# 組み合わせ数え上げとランダム 写像の漸近解析 (公開用)

でぃぐ (@fujidig)

2019年10月27日

## 自己紹介

- 名前: でぃぐ
- 集合論を勉強している B4
- 今日話す内容は集合論とは全く関係がありません

- 2 ラベル付きクラスと指数型母関数
- 3 特異点解析と木関数の漸近展開
- 4 多変数の指数型母関数
- ⑤ 表の結果の証明 (一番短い周期の期待値を除く)
- 一番短い周期の期待値
- **⑦** まとめ

#### イントロダクション

- ② ラベル付きクラスと指数型母関数
- ③ 特異点解析と木関数の漸近展開
- 多変数の指数型母関数
- ⑤ 表の結果の証明 (一番短い周期の期待値を除く)
- 一番短い周期の期待値
- **の**まとめ

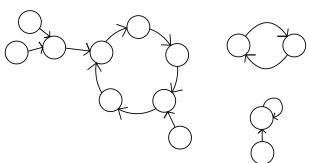
#### 話すこと

- $n := \{0, 1, ..., n-1\}$
- <u>n</u>から<u>n</u>への写像全体をMap(<u>n</u>, <u>n</u>)とおく
- $\mathsf{Map}(\underline{n},\underline{n})$  に一様分布を入れる. つまり、 $\mathsf{Map}(\underline{n},\underline{n})$  の要素数は $n^n$  なので、そのどれもが等確率で出るものとする
- n次の置換全体の集合 S<sub>n</sub> にも一様分布を入れる.これは n! 個が等確率
- なお、n次の置換の定義域はここでは $\underline{n}$ とする. よって $\mathfrak{S}_n \subseteq \mathsf{Map}(\underline{n},\underline{n})$

#### 話すこと

 $f: \underline{n} \to \underline{n}$  に対して,頂点集合を  $\underline{n}$  とし,f(i) = j なら頂点 i から頂点 j に有向辺を貼ったグラフを f の functional グラフと呼ぶことにする.

これは以下のようにサイクルの各頂点に木をくっつけたようなグラフを連結成分とするようなグラフになっている:



#### 話すこと

このとき,ランダムに元をとったときの次のような量の $n \to \infty$  での漸近的振る舞いが知られている:

量	Map( <u>n</u> , <u>n</u> )	$\mathfrak{S}_n$
長さ 1 のサイクルを持つ確率	$\rightarrow 1 - 1/e \doteq 0.63$	
連結成分の個数の期待値	$\sim \frac{1}{2} \log n$	= <i>H</i> <sub>n</sub>
ランダムな点をとったときそれが属する 連結成分が持つサイクルの長さの期待値	$\sim \sqrt{\pi n/8} \doteq 0.62\sqrt{n}$	=(n+1)/2
一番長いサイクルの長さの期待値	$\sim c_1 \sqrt{n} \doteq 0.78 \sqrt{n}$	$\sim \lambda n \doteq 0.62n$
一番短いサイクルの長さの期待値	$\sim \frac{e^{-\gamma}}{2} \log n \doteq 0.28 \log n$	$\sim e^{-\gamma} \log n \doteq 0.56 \log n$

表の左側の結果について, 証明を与えたい.

#### このテーマのモチベーション

- このテーマは疑似乱数の発生と関係してくる
- 一般的にコンピュータは(外部からノイズを入力 しない限り)乱数を発生させることはできない
- そこで確定的な計算によって乱数っぽいもの (疑 似乱数) を作っている
- 疑似乱数は固定された写像  $f \in \mathsf{Map}(\underline{n},\underline{n})$  を使った 反復列  $(x_n)_{n \in \mathbb{N}}$  すなわち
  - $x_0 = (適当な値), x_{n+1} = f(x_n)$  で定義される数列によって定められると考えることができる

#### このテーマのモチベーション

- 疑似乱数の生成に使われる f は「でたらめなアルゴリズム」を持ってくれば良いだろうという誤解がコンピュータの最初期にはあった
- しかし、「でたらめなアルゴリズム」は短い周期を 持ちやすいので、疑似乱数にふさわしくない

デモ.

#### このテーマのモチベーション

以上で見せたような「周期が非常に短くなりうる」という現象はどのくらい一般的なのだろうか?

→ 実はめっちゃ一般的!

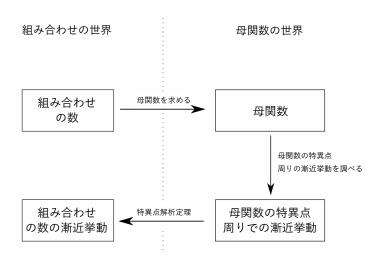
ということを証明する. 実際,「でたらめなアルゴリズム」を選ぶことが  $Map(\underline{n},\underline{n})$  に一様分布を入れて選ぶことでうまくモデル 化できているとすれば,表の左下の結果:  $\frac{e^{-\gamma}}{2}\log n$  を使 い,たとえば $n=2^{32}$  のときで,一番短い周期の期待値 は 6.2 程度, $n=2^{64}$  でも 12.4 程度であるとわかる

## 問題を解く流れ

これから表の左側の証明をしていくが, その流れは次 の通り

- 漸近挙動を得たい量に対応する母関数を求める
- ② その母関数の特異点周りの振る舞いを調べる
- ⑤ 「特異点解析定理」よりもとの量の漸近挙動がわ かる

## 問題を解く流れ



次節では母関数を求めるための道具を紹介しよう

- **1** イントロダクション
- 2 ラベル付きクラスと指数型母関数
- ③ 特異点解析と木関数の漸近展開
- 4 多変数の指数型母関数
- ⑤ 表の結果の証明 (一番短い周期の期待値を除く)
- 一番短い周期の期待値
- **の**まとめ

