2016年　計量国語学会　予稿

数理漢文学への道(2)

－線形と非線形のはざまで－

島野達雄

１．漢文数列入門

中国語の「百聞不如一見」にレ点と一二点を加えた「百聞不ㇾ如㆓一見㆒」は，「先に読む部分を＜　＞でくくる」というルールのもとで，「百聞不＜如＜一見＞＞」となる．

これを「百聞不ㇾ如㆓一見㆒」の**括弧表示**とよぶ．括弧表示では，上の例のほか，上下点のなかの一二点，三点のある一二点やレ点の連続などでも，括弧が**入れ子**になる．

括弧表示は，「百聞不如一見」から「百聞一見如不」へ，漢字の並び替えを指示しているともいえる．括弧をはずして漢字を並び替えることを**括弧表示の展開**とよぶ．

ここでは，n字の漢字列に対して，「返り点のつけ方」，言いかえれば「括弧表示のパターン」はいくつあるか，を考察する．なお，いわゆる異点同順のばあいは，同じ括弧表示としてあつかう．

まず，漢字0字のばあいを想定し，パターンは1通り，と約束する．

漢字1字はそのまま読むほかないので，パターンは1通りしかない．漢字2字では，レ点の有無により，「少年」と「有ㇾ朋」のように2通りのケースが考えられる．漢字3字では，返り点をつけない「我独醒」，レ点を使った「可ㇾ妻也」，「朝聞ㇾ道」，「不ㇾ踰ㇾ矩」，一二点を使った「楽㆓其俗㆒」のように返り点のつけ方のパターンが5通りある．漢字4字では14通り．5字では42通り，6字では132通り，7字では429通りある．

こうして$\{1,1,2,5,14,42,132,429,\cdots\}$という数列が生まれる．この数列は**カタラン数**とよばれ，**組合せ論**（Combinatorics）ではフィボナッチ数列についで（？）よく知られている．

カタラン数の一般項（このばあいは漢字n字に対する返り点のつけ方，つまり括弧表示のパターンの数）をc_nであらわす．数列としてのカタラン数は$\{c_n\}$と書く．

「俳㆓－笑　之㆒」のような竪点（たててん，ハイフン）の使用を認めると，カタラン数より多い新たな数列が生じる．竪点2個まで，竪点3個まで，…竪点を最大限まで使用することを認める数列も考えられる．竪点の括弧表示は，「俳－笑＜之＞」のようにする．

「再読」という現象を加味するときも，同じように竪点の数に応じて，数列が生まれる．再読の括弧表示は，たとえば，過猶ㇾ不ㇾ及＝過猶－＜不＜及＞＞＝過猶及不猶のように，－＜　＞で表示する．

以上の数列たちを総称して**漢文数列**とよぶ．

括弧表示のパターンと展開したあとの並び替え（語順変換）は 1 **対** 1 **に対応**することが数学的に証明できる．言いかえれば，括弧表示のパターン数と並び替えの数は一致する．

２．カタラン数の性質

カタラン数$\{c_n\}$は，竪点を使わず（竪点の使用 0 個），返り点だけの使用を認め，再読を認めない，もっとも単純な条件のもとで生じる．

漢字 0 字および 1 字では，パターン数は 1 とする．つまり，$c_0 = c_1 = 1$とする．

漢字 2 字以上のばあい，先頭の 2 つの漢字 A，B の関係は，①AB のように連続しているか，②A＜B のように＜があるか，の 2 通りのタイプしかない．

①のとき，B 以下（B の右側）に$\mathrm{n}-1$字の括弧表示がある．形式的に，

A ＋〔n－1 字の括弧表示〕

と書け，先頭の A はパターン数には無関係で，この場合のパターン数はc_{n-1}になる．

②のとき，A＜B…＞C… のかたちをしている．

A ＜〔i 字の括弧表示〕＞〔n－1－i 字の括弧表示〕

括弧のなかの i は 1 以上，n－1 以下の可能性があるので，総数は$\mathrm{i}=1$から$\mathrm{i}=\mathrm{n}-1$までの，c_iとc_{n-1-i}の積の総和になる．つまり，$c_1c_{n-2}+c_2c_{n-3}+\cdots+c_{n-2}c_1+c_{n-1}c_0$となる．

