数学科

テングサグラフでの原点回帰の確率

福谷　敏

【概要】　問題の内容は、次の通りである：閉路のない無限無向単純グラフ（これは、難しく言うと、格子点の普遍被覆）で、各頂点からn(一定数）本の辺が出ているグラフ上で、頂点からrandom walkを考え、原点に戻ってくる確率を計算する。実験と予想の後、証明が得られた確率は$1/(n-1)$である。
【キーワード】　実験数学教材　テングサグラフ　ランダムウォーク　非可換代数　回帰確率　組み合わせ　カタラン数　母関数

1．はじめに

1993年秋に数学コンクールの審査を担当し、解説を考えていた。このときの副産物として見つけた小さな定理を紹介したい。楽しい実験数学教材を経験できた。ここに述べる定理は教員の世界では良く知られた教材ではないかと想像していた。しかし、そう自明な事実ではなさそうなので、大分時間がたったが、敢えて記事にした。

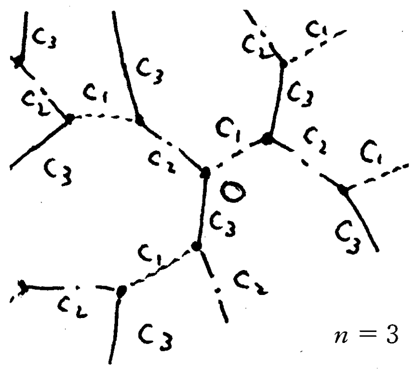
オリジナルの時点では、UBASICという言語で、数え上げ実験、シミュレーションと数式処理ソフトDERIVEで無限級数の数値予測などを行った。改めて記事にするにあたって、数え上げ実験を「非可換代数上の母関数」の考え方を変えて挑戦してみた。すなわち、「非可換版の二項展開公式」で求めてみる。数式処理ソフトMATHEMATICAの２文の入力で定義ですむ。他にも応用がありそうな代数的方法と思う。

定理の証明はアイデアは組み合わせ数の母関数を利用する。

2．定理（問題とその結論）

定義１　閉路（輪になった部分）のない無限無向単純グラフ（無限木）で、各頂点からn本（一定数）辺が出ているグラフをn**テングサグラフ**と呼ぶことにする。また、１つの頂点を固定して原点と呼ぶことにする。

$A^0(2l)$で原点を出発して、$2l$回辺を順に移動して（$2l$歩進むということにする）、はじめて原点に戻ってくるような経路の総数を表す。



$n=3$

定義２　$p(n)$はnテングサグラフ上の各頂点でn個の辺のいずれをとるか**同様に確からしい**場合の、原点を出発して原点に帰ってくる確率を表すことにする。

命題

$$A^0(2l)=n(n-1)^{l-1}\left(\frac{1}{l}\right){}_{(2l-2)}C_{l-1}$$

定理

$$p(n)=\frac{1}{n-1}$$

3．有名な類似の事実

フェラーの有名な確率論の教科書「確率論とその応用」の第14章第７節には、定理として「１次元および２次元の対称な（隣への移動が等確率のこと）格子上のランダム・ウォークでは、粒子は遅かれ早かれ初めの位置に戻る確率は１である。しかし、３次元の場合は、その確率は0.35(復帰の期待数は0.53にすぎない)」と述べている。スターリングの公式を利用した証明が述べられており、高校数学では手に負えない。

もっと簡単なモデルを考えて、このような事実を高校生程度の知識（組み合わせ数の母関数だけは無限の

処理に使う）で明晰に理解することを試みた。ここでは、「普遍被覆」といわれるもので代用すると、N＝1次元の格子は、$n=2$のテングサグラフと全く同じものである。N（$N \geqq 2$）次元の格子の場合は、（その普遍被覆グラフである）$n=2$Nのテングサグラフを代わりに考える。我々の定理は、1次元では「1次元格子のランダム・ウォークの原点回帰の確率は1」そのものである。平面では、「原点に一方の端を固定した無限に伸びるゴム紐をつけた粒子が平面の格子点上をランダム・ウォークし、ゴム紐の長さがが0に戻ることが起こる確率は1／3」という事実が我々の定理の1つの解釈である。母関数を用いるほかには代数的な計算で証明できる。初めは、UBASICというBASIC言語で書いたプログラムで実験した。すべての場合を数え上げる生成判定の方法とランダム関数を用いてシュミレーション的方法で実験したところ、余りに綺麗な予想値が見つかった。さらに、証明を目指した式を見ると、式の数値計算でも予想値は整数分の1が予想されたので、簡単な証明があるのではと確信し証明を試みたら、比較的簡単に導くことができた。