#### ラベル付きクラス

#### 定義(ラベル付きクラス)

- ・サイズnの弱ラベル付きオブジェクトとは頂点数nのグラフでその頂点集合が $\mathbb Z$ の部分集合であるもの。
- ・サイズnのよくラベル付けされたオブジェクトとは弱ラベル付きオブジェクトでかつ頂点集合が $\underline{n}$ であるもののこと
- 弱ラベル付きオブジェクト  $\alpha$  に対してそのサイズ  $\delta$   $|\alpha|$  と書く
- よくラベル付けされたオブジェクトからなる集合 をラベル付きクラスという

#### 指数型母関数

#### 定義 (指数型母関数)

ラベル付きクラス $\mathcal{A}$ に対して次で定義される形式的冪級数を $\mathcal{A}$ の指数型母関数という:

$$A(z) = \sum_{n>0} |\mathcal{A}_n| \frac{z^n}{n!}.$$

ここに $\mathcal{A}_n$ は $\mathcal{A}$ の要素でサイズがnのもの全体の集合. (これが各nに対して有限なものしか考えないことにする)

## ラベル付きクラスの同型

#### 同型

 $\mathcal{A}, \mathcal{B}$  をラベル付きクラスとする.

全単射  $f: \mathcal{A} \to \mathcal{B}$  でオブジェクトのサイズを保つものを同型写像といい,同型写像が存在するとき  $\mathcal{A}, \mathcal{B}$  は同型であるといい, $\mathcal{A} \simeq \mathcal{B}$  で表す.

二つのラベル付きクラスが同型ならば、それらに対応する指数型母関数は等しい.

## ラベル付きクラスの定義のまずさとそ れの解決策としての **species**

- ラベル付きクラスを定義したのはいいが、実はそこには問題がある.
- 今後, ラベル付きオブジェクトの順序対やラベル付きオブジェクトの同値類もラベル付きオブジェクトとみなす場面がある. しかしそれはラベル付きオブジェクトはグラフの一部だという定義に反している.
- ここでは詳しく取り上げないが、ラベル付きクラスを 「species」という概念で定義すれば問題は解消する. [10] を 参照.
- しかし本稿ではラベル付きオブジェクトはグラフの一部だ という定義を採用しつつも、拡大解釈に寛容でいる、という 立場をとりたい。

## ゼロオブジェクトと単位オブジェクト

#### 定義(ゼロオブジェクトと単位オブジェクト)

- 頂点集合が空集合のグラフがなすラベル付きオブジェクトを  $\epsilon$  と書く. 頂点集合が  $\underline{1}$  =  $\{0\}$  の辺がないグラフがなすラベル付きオブジェクトを  $\underline{0}$  と書く.
- ⑩のみからなるラベル付きクラスを Unit と書く.
- $\{\epsilon\}$ , Unit のそれぞれの指数型母関数は

$$E(z) = 1, U(z) = z$$

である.

## 例: 置換のラベル付きクラス

置換  $\sigma \in \mathcal{S}_n$  が

$$\sigma = \begin{pmatrix} 0 & 1 & 2 & \cdots & n-1 \\ \sigma_0 & \sigma_1 & \sigma_2 & \cdots & \sigma_{n-1} \end{pmatrix}$$

だったとき、それのグラフ表現とは



のこととする. 置換のグラフ表現全体の集合はラベル付きクラスをなし、それを PERM と書く. ただし空写像も置換とみなすことにする.

## 例: 置換のラベル付きクラス

PERM の指数型母関数は

$$P(z) = \sum_{n \ge 0} n! \frac{z^n}{n!} = \sum_{n \ge 0} z^n = \frac{1}{1 - z}$$

となる.

#### 例: urn

完全不連結なグラフを **urn** といい,よくラベル付けされた urn 全体のなすラベル付きクラスを URN で書く:

サイズnのよくラベル付けされたurnはそれぞれ1個なので、URNの指数型母関数は

$$U(z) = \sum_{n>0} 1 \cdot \frac{z^n}{n!} = \sum_{n>0} \frac{z^n}{n!} = \exp z.$$

#### 例: 円形グラフ

よくラベル付けされた円形グラフのなすラベル付きクラスを Cyc で表す:

$$\mathsf{Cyc} = \left\{ \textcircled{0}, \textcircled{0}, \textcircled{0}, \textcircled{0}, \textcircled{0}, \ldots \right\}$$

サイズnの円形グラフは円順列と1対1対応し、それは(n-1)! 個あるので、Cycの指数型母関数は

$$C(z) = \sum_{n \ge 1} (n-1)! \frac{z^n}{n!} = \sum_{n \ge 1} \frac{z^n}{n} = \log \frac{1}{1-z}.$$

#### ラベル付きクラスの間の積

これから、ラベル付きクラスの間に積の演算を定義する.これは、指数型母関数の積に対応することになる.

$$\mathcal{A} = \mathcal{B} \star C \iff A(z) = B(z)C(z)$$

## ラベル付きオブジェクトの削減

#### 定義(削減)

サイズnの弱ラベル付きオブジェクトに対してその<mark>削減</mark>とは,ラベルの相対的な順序は保ったまま,ラベルの集合をnにしたものをいう.



## ラベル付きオブジェクトの積

#### 定義(積)

よくラベル付けされたオブジェクト $\beta, \gamma$ に対して,

$$\beta \star \gamma = \{(\beta', \gamma') | (\beta', \gamma') \text{ はよくラベル付けされていて}$$
  $\beta \text{ は} \beta' \text{ の削減かつ} \gamma \text{ は} \gamma' \text{ の削減} \}$ 

とおく.

