2, \ldots$$

で与えられる。n が奇数のとき、被積分関数は奇関数なので、$\mu_n = 0$ である。

さて、μ_n の値は、漸化式を使って求められる。

$$(x^n(4-x^2)^{3/2})' = 4nx^{n-1}\sqrt{4-x^2} - (n+3)x^{n+1}\sqrt{4-x^2}$$

を積分して、

$$\mu_{2k} = 2\frac{2k-1}{k+1}\mu_{2k-2}, \qquad k = 1, 2, \ldots.$$

これから、μ_{2k} は帰納的に

$$\begin{aligned}
\mu_2 &= 2\frac{2-1}{2}\mu_0 = 1,\\
\mu_4 &= 2\frac{4-1}{3}\mu_2 = 2,\\
\mu_6 &= 2\frac{6-1}{4}\mu_4 = 5,\\
\mu_8 &= 2\frac{8-1}{5}\mu_6 = 14,\\
&\vdots
\end{aligned}$$

と求められ、公式

$$\mu_{2n} = C_n$$

が予想される。実際これが正しいことは、

$$\frac{C_n}{C_{n-1}} = 2\frac{2n-1}{n+1}$$

より、C_n と μ_{2n} が同じ漸化式をみたすことからわかる。

8. 漸近挙動

上でも使った比の式から

$$\frac{C_n}{C_{n-1}} = 4\frac{n-1/2}{n+1} \le 4 \qquad (n \ge 1)$$

がわかるので、カタラン数についての評価式

$$C_n \le 4^n$$

が得られる。これから、

$$\left|\sum_{k=0}^{\infty} C_k z^k\right| \leq \sum_{k=0}^{\infty} C_k |z|^k \leq \sum_{k=0}^{\infty} (4|z|)^k = \frac{1}{1-4|z|}$$

は、$|z| < 1/4$ のとき収束して良い性質（絶対収束）をもつことがわかる。とくに、

$$\left(\sum_{k=0}^{\infty} C_k z^k\right)^2 = C_0^2 + (C_0C_1 + C_1C_0)z + (C_0C_2 + C_1^2 + C_2C_0)z^2 + \dots$$

が示せるので、これと二項展開式

$$\sqrt{1-4z} = 1 - 2\sum_{k=1}^{\infty} C_{k-1} z^k, \qquad |z| < \frac{1}{4}$$

とを組み合わせると、$\{C_{n-2}\}_{n\geq 2}$ が $\{E_n\}_{n\geq 2}$ と同じ漸化式をみたし、$C_0 = 1 = E_2$ であることから、$E_n = C_{n-2}$ という結論が厳密に得られる。

さて、$\lim_{n\to\infty} C_n/C_{n-1} = 4$ であるから、カタラン数の増大度のスピードの目安として $C_n \sim 4^n$ を得る。これをもっと精密化させると

$$C_n \approx \frac{1}{\sqrt{\pi}} \frac{4^n}{n^{3/2}}$$

となる。ここで、記号 $\approx$ は、2 つの量の比が $n \to \infty$ のとき 1 に近づく、ということを意味する。

この結果を導くために、階乗の増大度のスピードを記述する**スターリングの公式** (J. Stirling, 1692–1770)

$$n! \approx \sqrt{2\pi} n^n \sqrt{n} e^{-n}$$

を援用する。これを使えば、

$$C_n = \frac{(2n)!}{n!(n+1)!}$$

の増大度のスピードは、

$$\frac{\sqrt{2\pi}(2n)^{2n}\sqrt{2n}e^{-2n}}{(\sqrt{2\pi}n^n\sqrt{n}e^{-n})^2 n} = \frac{1}{\sqrt{\pi}} \frac{4^n}{n^{3/2}}$$

のそれに等しいことがわかる。

9. Random Walks

原点を出発点として直線上の整数点 $\{0, \pm 1, \pm 2, \dots\}$ を確率 p で右に 1、確率 $q = 1 - p$ で左に 1 移動する確率現象 (**random walk**) を考える。n ステップ後に $x = k$ という点の上にいる確率 p_k は、

$$p_k = \begin{cases} \binom{n}{(n-k)/2} p^{(n+k)/2} q^{(n-k)/2} & |k| \leq n \text{ で } n-k \text{ が偶数のとき、} \\ 0 & \text{それ以外のとき} \end{cases}$$

である。とくに

$$p_0 = \binom{n}{n/2} (pq)^{n/2}$$

は n ステップ後に原点に戻っている確率を表す。$2m$ ステップで初めて原点に戻る確率を r_m で表せば、p_0 はさらに

$$p_0 = \sum_{m=1}^{n/2} r_m$$

と分解される。

r_m を具体的に求めてみよう。n ステップの random walk の履歴を表すのに長さ n の lattice path を用いれば、$2m$ ステップ目ではじめて原点にもどる lattice path は、$(0, 0)$ と (m, m) 以外の点で直線 $y = x$ と交わらないものに他ならない。したがって、その総数は、$2C_{m-1}$. 一方、random walk が与えられた lattice path をたどる確率は $p^m q^m$ であるから、

$$r_m = 2C_{m-1}(pq)^m$$

となる。

これを使えば、いつかは原点に戻る確率 (**再帰確率**–recurrence probability) r は、

$$r = \sum_{m=1}^{\infty} r_m = 2 \sum_{m=1}^{\infty} C_{m-1}(pq)^m = 1 - \sqrt{1 - 4pq}$$

である。とくに、$p = q = 1/2$ のときには、$r = 1$ となって、絶対いつかは原点に戻ってくることを意味する。

ところが、原点に戻るまでに平均何ステップ必要かというと、

$$\sum_{m=1} 2m r_m = 4 \sum_{m=1}^{\infty} m C_{m-1} \frac{1}{4^m} = +\infty$$

となって、永遠に待たなければならない。

長い人生において、運というものは、$+$ にも $-$ にも等しい頻度で我々に降り懸る。その意味で、良いことも悪いことも長い目でみれば、

同じ数だけ起るのだと言えよう ($r=1$)。ところが、そのための「長い目」というのは、永遠の時を要する。
　有限の人生において、これは、運の偏在を意味しないだろうか。神や仏の力を借りずとも。

10. さらに勉強したい人のために

寺元 英 「ランダムな現象の数学」、吉岡書店
斎藤 正彦 「微分積分教科書」、東京図書
草場 公邦 「線型代数」、朝倉書店
セール 「有限群の線型表現」、岩波書店
ファインマン 「光と物質のふしぎな理論」、岩波書店