今回この記事では、実験自体を工夫し、大げさに言うと非可換代数の計算を利用する。実際は、素直な母関数の方法なので、数式処理ソフトMathematicaで、「非可換版の二項定理のアルゴリズム」として、2つの文で表現される。下に、$n=3$の場合の入力例In［1］、In［2］を載せる。各頂点からでた3本の辺を3色C1、C2、C3と順に色づけしておく。

```
In[1]
  ans[1]:=ans[1]=s**(c1+c2+c3)
  //.{zz_**(zz_+yy_)->zz**xx+zz**yy}

In[2]
  ans[n_]:ans[n]=ans[n-1]**(c1+c2+c3)
  //.{s**id->goal,(xx_+yy_)**zz_->xx**zz+yy**zz
  zz_**(xx_+yy_)->zz**xx+zz**yy,xx_**id->xx
  c1**c1->id, c2**c2->id, c3**c3->id,
  (k_Integer*yy_)**xx_->k yy**xx,
  (xx_**yy_)**zz_->xx**yy**zz, goal**xx_->goal
```

$n=3$のテングサグラフで、粒子が4回移動するのは3^4通り可能である。このうち原点に戻る回数は、Out［4］における「goal」の係数「33」また、長さが2に縮むの例として、ゴムの色C1の辺から色C2の辺に伸びたものなるのは、c1、c2の係数4、つまり4通りあることを示している（コンピュータ画面では、式の各項はもう少し複雑な形になるが簡単に直した）。

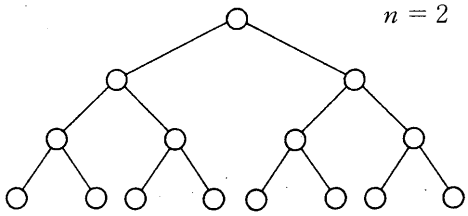
```
In[3]  ans[4]
  Out[3]  33  goal+4  c1*c2+4 c1*c3+4 c2*c1
  +4  c2*c3+4  c3*c1+4  c3*c2
  +c1*c2*c1*c2+c1*c2*c1*c3+c1*c2*c3*c1
  +c1*c2*c3*c2+c1*c3*c1*c2+c1*c3*c1*c3
  +c1*c3*c2*c1+c1*c3*c2*c3+c2*c1*c2*c1
  +c2*c1*c2*c3+c2*c1*c3*c1+c2*c1*c3*c2
  +c2*c3*c1*c2+c2*c3*c1*c3+c2*c3*c2*c1
  +c2*c3*c2*c3+c3*c1*c2*c1+c3*c1*c2*c3
  +c3*c1*c3*c1+c3*c1*c3*c2+c3*c2*c1*c2
  +c3*c2*c1*c3+c3*c2*c3*c1+c3*c2*c3*c2
```

確率は　$33/3^4$でおよそ0.41となる。スペースの関係で載せられないが、粒子が10歩移動する場合、goalの係数は27753となる。確率は$27753/3^{10}$でおよそ0.47となる。かなり予想値1／2に近い。また、この計算は非可換の文字式計算を利用している（この代数の積は結合的ですらない）。この非可換代数を利用した母関数的方法についてはそれ自体魅力的な考え方ではないかと思う。ただし、数え上げそのものを考えると、（すべての場合を数え上げる）生成判定やバックトラックの方が効率はよい。また、感じをつかむには、ランダム関数を用いてシミュレーションをするのも証明をつける意欲を強くさせるので有力であるが、紙面をとるのでプログラムは省略する。

4．定理の証明

定義3　nテングサグラフの原点から1つの辺を選んで、その辺とその先を消去したグラフを$n-1$分木と呼ぶことにする（$n=3$のときなら、いわゆる2分木になる）。このとき、原点の位置の点をこの木の端点と呼ぶ。