# ラベル付きオブジェクトの積

例: 
$$\beta = 0$$
 1,  $\gamma = 0$  1 に対して

$$\beta \star \gamma = \left\{ \begin{pmatrix} 0 & 1 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 2 & 4 \\ 3 & 1 & 2 & 4 \end{pmatrix}, \begin{pmatrix} 0 & 1 & 2 & 3 \\ 4 & 1 & 2 & 3 \end{pmatrix}, \begin{pmatrix} 0 & 2 & 1 & 3 \\ 4 & 3 & 1 & 2 \end{pmatrix}, \begin{pmatrix} 0 & 2 & 1 & 3 \\ 4 & 3 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 & 3 \\ 4 & 3 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 & 3 \\ 4 & 3 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 & 3 \\ 4 & 3 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 & 3 \\ 4 & 3 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 & 3 \\ 4 & 3 & 0 & 2 \end{pmatrix}, \begin{pmatrix} 1 & 2 & 0 & 3 \\ 4 & 3 & 0 & 2 \end{pmatrix}$$

要素数が  $\binom{5}{3}$  = 10 になっていることに注意しておく.

#### ラベル付きクラスの積

 $\mathcal{B}$ , C をラベル付きクラスとする. その積  $\mathcal{B} \star C$  を

$$\mathcal{B} \star C = \bigcup_{\beta \in \mathcal{B}, \gamma \in C} \beta \star \gamma$$

で定める.

## ラベル付きクラスの積と指数型母関数

#### 定理

 $\mathcal{B}, C$  をラベル付きクラスとする.その積を  $\mathcal{A} = \mathcal{B} \star C$  として, $\mathcal{A}, \mathcal{B}, C$  の指数型母関数をそれぞれ A, B, C とすれば

$$A(z) = B(z)C(z)$$

が成り立つ.

略証.  $\mathcal{A}$ ,  $\mathcal{B}$ ,  $\mathcal{C}$  の元のなかでサイズ n のものの個数をそれぞれ  $A_n$ ,  $B_n$ ,  $C_n$  と書くと

$$A_n = \sum_{k=0}^n \binom{n}{k} B_k C_{n-k}$$

が成り立つ. これは B(z)C(z) の  $z^n$  の係数に n! をかけたものそのものである.

## 列クラス・集合クラス・サイクルク ラス

ここからラベル付きクラスが与えられたときにそこから新しいラベル付きクラスを作る操作

- SEQUENCE
- SET
- CYCLE

を定義する.

## 列クラス

#### 定義 (列クラス)

ラベル付きクラス $\mathcal{B}$ に対して、

SEQUENCE<sub>k</sub>(
$$\mathcal{B}$$
) =  $\underbrace{\mathcal{B} \star \mathcal{B} \star \cdots \star \mathcal{B}}_{k \text{ (fill)}}$ 

と定義し,

$$\mathsf{Sequence}(\mathcal{B}) = \bigcup_{k \geq 0} \mathsf{Sequence}_k(\mathcal{B})$$

と定義する.

#### 列クラス

ラベル付きクラスの積が指数型母関数の積に対応した ので

$$\mathcal{A} = \text{Sequence}_k(\mathcal{B}) \Rightarrow A(z) = B(z)^k$$

$$\mathcal{A} = \text{Sequence}(\mathcal{B}) \Rightarrow A(z) = \sum_{k>0} B(z)^k = \frac{1}{1 - B(z)}$$

となる. ここにA,Bは $\mathcal{A},\mathcal{B}$ の指数型母関数.

## 集合クラス

#### 定義(集合クラス)

ラベル付きクラス $\beta$ に対して、

$$Set_k(\mathcal{B}) = Sequence_k(\mathcal{B})/\sim$$

と定義する. ここに~は並び替えて等しくなるという 同値関係. また

$$\mathsf{Set}(\mathcal{B}) = \bigcup_{k \geq 0} \mathsf{Set}_k(\mathcal{B})$$

と定義する.

#### 集合クラス

サイズkの Sequence( $\mathcal{B}$ ) のオブジェクトに対し並び替えて等しくなるものはk! 個あるのだから

$$\mathcal{A} = \operatorname{Set}_{k}(\mathcal{B}) \Rightarrow A(z) = \frac{1}{k!}B(z)^{k}$$

$$\mathcal{A} = \operatorname{Set}(\mathcal{B}) \Rightarrow A(z) = \sum_{k>0} \frac{1}{k!}B(z)^{k} = \exp B(z)$$

となる.ここにA,Bは $\mathcal{A},\mathcal{B}$ の指数型母関数.

#### サイクルクラス

#### 定義 (サイクルクラス)

ラベル付きクラス $\beta$ に対して、

$$Cycle_k(\mathcal{B}) = Sequence_k(\mathcal{B})/\sim$$

と定義する. ここに~は回転させるような並び替えで 等しくなるという同値関係. また

$$\mathsf{CYCLE}(\mathcal{B}) = \bigcup_{k \geq 1} \mathsf{CYCLE}_k(\mathcal{B})$$

と定義する.

#### サイクルクラス

$$\mathcal{A} = \text{Cycle}_k(\mathcal{B}) \Rightarrow A(z) = \frac{1}{k}B(z)^k$$

$$\mathcal{A} = \text{Cycle}(\mathcal{B}) \Rightarrow A(z) = \sum_{k \ge 1} \frac{1}{k}B(z)^k = \log \frac{1}{1 - B(z)}$$

となる.ここにA,Bは $\mathcal{A},\mathcal{B}$ の指数型母関数.

#### PERM, URN, CYC との関係

PERM, URN, CYC との関係は次の通り:

Perm = Sequence(Unit)

URN = SET(UNIT)

Cyc = Cycle(Unit)

#### 置換のクラス

置換全体のクラス PERM を思い出そう.その指数型母 関数は

$$P(z) = \sum_{n>0} n! \frac{z^n}{n!} = \sum_{n>0} z^n = \frac{1}{1-z}$$

であった. ところで置換というのは互いに重なりのない巡回置換の積として表せることを思い出すと同型

$$Perm \simeq Set(Cycle(Unit))$$

が成り立つ. 右辺の指数型母関数を計算してみると

$$\exp\left(\log\frac{1}{1-z}\right) = \frac{1}{1-z}$$

となってP(z)に一致する. 矛盾しない結果が得られた.