よって①のc_{n-1}をc_0c_{n-1}とし，②とあわせて，次の式をえる．

[2.1]　$c_n = c_0c_{n-1}+c_1c_{n-2}+\cdots+c_{n-2}c_1+c_{n-1}c_0$

この$c_0c_{n-1}+c_1c_{n-2}+\cdots+c_{n-2}c_1+c_{n-1}c_0$を数列$\{c_n\}$の**畳み込み**（convolution）とよぶ．

ここで，x^nの係数がc_nである無限級数$C(x)=c_0+c_1x+c_2x^2+\cdots+c_nx^n+\cdots$を考える．

数列$\{c_n\}$を順に書き出し，両辺に$x^n\ (n=0,1,2,\cdots)$をかけると，下のようになる．

$$
\begin{aligned}
&c_0 = 1\\
&c_1x = c_0c_0x\\
&c_2x^2 = (c_0c_1+c_1c_0)x^2\\
&c_3x^3 = (c_0c_2+c_1c_1+c_2c_0)x^3\\
&\ \vdots\\
&c_nx^n = (c_0c_{n-1}+c_1c_{n-2}+\cdots+c_{n-2}c_1+c_{n-1}c_0)x^n\\
&\ \vdots
\end{aligned}
$$

この左辺たちの和は，まさしく$C(x)$である．右辺たちの和は，

$$
\begin{aligned}
C(x)^2 &= (c_0+c_1x+c_2x^2+\cdots+c_nx^n+\cdots)(c_0+c_1x+c_2x^2+\cdots+c_nx^n+\cdots)\\
&= c_0c_0+(c_0c_1+c_1c_0)x+(c_0c_2+c_1c_1+c_2c_0)x^2+(c_0c_3+c_1c_2+c_2c_1+c_3c_0)x^3+\cdots
\end{aligned}
$$

に注意すると，$C(x)^2$にxをかけ，1 を加えたものになっている．$C(x)$をCと略記して，

[2.2]　$xC^2-C+1=0$

をえる．$C(x)$を数列$\{c_n\}$の**母関数**とよび，[2.2]をその**関数方程式**とよぶ．

[2.2]をxについて微分すると，

[2.3]　$C^2+2xCC'-C'=0$

これにxをかけ，[2.2]の$xC^2=C-1$を代入すると，

[2.4]　$2x^2CC'+C-xC'-1=0$

さらに[2.4]にxCをかけ，[2.2]の$xC^2=C-1$を代入すると，

[2.5]　$x^2CC'+(1-x)C-2x^2C'-1=0$

[2.4]と[2.5]からCC'を消去すると，次の 1 **階線形微分方程式**をえる．

[2.6]　$(1-2x)C+x(1-4x)C'-1=0$

ここで，$C=\sum_{n=0}^{\infty}c_nx^n$，$C'=\sum_{n=0}^{\infty}nc_nx^{n-1}$であることから$(1-2x)C$の$x^n$の係数は$c_n-2c_{n-1}$．$x(1-4x)C'$の$x^n$の係数は$\mathrm{n}c_n-4(n-1)c_{n-1}$．これらを足し合わせると 0 になるはずなので，

[2.7]　$(n+1)c_n=2(2n-1)c_{n-1}$

という 1 **階線形差分方程式**をえる．

次に，[2.7]の両辺を$(n+1)c_{n-1}$で割り，$n\to\infty$とすると，数列$\{c_n\}$の**項比の極限** rは，

[2.8]　$$r=\lim_{n\to\infty}\frac{c_n}{c_{n-1}}=\lim_{n\to\infty}\frac{2(2n-1)}{n+1}=4$$