$n=2$



$n-1$分木で、原点から$2l$歩ではじめて元に戻る場合の数をB（$2l$）と表す。

補題1

$$B(2l)=\frac{(n-1)^l}{l+1}\,{}_{2l}C_l$$

証明

$n-1$分木の原点を1番上に、次に1つ離れた頂点を1段下に並べて、さらに2つ離れた点を2段下にというように置くことにする。まず、道筋を頂点のある

段の移動で分類する。下へ l 歩上に l 歩進みかつ原点より上には行かないので、ちょうどカタラン数の数だけ、段の移動があり。また、このそれぞれの場合に、下の段に１歩進む方法の数は１通りしかない。しかも、上に１下に１進むのであるから、段の移動回数に $(n-1)^{l}$ を掛けたものが求める場合の数となる。∎

補題２

$$A^{0}(2l)=nB(2l-2)$$

証明　n テングサグラフは原点から出た n 本の辺の先に n 個の $n-1$ 分木の原点がついたものと考えられる。ここで、$2l$ 歩目にはじめて原点に戻ってくるのは、最初と最後の１歩以外の $2l-2$ 歩は $n-1$ 分木にある場合である。∎

以上２つの補題をまとめると次の式が導ける。

命題

$$A^{0}(2l)=n(n-1)^{l-1}\frac{1}{l}\,{}_{(2l-2)}C_{l-1}$$

次に、原点に戻る確率について考える。

補題３

$$p(n)=\sum_{l=1}^{\infty}\frac{A^{0}(2l)}{n^{2l}}$$

証明　$2l$ 歩進むとき、各１歩ごと n 通りの選択があり同様に確からしいことから、各項の分母については正しい。また、原点にかえってくるのは、何歩目であるかを利用し、排反事象に分類して、その総和を考えると導かれる。

具体的な計算式は次の通りである。

命題

$$p(n)=\sum_{l=1}^{\infty}n(n-1)^{l-1}\,{}_{(2l-2)}C_{l-1}\frac{1}{n^{2l}}$$

証明　上の補題１から補題３までをまとめただけである。

有名な母関数の関係式を復習する。

補題４　$f(z)=\sum_{l=1}^{\infty}\frac{1}{l+1}\,{}_{2l}C_{l}z^{l}$　と置くと、$f(z)=\frac{1-\sqrt{1-4z}}{2z}$ となる。

証明　z の係数を比較して

$$f(z)=1+zf(z)^{2}$$

を満たしていることが確かめられ、根の公式でとく。根号の前の符合は最初の２項で定まる。∎

定理

$$p(n)=\frac{1}{n-1}$$

証明

$$\begin{aligned}
p(n)&=\sum_{l=1}^{\infty}\frac{A^{0}(2l)}{n^{2l}}\\
&=\frac{n}{n-1}\sum_{l=0}^{\infty}{}_{(2l-2)}C_{(l-1)}\frac{1}{l}\left(\frac{n-1}{n^{2}}\right)^{l}\\
&=\frac{n}{n-1}\quad{}_{(2l)}C_{l}\frac{1}{l+1}\left(\frac{n-1}{n^{2}}\right)^{l+1}\\
&=\frac{1}{n}\sum_{l=0}^{\infty}{}_{(2l)}C_{l}\frac{1}{l+1}\left(\frac{n-1}{n^{2}}\right)^{l}\\
&=\frac{1}{n}f\left(\frac{n-1}{n^{2}}\right)\\
&=\frac{1}{n}\,\frac{n-\sqrt{n^{2}-4n+4}}{2\frac{n-1}{n}}\\
&=\frac{1}{n}\,\frac{n-(n-2)}{2\frac{n-1}{n}}\\
&=\frac{1}{n-1}
\end{aligned}$$

∎

5．おわりに

生物界や物理に対応する現象がありそうな気がする。見つかると、教科横断的なオリジナル教材にできるかもしれない。また時間があったら、原点回帰だけでない一般的な状況でグラフ上のランダムウォークを調べてみたい。深みと考える楽しみが生徒と共有できる題材に出会えて幸せな1993年のコンクールであった。