#### 木

#### 定義(木)

連結でサイクルを持たないグラフを木という.特定の頂点が指定された木を根付き木といい、その頂点を根という.以降、単に木といえば根付き木を指す.よくラベル付けられた木全体のなすラベル付きクラスをTREEと書く.

根を二重丸で表している。 (コンピュータサイエンスの一般的な流儀と違って 根を一番下に描いている)

#### 木

サイズnの TREE の元の個数は $n^{n-1}$  である (Cayley の定理). よって TREE の指数型母関数は

$$T(z) = \sum_{n \ge 1} \frac{n^{n-1}}{n!} z^n$$

である.

#### 木

木というのは根と部分木の集合の組だとみなせるので 同型:

Tree  $\simeq$  Unit  $\star$  Set(Tree)

を得る. よって木の指数型母関数は

$$T(z) = z \exp T(z)$$

を満たす.

なお,この関数方程式を解いて T(z) を初等関数で表すことはできないことが 知られている

#### functional グラフ

functional グラフのなすラベル付きクラスを Func と書く. functional グラフは各頂点に木がくっついたサイクルを連結成分とするグラフだったので,

が成立. したがって Func, Component の指数型母関数  $E_{I,K}$  とすれば,

$$F(z) = \exp K(z), \ K(z) = \log \frac{1}{1 - T(z)}$$

を得る. したがって,

$$F(z) = \frac{1}{1 - T(z)}.$$

#### この節のまとめ

ラベル付きクラスからラベル付きクラスを作る4つの 演算を紹介した. それらと指数型母関数の対応は次の 通り:

構成		指数型母関数
積	$\mathcal{A} = \mathcal{B} \star C$	A(z) = B(z)C(z)
列	$\mathcal{A} = Sequence(\mathcal{B})$	$A(z) = \frac{1}{1 - B(z)}$
集合	$\mathcal{A} = Set(\mathcal{B})$	$A(z) = \exp B(z)$
サイクル	$\mathcal{A} = Cycle(\mathcal{B})$	$A(z) = \log \frac{1}{1 - B(z)}$

これらを利用して functional グラフの指数型母関数を求めた.

- イントロダクション
- ② ラベル付きクラスと指数型母関数
- 3 特異点解析と木関数の漸近展開
- 多変数の指数型母関数
- ⑤ 表の結果の証明 (一番短い周期の期待値を除く)
- 一番短い周期の期待値
- **⑦** まとめ

#### この節でやること

この節「特異点解析と木関数の漸近展開」では、母関数を解析することでもとの数列の漸近挙動を知ることができることを見る.

### 特異点解析

#### 定理 (特異点解析; Flajolet & Odlyzko)

*f(z)* を領域

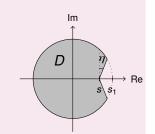
$$D = \{ z \in \mathbb{C} \mid |z| < s_1, |\arg(z - s)| > \frac{\pi}{2} - \eta, z \neq s \}$$

上で解析的な関数とする.ここに $s, s_1, \eta$  は正の実数で $s_1 > s$  を満たす. $\sigma(u) = u^{\alpha} \log^{\beta} u$  と  $\alpha \notin \{0, -1, -2, ...\}$  と  $c \in \mathbb{C}$  について,

$$f(z) \sim c\sigma\left(\frac{1}{1-z/s}\right) \text{ (as } z \to s \text{ in } D)$$

と仮定する. このとき

$$[z^n]f(z) \sim cs^{-n}\frac{\sigma(n)}{n\Gamma(\alpha)} \text{ (as } n \to \infty).$$



#### 木関数の漸近展開

#### 命題

TREE の指数型母関数 T(z) は  $D = \mathbb{C} - [e^{-1}, \infty)$  上に解析接続でき,D内で $z \rightarrow e^{-1}$  としたとき

$$T(z) = 1 - \sqrt{2}\sqrt{1 - ez} - \frac{1}{3}(1 - ez) + O((1 - ez)^{3/2})$$

- イントロダクション
- ② ラベル付きクラスと指数型母関数
- ③ 特異点解析と木関数の漸近展開
- 4 多変数の指数型母関数
- ⑤ 表の結果の証明 (一番短い周期の期待値を除く)
- 一番短い周期の期待値
- **⑦** まとめ

# 多変数の指数型母関数

この節では、一般にラベル付きクラスのあるパラメータの期待値を求めるために役立つ二変数指数型母関数を扱う.

### 二変数指数型母関数

#### 定義 (二変数指数型母関数)

 $\mathcal{A}$  をラベル付きクラス, $\chi:\mathcal{A}\to\mathbb{N}$  をパラメータとする.  $\mathcal{A}$  のパラメータ  $\chi$  に関する二変数指数型母関数とは

$$A(z,u) = \sum_{n,k \in \mathbb{N}} A_{n,k} \frac{z^n}{n!} u^k$$

のことである. ここに

$$A_{n,k} = \#\{\alpha \in \mathcal{A} \mid |\alpha| = n, \chi(\alpha) = k\}.$$

注. A(z,1) は  $\mathcal{A}$  の指数型母関数に一致するので  $\#\mathcal{A}_n = n![z^n]A(z,1).$ 

#### パラメータの期待値

各  $\mathcal{A}_n$  に一様分布を入れる. パラメータ  $\chi:\mathcal{A}\to\mathbb{N}$  を 各  $\mathcal{A}_n$  の確率変数と見る. このとき

#### 補題

 $\chi$ の期待値は

$$\mathbb{E}_{\mathcal{A}_n}(\chi) = \frac{[z^n]\partial_u A(z,u)|_{u=1}}{[z^n]A(z,1)}.$$

証明.分母は $\#\mathcal{A}_n$  をn! で割ったものである.分子を見る.