となる．

ダランベールの判定法により，rの逆数$1/4$は$C(x)$の**収束半径**になる．$x=1/4$のとき，[2.2]の関数 2 次方程式の判別式は 0 で，[2.2]は$(C-2)^2=0$，つまり$C=2$となる．すなわち

[2.9]　$$C\left(\frac{1}{4}\right)=1+\frac{1}{4}+2\cdot\left(\frac{1}{4}\right)^2+5\cdot\left(\frac{1}{4}\right)^3+14\cdot\left(\frac{1}{4}\right)^4+42\cdot\left(\frac{1}{4}\right)^5+132\cdot\left(\frac{1}{4}\right)^6+\cdots=2$$

３．漢文数列への拡張

ここからは，カタラン数の性質がすべての漢文数列で成り立つことを示す．

以下，数列$\{a_n\}$を$\{a(n)\}$のようにも書く．

返り点と竪点 m 個（$0\leq \mathrm{m}\leq \mathrm{n}-2$）までの使用を認め，再読をふくめない数列を$\{k^m(n)\}$，再読を加味した数列を$\{p^m(n)\}$とする．$\{k^0(n)\}$は竪点を使わない m＝0 のばあいで，つまりカタラン数$\{c_n\}$である．同様に$\{p^0(n)\}$を$\{d_n\}$と，$\{k^{n-2}(n)\}$を$\{k_n\}$と，$\{p^{n-2}(n)\}$を$\{p_n\}$と書く．

数列$\{a(n)\}$の畳み込み$a_0a_{n-1}+a_1a_{n-2}+\cdots+a_{n-1}a_0$を$a(0*n-1)$と書く．

たとえば竪点 1 個の使用を認める$\{k^1(n)\}$の先頭の 2 字は，①AB…，②A＜B… に加えて，③A－B＜… つまり

A－B ＜〔i 字の括弧表示〕＞〔n－2－i 字の括弧表示〕　（$1\leq \mathrm{i}\leq \mathrm{n}-2$）

を数え上げるので，

$k^1(\mathrm{n})=k^1(0*n-1)+k^1(0*n-2)-k^1(n-2)$

と定義するのが妥当であろう．

一般に$\{k^m(n)\}$をつぎのように**定義する**．（$k^m(0) = k^m(1) = 1$）

$$k^m(\mathrm{n}) = k^m(0*n-1) + k^m(0*n-2) - k^m(n-2) + k^m(0*n-3) - k^m(n-3) + \cdots + k^m(0*n-m-1) - k^m(n-m-1)$$

すると，$\{k_n\}$は$m = n-2$のときなので，

$$k_n = k(0*n-1) + k(0*n-2) - k_{n-2} + k(0*n-3) - k_{n-3} + \cdots + k(0*1) - k_1$$

で定義されるが，上式は

[3.1]　$k_n = k_{n-1} - k_{n-2} + k(0*n-1)$

と簡略化できる．

[3.1]からは，カタラン数の[2.2]とほぼ同じ手順で母関数$K(x)$の関数方程式

[3.2]　$xK^2 - (1-x+x^2)K + 1 - x = 0$

を求めることができる．

同様に再読を認める$\{p^m(n)\}$では，④A－＜B…　つまり②A＜B…　と同数が増えるので，

$$p^m(\mathrm{n}) = 2p^m(0*n-1) - p^m(n-1) + p^m(0*n-2) - p^m(n-2) + p^m(0*n-3) - p^m(n-3) + \cdots + p^m(0*n-m-1) - p^m(n-m-1)$$