# パラメータの期待値

#### 補題

$$\mathbb{E}_{\mathcal{A}_n}(\chi) = \frac{[z^n]\partial_u A(z,u)|_{u=1}}{[z^n]A(z,1)}.$$

$$A(z, u) = \sum_{\alpha \in \mathcal{A}} u^{\chi(\alpha)} \frac{z^{|\alpha|}}{|\alpha|!}$$

なので

$$\left. \frac{\partial}{\partial u} A(z, u) \right|_{u=1} = \sum_{\alpha \in \mathcal{A}} \chi(\alpha) \frac{z^{|\alpha|}}{|\alpha|!} = \sum_{n \geq 0} \left( \sum_{|\alpha| = n} \chi(\alpha) \right) \frac{z^n}{n!}$$

よりよい.



### ラベル付きクラスの演算と二変数指数 型母関数

通常の指数型母関数と同様の対応がラベル付きクラスと二変数指数型母関数の間にもつく:

構成		二変数指数型母関数
積	$\mathcal{A} = \mathcal{B} \star C$	A(z,u) = B(z,u)C(z,u)
列	$\mathcal{A} = Sequence(\mathcal{B})$	$A(z,u) = \frac{1}{1-B(z,u)}$
集合	$\mathcal{A} = Set(\mathcal{B})$	$A(z, u) = \exp B(z, u)$
サイクル	$\mathcal{A} = Cycle(\mathcal{B})$	$A(z, u) = \log \frac{1}{1 - B(z, u)}$

- イントロダクション
- ② ラベル付きクラスと指数型母関数
- ③ 特異点解析と木関数の漸近展開
- 4 多変数の指数型母関数
- ⑤表の結果の証明 (一番短い周期の期待値を除く)
- 一番短い周期の期待値
- **の**まとめ

### 表の結果の証明

以上の準備のもとで,以下の表の左側の結果を証明していく.

量	Map( <u>n</u> , <u>n</u> )	$\mathfrak{S}_n$
長さ 1 のサイクルを持つ確率	$\rightarrow 1 - 1/e \doteq 0.63$	
連結成分の個数の期待値	$\sim \frac{1}{2} \log n$	$=H_n$
ランダムな点をとったときそれが属する 連結成分が持つサイクルの長さの期待値	$\sim \sqrt{\pi n/8} \doteq 0.62\sqrt{n}$	=(n+1)/2
一番長いサイクルの長さの期待値	$\sim c_1 \sqrt{n} \doteq 0.78 \sqrt{n}$	$\sim \lambda n \doteq 0.62n$
一番短いサイクルの長さの期待値	$\sim \frac{e^{-\gamma}}{2} \log n \doteq 0.28 \log n$	$\sim e^{-\gamma} \log n \doteq 0.56 \log n$

連結成分の個数をパラメータとする Func の二変数指数型母関数  $\xi(z,u)$  は

$$\xi(z, u) = \exp(u \log \frac{1}{1 - T(z)})$$

となる.これについて  $\Xi = \frac{\partial}{\partial u} \xi(z,u) \Big|_{u=1}$  を計算すると

$$\Xi(z) = \frac{1}{1 - T(z)} \log \frac{1}{1 - T(z)}$$

木関数の近似

$$T(z) = 1 - \sqrt{2}\sqrt{1 - ez} + O(1 - ez) \quad (z \to 1/e)$$
 を使うと  
 $\Xi(z) \sim \frac{1}{\sqrt{2}\sqrt{1 - ez}} \log \frac{1}{\sqrt{2}\sqrt{1 - ez}}$   
 $\sim \frac{1}{2\sqrt{2}} \frac{1}{1 - ez} \log \frac{1}{1 - ez} \quad (z \to 1/e).$ 

特異点解析定理を使うと

$$[z^n] \Xi(z) \sim \frac{1}{2\sqrt{2}} e^n \frac{n^{1/2} \log n}{n\sqrt{\pi}}$$
$$= \frac{1}{2\sqrt{2\pi n}} e^n \log n$$

よって連結成分の個数の期待値は

$$\frac{n![z^n]\Xi(z)}{n^n} \sim \frac{n!}{n^n} \frac{1}{2\sqrt{2\pi n}} e^n \log n$$
$$\sim \frac{1}{2} \log n \quad (\because \text{ Stirling } の公式)$$

となる.

量	Map( <u>n</u> , <u>n</u> )	$\mathfrak{S}_n$
長さ 1 のサイクルを持つ確率	$\rightarrow 1 - 1/e \doteq 0.63$	
連結成分の個数の期待値	~ ½ log n ✔	$=H_n$
ランダムな点をとったときそれが属する 連結成分が持つサイクルの長さの期待値	$\sim \sqrt{\pi n/8} \doteq 0.62\sqrt{n}$	=(n+1)/2
一番長いサイクルの長さの期待値	$\sim c_1 \sqrt{n} \doteq 0.78 \sqrt{n}$	$\sim \lambda n \doteq 0.62n$
一番短いサイクルの長さの期待値	$\sim \frac{e^{-\gamma}}{2} \log n \doteq 0.28 \log n$	$\sim e^{-\gamma} \log n \doteq 0.56 \log n$

次にランダムな点をとったときそれが属する連結成分が持つサイクルの長さの期待値を考察する.

# 

 $\varphi \in \mathsf{FUNC}_n, i \in n$  に対して

NCycPts $(\varphi, i) = (\varphi$ においてi が属する 連結成分のサイクルの長さ)

とおく. この量に関する指数型母関数を

$$\xi(z) = \sum_{n \ge 0} \Big( \sum_{\substack{\varphi \in \mathsf{FUNC}_n \\ i \in n}} \mathsf{NCycPts}(\varphi, i) \Big) \frac{z^n}{n!}$$

とする.

すると欲しい量であるランダムな点からのサイクルの 長さの期待値は

$$E = \frac{n![z^n]\xi(z)}{n^{n+1}}$$

で求まる.  $\xi$ を求めよう.