で**定義する**．（$p^m(0) = p^m(1) = 1$）

$\{p_n\}$についても，

$$p_n = 2p(0*n-1) - p_{n-1} + p(0*n-2) - p_{n-2} + \cdots + p(0*1) - p_1$$

は，次のように簡略化できる．

[3.3]　$p_n = 2p(0*n-1) - p(0*n-2)$

母関数$P(x)$の関数方程式は，次のようになる．

[3.4]　$x(2-x)P^2 - P + 1 - x = 0$

$\{k^m(n)\}$の母関数の関数方程式は，

$\{c_n\}$	$xC^2 - C + 1 = 0$
$\{k^1(n)\}$	$x(1+x){K_1}^2 - (1+x^2)K_1 + 1 = 0$
$\{k^2(n)\}$	$x(1+x+x^2){K_2}^2 - (1+x^2+x^3)K_2 + 1 = 0$
$\{k^3(n)\}$	$x(1+x+x^2+x^3){K_3}^2 - (1+x^2+x^3+x^4)K_3 + 1 = 0$
$\{k^m(n)\}$	$x(1+x+x^2+\cdots+x^m){K_m}^2 - (1+x^2+x^3+\cdots+x^{m+1})K_m + 1 = 0$

ここで$\mathrm{m} \to \infty$としたときのK_mを$\widetilde{K}$とおくと，$|x| < 1$のとき

[3.5]　$\dfrac{x}{1-x}\widetilde{K}^2 - \left(\dfrac{1}{1-x} - x\right)\widetilde{K} + 1 = 0$

この式は[3.2]と一致している．すなわち$\{k^m(n)\}$の母関数の関数方程式の極限は，$\{k_n\}$の関数方程式に等しい．（D. G. Rogers, Mar 2012）

同様に，$\{p^m(n)\}$についても，

$\{p^m(n)\}$　$x(2+x+x^2+\cdots+x^m){P_m}^2-(1+x+x^2+\cdots+x^{m+1})P_m+1=0$
となり，$m\to\infty$としたとき，関数方程式の極限は$\{p_n\}$の[3.4]と一致する．

母関数の関数方程式を微分すると，カタラン数と同じ手順で，1 階線形微分方程式が求められる．たとえば$\{k_n\}$の母関数$K(x)$の関数方程式$xK^2-(1-x+x^2)K+1-x=0$からは，[2.4]と[2.5]に相当する 2 つの式がえられる．

[3.6]　$2x^2KK'-x(1-x+x^2)K'+(1-x^2)K-1=0$

[3.7]　$x^2(1-x+x^2)KK'-2x^2(1-x)K'+(1-2x+x^3-x^4)K-(1-x)(1-x^2)=0$

両式から，KK' を消去し，[2.6]に相当する微分方程式は次のようになる．

[3.8]　$(1-3x+x^3-x^4)K+x(1-6x+7x^2-2x^3+x^4)K'-1+x+3x^2-2x^3=0$

この式で，x^nの係数を調べると，

[3.9]　$(n+1)k_n=3(2n-1)k_{n-1}-7(n-2)k_{n-2}+(2n-7)k_{n-3}-(n-5)k_{n-4}$

という 4 階線形差分方程式をえる．

同様にして，$\{k^1(n)\}$は 5 階，$\{k^2(n)\}$は 8 階，$\{k^3(n)\}$は 12 階，再読を加味した$\{p^1(n)\}$は 5 階，$\{p^2(n)\}$は 8 階，$\{p^3(n)\}$は 11 階の線形差分方程式を満たす．下に$\{d_n\}$の 2 階，$\{p_n\}$の 4 階線形差分方程式を示す．

[3.10]　$(n+1)d_n=3(2n-1)d_{n-1}-(n-2)d_{n-2}$

[3.11]　$2(n+1)p_n=(17n-7)p_{n-1}-4(8n-13)p_{n-2}+4(5n-13)p_{n-3}-2(2n-7)p_{n-4}$

このようにして得た差分方程式からは，項比の極限および母関数の収束半径を求めることができる．$\{k_n\}$のばあい，[3.9]の両辺を$(n+1)k_{n-1}$で割ると，

[3.12]　$$\frac{k_n}{k_{n-1}}=\frac{1}{n+1}\left\{3(2n-1)-\frac{7(n-2)}{\frac{k_{n-1}}{k_{n-2}}}+\frac{2n-7}{\frac{k_{n-1}}{k_{n-2}}\cdot\frac{k_{n-2}}{k_{n-3}}}-\frac{n-5}{\frac{k_{n-1}}{k_{n-2}}\cdot\frac{k_{n-2}}{k_{n-3}}\cdot\frac{k_{n-3}}{k_{n-4}}}\right\}$$

よって，項比k_n/k_{n-1}の$n\to\infty$としたときの極限rは，

[3.13]　$r^4-6r^3+7r^2-2r+1=0$

の解になる．また，ダランベールの判定法により，rの逆数sが母関数の収束半径になる．

[3.13]から，sの方程式は次のようになる．

[3.14]　$(1-s+s^2)^2-4s(1-s)=0$

これは[3.2]の関数 2 次方程式の判別式が 0 で，[3.2]が重根をもつことを示している．

一般に$\{k^m(n)\}$のばあい，sの方程式は

[3.15]　$(1+s+\cdots+s^m)\{(1+s^2+\cdots+s^{m+1})^2-4s(1+s+\cdots+s^m)\}=0$

となり，$m\to\infty$とすると，$0<s<1$だから，当然のことながら[3.14]と一致する．

むろん$\{p^m(n)\}$についても，sの方程式の極限は$\{p_n\}$のsの方程式と一致する．

４．線形と非線形のはざまで

これまでの議論を一般化すると，

$$[4.1]\quad a_n = \sum_{i=1}^{k} P_i a_{n-i} + P_{k+1} \sum_{i=0}^{n-1} a_i a_{n-1-i} \qquad (P_1, P_2 \cdots \text{は定数})$$

タイプの，線形部分$\sum P_i a_{n-i}$と畳み込みの非線形部分$P_{k+1} \sum a_i a_{n-1-i}$からなる数列$\{a_n\}$は，母関数が 2 次の関数方程式を満たすことがわかる．このような数列を**畳み込み数列**とよぶ．

もし母関数 L が 2 次の関数方程式,

$$[4.2]\quad f(x)L^2 + g(x)L + h(x) = 0$$

を満たすならば，L'に関する次の微分方程式が成り立つ．（D. G. Rogers, Apr 2016）

$$[4.3]\quad (2fh' - 2ff'h - fgg' + f'g^2)L - (4f^2h - fg^2)L' - 2fg'h + fgh' + f'gh = 0$$

$f(x), g(x), h(x)$が多項式のばあい，[4.3]を書きかえ，

$$[4.4]\quad F(x)L + G(x)L' + H(x) = 0$$

とすると，$F(x), G(x), H(x)$も多項式となり，$F(x)$の次数が Nなら$G(x)$の次数は$N+1$になる．

このとき，これまでの議論で見たように，母関数 L の数列は N 階の差分方程式を満たす．つまり，Lが多項式を係数にもつ 2 次の関数方程式を満たすばあい，もとの数列は線形の差分方程式を満たす．すなわち，ある自然数 N が存在し，畳み込み数列$\{a_n\}$に関して，各Q_{n-j}とR_{n-j}を定数とする N 階線形差分方程式が成り立つ．

$$[4.5]\quad \sum_{j=0}^{N} (Q_{n-j} n + R_{n-j}) a_{n-j} = 0$$

この定理をつかうと，畳み込み数列の各項を計算するスピードが大幅にあがる．

また，項比の極限および母関数の収束半径がこの定理から容易に求められる．

漢文数列は，非線形の**入れ子構造**をもつ括弧表示から生じる．にもかかわらず，線形の差分方程式を満たす．漢文数列は，まさに**線形と非線形のはざま**に存在している．

今後は，数学および情報科学の課題として，漢文数列の一般項や近似値をもとめるほか，ヤング図形やランダウ関数に関連した,整数の分割やグラフ理論の問題などが残っている．

また，漢文を対象にしたモデル化をおこなう上で,「再読の運用」「漢字ごとの四声表記」「添え仮名」「品詞分類」など，より史料に即した概念の導入をはからねばならない．

現代の最先端の数学と漢文学への寄与は，今後の数理漢文学が目指すひとつの方向といえるだろう．