まず、サイクルに属する要素の個数をパラメータとする連結成分の指数型母関数は

$$f(z, u) = \log \frac{1}{1 - uT(z)}$$

で与えられる.次の関数を考えよう:

$$g(z) = z \frac{\partial^2}{\partial z \partial u} f(z, u) \bigg|_{u=1}.$$

#### 期待值

g(z) は重みのついた連結成分の指数型母関数であって、サイズnの連結成分でk個のサイクル要素を持つものは重さ $n \cdot k$ を持つものになる。

$$f(z,u) = \sum_{n,k} a_{n,k} \frac{z^n}{n!} u^k$$
 とすると

$$\frac{\partial}{\partial u}f(z,u) = \sum_{n,k} a_{n,k} \frac{z^n}{n!} k u^{k-1}$$

$$\therefore \left. \frac{\partial}{\partial u} f(z, u) \right|_{u=1} = \sum_{n,k} (a_{n,k} \cdot k) \frac{z^n}{n!}$$

$$\therefore z \frac{\partial}{\partial z} \frac{\partial}{\partial u} f(z, u) \bigg|_{u=1} = \sum_{n,k} (a_{n,k} \cdot n \cdot k) \frac{z^n}{n!}$$

そこで 
$$g(z) = z \frac{\partial^2}{\partial z \partial u} f(z, u) \Big|_{u=1}$$
 を計算すると

$$g(z) = \frac{zT'(z)}{(1-T(z))^2}.$$

g(z) と Func の指数型母関数  $\frac{1}{1-T(z)}$  を掛けたものが求める  $\xi(z)$  になる:

$$\xi(z) = \frac{zT'(z)}{(1 - T(z))^3}.$$

#### 期待值

g(z) と Func の指数型母関数  $\frac{1}{1-T(z)}$  を掛けたものが求める  $\mathcal{E}(z)$  になる.

#### 期待值

 $CO(z) = (zT'(z))/(1 - T(z))^3$  の特異点 z = 1/e における展開は

$$\xi(z) \sim \frac{1}{4(1-ez)^2}$$

になる. これに特異点解析定理を使うと

$$[z^n]\xi(z)\sim \frac{1}{4}e^n n.$$

$$\therefore E = \frac{n![z^n]\xi(z)}{n^{n+1}} \sim \frac{n!}{n^{n+1}} \frac{1}{4} e^n n \sim \sqrt{\frac{\pi n}{8}}.$$

#### 期待值

量	Map( <u>n</u> , <u>n</u> )	$\mathfrak{S}_n$	
長さ 1 のサイクルを持つ確率	$\rightarrow 1 - 1/e \doteq 0.63$		
連結成分の個数の期待値	~ ½ log n ✔	$=H_n$	
ランダムな点をとったときそれが属する 連結成分が持つサイクルの長さの期待値	$\sim \sqrt{\pi n/8} \doteq 0.62\sqrt{n} \checkmark$	=(n+1)/2	
一番長いサイクルの長さの期待値	$\sim c_1 \sqrt{n} \doteq 0.78 \sqrt{n}$	$\sim \lambda n \doteq 0.62n$	
一番短いサイクルの長さの期待値	$\sim \frac{e^{-\gamma}}{2} \log n \doteq 0.28 \log n$	$\sim e^{-\gamma} \log n \doteq 0.56 \log n$	

$$\xi^{\max}(\varphi) = \max_{i \in \underline{n}} \xi(\varphi, i)$$

と母関数

$$\Xi(z) = \sum_{n \ge 0} \xi_n \frac{z^n}{n!}$$

where  $\xi_n = \sum_{\varphi \in \mathsf{FUNC}_n} \xi^{\mathsf{max}}(\varphi)$  を考えよう.

このとき

$$F^{[k]}(z) := \sum_{\varphi \in \mathsf{Func}_n^{[k]}} \frac{z^{|\varphi|}}{|\varphi|!}$$

where  $\operatorname{Func}_n^{[k]} = \{ \varphi \in \operatorname{Func}_n \mid \xi^{\max}(\varphi) \leq k \}$  とおくと

$$\Xi(z) = \sum_{k>0} (F(z) - F^{[k]}(z))$$

となる.

これは離散確率変数 X に関する次の式の母関数バージョンである:

$$\mathbf{E}[X] = \sum_{k>1} k \mathbf{P}(X = k) = \sum_{k>1} \mathbf{P}(X \ge k) = \sum_{k>0} (1 - \mathbf{P}(X \le k))$$

さて, $\xi(\varphi,i)$ が $\varphi$ においてiが属する成分のサイクルの長さだったとしよう.このとき

$$F^{[k]}(z) = \exp(I_k(t(z)))$$

where  $I_k(z) = \sum_{j=1}^k \frac{u^j}{j}$  rand.

$$\Xi(z) = \sum_{k>0} \left( \frac{1}{1 - T(z)} - F^{[k]}(z) \right) = \frac{1}{1 - T(z)} \sum_{k>1} (1 - e^{-r_k(t(z))})$$

where 
$$r_k(u) = \log \frac{1}{1-u} - I_{k-1}(u) = \sum_{j \ge k} \frac{u^j}{j}$$
.

T(z) は z = 1/e の周りで |T(z)| < 1 を満たす.よって, $T(z) = e^{-x}$  とおけば Re x > 0 のときに  $x \to 0$  の挙動を調べればよい.和を積分で近似して次を得る:

$$\Xi(z) = \frac{1}{1 - T(z)} \sum_{k \ge 1} \left( 1 - \exp(-\sum_{j \ge k} \frac{e^{-jx}}{j}) \right) 
\sim \frac{1}{1 - T(z)} \sum_{k \ge 1} \left( 1 - \exp(-\int_{kx}^{\infty} \frac{e^{-v}}{v} dv) \right) 
\sim \frac{1}{(1 - T(z))x} \int_{0}^{x\infty} \left( 1 - \exp(-\int_{u}^{\infty} \frac{e^{-v}}{v} dv) \right) du 
\sim \frac{1}{(1 - T(z))^{2}} \int_{0}^{\infty} \left( 1 - \exp(-\int_{u}^{\infty} \frac{e^{-v}}{v} dv) \right) du$$

### 一番長いサイクルの長さの期待値

特異点解析定理を  $\frac{1}{(1-T(z))^2}$  に適用すると

$$[z^n] \frac{1}{(1-T(z))^2} \sim \frac{1}{2}e^n$$

なので,一番長いサイクルの長さの期待値 E は

$$E = \frac{n!}{n^n} [z^n] \Xi(z) \sim \sqrt{\frac{\pi n}{2}} \int_0^\infty \left( 1 - \exp(-\int_u^\infty \frac{e^{-v}}{v} dv) \right) du$$

ここに√πの比例係数は

$$c_1 = \sqrt{\frac{\pi}{2}} \int_0^\infty \left( 1 - \exp(-\int_u^\infty \frac{e^{-v}}{v} dv) \right) du = 0.782481600991657 \dots$$

# 一番長いサイクルの長さの期待値

量	Map( <u>n</u> , <u>n</u> )	$\mathfrak{S}_n$
長さ 1 のサイクルを持つ確率	$\rightarrow 1 - 1/e \doteq 0.63$	
連結成分の個数の期待値	~ ½ log n ✔	= <i>H</i> <sub>n</sub>
ランダムな点をとったときそれが属する 連結成分が持つサイクルの長さの期待値	$\sim \sqrt{\pi n/8} \doteq 0.62\sqrt{n} \checkmark$	=(n+1)/2
一番長いサイクルの長さの期待値	$\sim c_1 \sqrt{n} \doteq 0.78 \sqrt{n} \checkmark$	$\sim \lambda n \doteq 0.62n$
一番短いサイクルの長さの期待値	$\sim \frac{e^{-\gamma}}{2} \log n \doteq 0.28 \log n$	$\sim e^{-\gamma} \log n \doteq 0.56 \log n$

- イントロダクション
- ② ラベル付きクラスと指数型母関数
- ③ 特異点解析と木関数の漸近展開
- 4 多変数の指数型母関数
- ⑤ 表の結果の証明 (一番短い周期の期待値を除く)
- 6 一番短い周期の期待値
- **⑦** まとめ

 $a_n$  を  $\varphi$  が Map( $\underline{n},\underline{n}$ ) 全体を走るときの  $\varphi$  の最小の周期の和とする.  $a_n$  の指数型母関数  $f(z) = \sum_n a_n z^n/n!$  は次で与えられる:

$$f(z) = \sum_{k \ge 1} \left( -1 + \exp\left(\sum_{j \ge k} \frac{t(z)^j}{j}\right) \right).$$

なぜならば

$$Set(Cycle_{\geq k}(Tree)) - \{\emptyset\} \Leftrightarrow exp\left(\sum_{j>k} \frac{T(z)^j}{j}\right) - 1.$$

ここで

$$A(w) = \sum_{k \ge 1} \left( -1 + \exp\left(\sum_{j \ge k} \frac{w^j}{j}\right) \right)$$

とおくと母関数 f(z) は

$$f(z) = A(T(z)).$$

一番長い周期のときと同様  $T(z) = e^{-x}$  とおけば Re(x) > 0 で  $A(e^{-x})$  の  $x \to 0$  の挙動を調べればよい.

関数 f : (0,∞) → ℂのメリン変換は次で定義される:

$$\mathcal{M}[f](s) = \int_0^\infty x^{s-1} f(x) dx.$$

$$A(e^{-x}) = \sum_{k \ge 1} \left( -1 + \exp \sum_{j \ge k} \frac{e^{-jx}}{j} \right)$$

の挙動を調べる. まず内側の和を積分で近似すると

$$A(e^{-x}) \sim \sum_{k \ge 1} \left(-1 + \exp \int_{kx}^{\infty} \frac{e^{-u}}{u} du\right)$$

この右辺を $\tilde{A}(x)$ とおく.また, $E_1(x)=\int_x^\infty \frac{e^{-u}}{u}du$ とおく.したがって

$$\tilde{A}(x) = \sum_{k>1} (-1 + \exp E_1(kx)).$$

計算をがちゃがちゃやれば  $M ilde{A}(s)$  は

$$M\tilde{A}(s) = \frac{e^{-\gamma}}{s^2(s-1)}\zeta(s)h(s)$$

と有理型接続されることが分かる. ここに h(s) は Re(s) > 0 で正則な関数で、h(1) = 1.

s=1 での  $1/s^2$ , h(s) のテイラー展開と  $\zeta(s)$  のローラン 展開を使うことで

$$M[\tilde{A}](s) \simeq \frac{e^{-\gamma}}{(s-1)^2} + \frac{e^{-\gamma}(-2+\gamma) + h'(1)}{s-1}$$

を得る. ここに × は差が正則関数の意味.

#### 定義 (fundamental strip)

実数a < bについて $\langle a, b \rangle = \{s \in \mathbb{C} \mid a < \operatorname{Re}(s) < b\}$ とおく. メリン変換 $\mathcal{M}f(s)$ の値が $\langle a, b \rangle$ 内で存在するような最大の $\langle a, b \rangle$ をこのメリン変換のfundamental stripという.

#### 定理 (メリン変換による漸近挙動の対応)

f(x) を  $(0,\infty)$  で連続な関数とし,そのメリン変換 Mf(s) は fundamental strip  $\langle \alpha,\beta \rangle$  を持つとする.ある  $\alpha' < \alpha$  について,Mf(s) が  $\langle \alpha',\beta \rangle$  に有理型接続できると仮定し,そこで極は有限個で, $\mathrm{Re}(s) = \alpha'$  で解析的とする.また, $\eta \in (\alpha,\beta)$  と r > 1 が存在して, $\alpha' \leq \mathrm{Re}(s) \leq \eta$  で  $|s| \to \infty$  としたとき  $Mf(s) = O(|s|^{-r})$  となることを仮定する.このときもし, $s \in \langle \alpha',\alpha \rangle$  について

$$\mathcal{M}f(s) \times \sum_{(\xi,k)\in A} d_{\xi,k} \frac{1}{(s-\xi)^k}$$

ならば $x \rightarrow +0$  で

$$f(x) = \sum_{(\xi,k) \in A} d_{\xi,k} \left( \frac{(-1)^{k-1}}{(k-1)!} x^{-\xi} (\log x)^{k-1} \right) + O(x^{-\alpha'}).$$

定理を $f = \tilde{A}, \alpha = 1, \beta = \infty, \alpha' = 0.5, \eta = 2$ で適用すると

$$\tilde{A}(x) = -e^{-\gamma}x^{-1}\log x + (e^{-\gamma}(-2+\gamma) + h'(1))x^{-1} + O(x^{-0.5})$$

を得る. したがって,

$$\tilde{A}(x) \sim e^{-\gamma} \frac{1}{x} \log \frac{1}{x} \quad (x \to 0, x > 0)$$

を得る. よって,

$$A(e^{-x}) \sim e^{-\gamma} \frac{1}{x} \log \frac{1}{x} \quad (x \to 0, x > 0).$$

頑張ればx を実数の範囲でなくて Re(x) > 0 の範囲で動かしたときも同様の結果:

$$A(e^{-x}) \sim e^{-\gamma} \frac{1}{x} \log \frac{1}{x} \quad (x \to 0, \operatorname{Re}(x) > 0).$$

が得られる.

そこでこれに特異点解析を使えば一番短い周期の期待値Eとして

$$E \sim \frac{e^{-\gamma}}{2} \log n$$

を得ることができる.

量	Map( <u>n</u> , <u>n</u> )	⊗ <sub>n</sub>
長さ 1 のサイクルを持つ確率	→ 1 − 1/e = 0.63	
連結成分の個数の期待値	~ ½ log n ✔	= H <sub>n</sub>
ランダムな点をとったときそれが属する 連結成分が持つサイクルの長さの期待値	$\sim \sqrt{\pi n/8} \doteq 0.62\sqrt{n} \checkmark$	=(n+1)/2
一番長いサイクルの長さの期待値	$\sim c_1 \sqrt{n} \doteq 0.78 \sqrt{n}  \checkmark$	~ <i>\lambda n</i> ≐ 0.62 <i>n</i>
一番短いサイクルの長さの期待値	$\sim \frac{e^{-\gamma}}{2} \log n \doteq 0.28 \log n \checkmark$	$\sim e^{-\gamma}\log n \doteq 0.56\log n$

- イントロダクション
- 2 ラベル付きクラスと指数型母関数
- ③ 特異点解析と木関数の漸近展開
- 4 多変数の指数型母関数
- ⑤ 表の結果の証明 (一番短い周期の期待値を除く)
- 一番短い周期の期待値
- **⑦** まとめ

### まとめ

- ・以下の流れでランダム写像に関するいろんな量の 漸近挙動を調べた:
  - 漸近挙動を得たい量に対応する母関数を求める
  - 2 その母関数の特異点周りの振る舞いを調べる
  - 3 特異点解析定理を使い、もとの量の漸近挙動がわかる
- その結果次の結果を得た

量	Map( <u>n</u> , <u>n</u> )
連結成分の個数の期待値	$\sim \frac{1}{2} \log n$
ランダムな点をとったときそれが属する 連結成分が持つサイクルの長さの期待値	$\sim \sqrt{\pi n/8} \doteq 0.62\sqrt{n}$
一番長いサイクルの長さの期待値	$\sim c_1 \sqrt{n} \doteq 0.78 \sqrt{n}$
一番短いサイクルの長さの期待値	$\sim \frac{e^{-\gamma}}{2} \log n \doteq 0.28 \log n$

### 参考文献

- Philippe Flajolet and Andrew M. Odlyzko. Random mapping statistics. Research Report RR-1114.
   INRIA, 1989.
- [2] L. A. Shepp and S. P. Lloyd. "Ordered Cycle Lengths in a Random Permutation". In: Transactions of the American Mathematical Society 121.2 (1966), pp. 340–357.
- [3] Paul W Purdom and JH Williams. "Cycle length in a random function". In: Transactions of the American Mathematical Society 133.2 (1968), pp. 547–551.
- [4] Philippe Flajolet and Andrew Odlyzko. "Singularity analysis of generating functions". In: SIAM Journal on discrete mathematics 3.2 (1990), pp. 216–240.
- [5] I.P. Goulden and D.M. Jackson. Combinatorial Enumeration. Dover Books on Mathematics. Dover Publications, 2004.
- [6] P. Flajolet and R. Sedgewick. Analytic Combinatorics. Cambridge University Press, 2009.
- [7] D.E. Knuth. The Art of Computer Programming: Fundamental algorithms. Addison-Wesley series in computer science and information processing. Addison-Wesley, 1997.
- [8] D.E. Knuth. The Art of Computer Programming: Seminumerical algorithms. Addison-Wesley series in computer science and information processing. Addison-Wesley Publishing Company, 1981.
- [9] A208231 OEIS. http://oeis.org/A208231.
- [10] Gilbert Labelle François Bergeron and Pierre Leroux. Introduction to the Theory of Species of Structures. http://bergeron.math.uqam.ca/wp-content/uploads/2013/11/book.pdf.

# 参考文献

- [1] には表の左側の結果のうち一番短い周期以外が載っている. (非常に参考にした)
- [2] はランダムな置換に対する一番短い周期の結果を証明している.
- [3] は [2] の結果を使ってランダムな写像に対する一番短い周期の結果を示している.
- [4] は特異点解析を証明した論文.
- [6] は解析的組み合わせ論の教科書、組み合わせの数の漸近 挙動を母関数の解析によって得るといった手法を学ぶこと ができる. (非常に参考にした)
- [9] は OEIS(Online Encyclopedia of Integer Sequences) のランダム写像の一番短い周期に対する項目.